Robust Localization with Wearable Sensors

Xinyu Hou
St Anne's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2023

Acknowledgements

I would like to express my deep and sincere gratitude to my supervisor Professor Jeroen Bergmann, for his continuous support, warmest encouragement, valuable trust, immense knowledge, patience and kindness. He is the one who gave me the opportunity to stand on the giants' shoulders and take a look at the beauty of science. His expert guidance led me through the transition from a student to an independent researcher. Without his insightful guidance and unwavering support through all the difficult times in my Dphil study, these works would never be possible. I learnt from him how to be a dedicated scientist, a diligent researcher, a passionate entrepreneur, and a supportive supervisor. He is a role model that I will deeply admire and respect throughout my life. No amount of thanks could truly express how grateful I am.

Second, I would like to extend my heartfelt thanks to my colleagues in Natural Interaction Lab who have provided me with invaluable support and encouragement: (in alphabetical order) Leonardo de Almeida e Bueno, Runbei Cheng, Dr. Rita Hendricusdottir, Dr Emily (Man Ting) Kwong, Dr. Vikranth Harthikote Nagaraja, Joseph Russell, Victoria Walls. I feel so lucky to have worked with such a wonderful and warm team.

I am also particularly grateful for all professors who provided insightful and constructive suggestions, feedback, to shape my research and refine my arguments: (in alphabetical order) Prof. David Clifton, Prof. Aldo Faisal, Dr. Jonathan Gammell, Prof. Konstantinos Kamnitsas, Prof. Paul Newman, Prof. Amy Zavatsky, and all anonymous reviewers who

gave comments to improve my papers.

I am forever grateful to my family, friends, and my cat for their unconditional love, support, encouragement, and belief in me. They are always my source of energy, motivation, and happiness.

Thank you the miraculous universe. Thank you all past and current scientists. Thank you University of Oxford, Department of Engineering Science, and St Anne's College.

Abstract

Measuring physical movements of humans and understanding human behaviour is useful in a variety of areas and disciplines. Human inertial tracking is a method that can be leveraged for monitoring complex actions that emerge from interactions between human actors and their environment. An accurate estimation of motion trajectories can support new approaches to pedestrian navigation, emergency rescue, athlete management, and medicine. However, tracking with wearable inertial sensors has several problems that need to be overcome, such as the low accuracy of consumer-grade inertial measurement units (IMUs), the error accumulation problem in long-term tracking, and the artefacts generated by movements that are less common. This thesis focusses on measuring human movements with wearable head-mounted sensors to accurately estimate the physical location of a person over time. The research consisted of (i) providing an overview of the current state of research for inertial tracking with wearable sensors, (ii) investigating the performance of new tracking algorithms that combine sensor fusion and data-driven machine learning, (iii) eliminating the effect of random head motion during tracking, (iv) creating robust long-term tracking systems with a Bayesian neural network and sequential Monte Carlo method, and (v) verifying that the system can be applied with changing modes of behaviour, defined as natural transitions from walking to running and vice versa. This research introduces a new system for inertial tracking with head-mounted sensors (which can be placed in, e.g. helmets, caps, or glasses). This technology can be used for long-term positional tracking to explore complex behaviours.

Acronyms

AKF Adaptive Kalman Filter.

ATE absolute trajectory error.

Bi-RNNs Bidirectional Recurrent Neural Networks.

BLE Bluetooth low energy.

CKF Cascaded Kalman Filter.

CNN Convolutional Neural Network.

CWT/DWT continuous/discrete wavelet transforms.

DenseNet Dense Convolutional Network.

DL Deep Learning.

DNN deep neural network.

EMG Electromyography.

EVA Ethylene-vinyl acetate.

FFT Fast Fourier Transformation.

GNSS Global navigation satellite system.

GPS Global Positioning System.

GRU Gated Recurrent Units.

HDR Heuristic Drift Reduction.

HMMs Hidden Markov Models.

IMU Inertial Measurement Unit.

INS Inertial Navigation System.

LBS Location Based Services.

LPF low pass filter.

LSTM Long Short-Term Memory.

MEMS Micro-Electro-Mechanical Systems.

MLP Muti-layer perceptrons.

MTS motion tracking system.

NMT neural machine translation.

PDR Pedestrian Dead Reckoning.

PNN Probabilistic neural network classifiers.

PRISMA Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

PSDs position-sensitive detectors.

QSF Quasi-Static magnetic Field.

RAKF Robust Adaptive Kalman Filter.

RBPF Rao Blackwellised Particle filter.

RFID radio frequency identification.

RMSE Root Mean Square Error.

RNN Recurrent Neural Network.

RTE relative trajectory error.

SHSs Step-and-Heading Systems.

SINS Strapdown Inertial Navigation System.

STFT Short-Time Fourier Transform.

SVM support vector machine.

TCN Temporal Convolutional Network.

UAV unmanned aerial vehicle.

UWB ultra wideband.

VIO visual inertial odometry.

VLC Visible Light Communication.

VR virtual reality.

WLAN wireless local area network.

ZARU zero-angular rate update.

ZUPT Zero velocity update.

ZVU zero velocity updates.

Contents

List of Figures

Chapter 1

Introduction

1.1 Human positioning

Finding the relative position of objects is an eternal question in our three-dimensional universe. How far is the boundary of the universe from us? What are the distances and relative motions between celestial bodies? Docking between spaceships and space stations requires highly accurate relative position estimations. As for us humans, positioning is a much more common problem and used in a variety of fields across our daily life. People use it to locate and navigate to unfamiliar destinations without getting lost. Athletes use it to track their movements for recording workload, estimating fatigue level, or optimising their competitive strategies. Rescue teams use it to search for survivors and ensure that their own safety is maintained in extreme environments, such as burning buildings or wild environments [50]. Smart homes use user position information to make better decisions on how to support their lives. Hospitals can apply it to localise patients in case of emergencies [139]. Location Based Services (LBS) is thus often being leveraged to provide specific services to users based on their location. And it has been suggested that LBS is becoming one of the most important sources of revenue for the wireless communications industry [31].

The wide variety of these applications is attributed to the rapid development of position tracking technology in recent years. Global navigation satellite system (GNSS) is the most common tracking system, which includes Global Positioning System (GPS), GLONASS, BeiDou, Galileo. GNSS allows small electronic receivers

to determine their locations with high precision using time data transmitted along a line of sight by radio signals from satellites. However, due to signal blockage or strong multipath propagation, satellite signals are often unavailable or degraded in critical environments, including indoor, underground, or urban canyons. Therefore, it is widely used in outdoor open field environments.

As for positioning when the GNSS is unavailable or when it cannot be trusted to accurately work within a given environment, various methods have been developed by researchers, which could be classified into building dependent and building independent methods. Building-dependent methods include techniques which require access to the building's infrastructures such as Wi-Fi [46], Cellular, or Bluetooth low energy (BLE) beacons [174], as well as approaches requiring dedicated infrastructures, such as radio frequency identification (RFID) [77] [182], ultra wideband (UWB) [84, 3], infrared, Radar [78], Zigbee, Visible Light Communication (VLC), ultrasound [193], and acoustic signal. Building independent methods include image-based technologies like cameras, and dead reckoning with inertial sensors.

Unfortunately, many of these methods still rely on external aiding signals, information or infrastructure, and thus are not applicable in scenarios where the signals are severely affected or when there is no specific infrastructure available. The construction of new infrastructures to facilitate indoor tracking also comes at an added cost, as it can be expensive to setup and maintain these frameworks. In addition, image-capture devices may face privacy or security threats when used in a private environment.

In that case, inertial tracking outperforms these other methods, as it is self-contained and does not rely on any external infrastructure or signals. It will not be affected by the natural environment or human disturbance and could be used in any environment such as indoor, outdoor, underground, underwater, or complex surroundings that combine different environments together. For example, it is suitable for environmental emergencies, even as the environment is changing frequently. It also works well in dim and smoky light conditions, making it ideal for firefighting or

rescue operations that could face similar challenges. It can also be adopted for the sports community since it could provide a scalable solution. Trajectories of athletes could be recorded indoors or outdoors, which coaches could then analyse to improve performance or manage the players. The trajectory length that the athletes have covered during the basketball or football game can also reflect the amount of work that the player has done, and fatigue levels could be estimated, which could be applied by the coach, as a training tool, to create a competitive advantage [106] and avoid injuries. The inertial tracking techniques relies on data from inertial sensors.

1.2 Inertial sensors

Inertial Measurement Unit (IMU) is composed of accelerometer and gyroscope. IMUs have first been adopted in planes, spacecraft, and submarines as core components of control and navigation systems since the last century. It used to be very expensive and the size of the systems was huge. However, the provided valuable information and could generate very precise and useful data. In recent years, with the development of Micro-Electro-Mechanical Systems (MEMS), IMUs tend to be significantly lower in cost, smaller and more energy efficient, thus becoming widely deployed in many products such as wearable devices, mobile phones, robots and unmanned aerial vehicle (UAV), to name just a few.

Accelerometers measure accelerations based on Newton's second law and gyroscopes measure the angular velocity. Data are normally obtained from three axes to provide a suitable 3D representation. Low-cost IMUs face the problem of limited accuracy, with error sources categorised into two types: deterministic errors and stochastic errors. The deterministic terms include scale factor and axis misalignment. Stochastic terms are composed of white noise that follows a Gaussian distribution and bias with parameters including in-run bias stability, turn-on bias stability or repeatability,

and bias over temperature. The error models of IMU measurements are:

$$\begin{aligned} \omega_m &= \omega_t + \eta_\omega + b_\omega \\ a_m &= a_t + \eta_a + b_a \end{aligned} \tag{1.1}$$

where ω_m and a_m are the measured values, ω_t and a_t are true values, η is noise with high frequency and b is bias with low frequency.

One key benefit of the IMU is that it does not need any external infrastructure or signals. There are other advantages in choosing the inertial sensor for positioning. They include the small form factor, which can make them discreet to use and thus it can be integrated in a variety of daily objects like glasses, mouthguards, earphones, etc. Of course, the low-cost nature of the IMUS makes them very popular, and this is shown by the fact that they are integrated in almost all "smart" mobile devices that are available on the market.

1.3 Strapdown inertial navigation system

In the basic Strapdown Inertial Navigation System (SINS), the inertial sensors are rigidly mounted on the measuring subject. The basic SINS algorithms is based on the physical models as follows.

1.3.1 Orientation

The gyroscope measures the angular velocity signal in the body frame.

$$\omega_b(t) = (\omega_{bx}(t), \omega_{by}(t), \omega_{bz}(t))^T$$

A vector $\mathbf{v}_b$ in the body frame is equivalent to the vector $\mathbf{v}_g$ in global frame, if $\mathbf{v}_g = \mathbf{C}\mathbf{v}_b$, where $\mathbf{C}$ is the 3×3 rotation matrix in the direction cosines representation of the body frame attitude.

The changing rate of rotation matrix is updated with the angular velocity:

$$\dot{\mathbf{C}}(t) = \mathbf{C}(t)\boldsymbol{\Omega}(t) \tag{1.2}$$

where

$$\mathbf{\Omega}(t) = \begin{pmatrix} 0 & -\omega_{bz}(t) & \omega_{by}(t) \\ \omega_{bz}(t) & 0 & -\omega_{bx}(t) \\ -\omega_{by}(t) & \omega_{bx}(t) & 0 \end{pmatrix} \tag{1.3}$$

is the skew symmetric form of the angular rate vector $\omega_b(t)$.

Thus the rotation matrix will be updated with:

$$\mathbf{C}(t) = \mathbf{C}(0) \cdot \exp\left(\int_0^t \mathbf{\Omega}(t)dt\right) \tag{1.4}$$

1.3.2 Position

The accelerometer measures the acceleration signal in the body frame.

$$\mathbf{a}_b(t) = (a_{bx}(t), a_{by}(t), a_{bz}(t))^T$$

Transfer it to the global frame:

$$\mathbf{a}_g(t) = \mathbf{C}(t)\mathbf{a}_b(t) \tag{1.5}$$

To obtain the displacement:

$$\mathbf{s}_g(t) = \mathbf{s}_g(0) + \int_0^t \left(\mathbf{v}_g(0) + \int_0^t (\mathbf{a}_g(t) - \mathbf{g}_g)\, dt\right) dt \tag{1.6}$$

where $\mathbf{v}_g$ is the velocity and $\mathbf{s}_g$ is the displacement both in the global frame.

Although there are many reasons to use inertial sensors for localisation, inertial sensors do have some drawbacks that cannot be ignored. MEMS IMUs get smaller in size and lower in cost, but this can come with a loss of precision. The orientation and displacement are calculated directly by the integration of angular velocity (Equation 1.4) and the double integration of acceleration (Equation 1.6). This works well with the traditional larger (but precise) IMUs. Yet, when adopting low-cost MEMS IMUs

with only SINS, errors could be exploding by the double integration and thus the system will collapse in a few seconds. To overcome this problem, scientists have proposed a lot of different methods for positioning with these less precise low-cost IMUs. Model-based inertial positioning systems were one of the first solutions that gained traction.

1.4 Model based inertial positioning systems

SINS faces the problem of error mushrooming. A possible solution to mitigate unbounded drifts could be based on the knowledge we have of the human walking pattern, each foot standing still for a small period during the gait cycle when people are walking. Zero velocity update (ZUPT) [47] and zero-angular rate update (ZARU) utilise this feature to eliminate drifts during this stationary phase. However, these methods are only suitable for tracking systems consisting of IMUs mounted on the foot, which is a rare placement in the current product market.

When we discard the basic integration concept and fully focus on leveraging the human walking patterns, the concept of Pedestrian Dead Reckoning can be considered. Pedestrian Dead Reckoning (PDR) is the process of calculating one's current location by using the previously known position, and advancing that position over time using established or estimated speeds and trajectories (or alternatively stride lengths and directions). PDR is based on the idea of Step-and-Heading Systems (SHSs), which divide the whole tracking problem into three specific questions consisting of the following aspects: (i) step detection, (ii) step length estimation, and (iii) orientation estimation. Figure 1.1 illustrates the phases that make up the SHSs. After estimating the step length L_t and the orientation of each step θ_t, the position x_t, y_t at the time step t could be calculated by the following.

$$\begin{aligned} x_t &= x_{t-1} + L_t \cdot \cos\theta_t \\ y_t &= y_{t-1} + L_t \cdot \sin\theta_t \end{aligned} \tag{1.7}$$

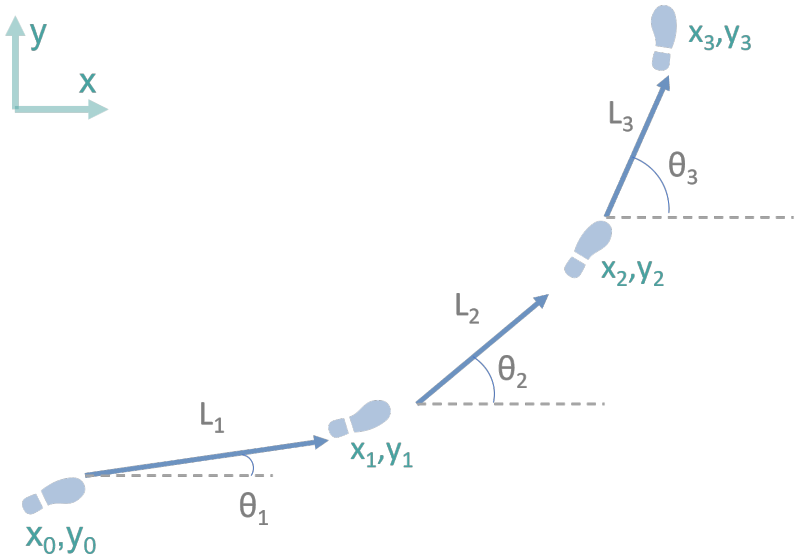


Figure 1.1: Step-and-Heading Systems (SHSs) with x and y representing the 2D global reference frame. L is the step length and θ the orientation of each step. A total of three steps is shown.

1.4.1 Step detection

There are a range of different approaches that can be used to detect whether a step has been taken. Step detection methods include, for example, peak detection, zero-crossing detection, Fast Fourier Transformation (FFT), and filtering techniques.

1.4.2 Step length estimation

Step length is the distance between the point of initial contact of one foot and the point of initial contact of the opposite foot. In normal gait, right and left step lengths are relatively similar. Step length does, however, vary between people according to height, gender, age, physical condition, etc. An inertial navigation system (INS) calculates the distance by double integration of the acceleration signal. Yet, the integration error due to noise, bias and other disturbances is not negligible and can increase rather quickly. There are various algorithms designed to correct or compensate for these errors, such as ZUPT, which utilises the stance and swing phase during walking. This implies that ZUPT is not suitable for PDR with head-mounted sensors.

Machine learning methods were also normally used in step length estimation, such as using stacked autoencoders [53], Long Short-Term Memory (LSTM) and denoising autoencoders [165], etc. There are other model-based step length estimators, which propose equations to capture the relationship between step length and other step characteristics, such as step frequency, maximum or minimum acceleration, or acceleration variation. For example, Do et al. [36] estimated the horizontal displacement using vertical acceleration, which is based on the double integration of vertical acceleration followed by the use of the inverted pendulum model. The most commonly used models consist of the linear model [103], Weinberg model [166], Kim model [91], Scarlett model [143] and Shin model [149].

1. The linear model represents the linear relationship between step length and walking frequency f (Hz).

$$step\_length = a \cdot f + b \tag{1.8}$$

2. Weinberg method utilises the difference of the vertical acceleration values within a step.

$$step\_length = k \cdot \sqrt[4]{a_{max} - a_{min}} \tag{1.9}$$

3. Kim method is only based on the average acceleration within a step.

$$step\_length = k \cdot \sqrt[3]{\frac{\sum_{i=1}^{N} |a_i|}{N}} \tag{1.10}$$

4. Scarlett method eliminates the spring effect of the human gait by using minimum, maximum, and average acceleration.

$$step\_length = k \cdot \frac{\frac{\sum_{i=1}^{N} |a_i|}{N} - a_{min}}{a_{max} - a_{min}} \tag{1.11}$$

5. In Shin model, not only the step frequency but also the variance during that step is involved. So, it is more precise than the frequency singly related model listed before.

$$step\_length = a \cdot f + b \cdot v + c \tag{1.12}$$

where a, b, c, k are coefficients, a_{max} and a_{min} (ms^{-2}) are the maximum and minimum accelerations in one step, N is the number of samples in one step, v is the variance of the accelerometer signal in one step, a_i stands for the accelerometer signal at time i.

1.4.3 Orientation estimation

Orientation can be integrated directly from gyroscope data using the method described in Section 1.3.1. However, the error of gyroscope will be integrated as well, especially the constant bias, which grows linearly with time when integrated, leading to non-negligible error in long-term estimation. Previously mentioned methods ZUPT and ZARU can also help correct the error of orientation estimation, but only works for foot-mounted sensors. Other condition-specific methods will be introduced in the next chapter with their conditions and limitations. The more general way is using model-based state estimation methods on both the estimated step length and orientation to optimise the result.

1.4.4 Model based state estimation methods

1.4.4.1 Kalman filter

A Kalman Filter (KF) is a recursive Bayesian filter, known to be an optimal filter for Gaussian linear systems. Bayesian models belong to probabilistic models. A probabilistic model is a mathematical quantitative description of an uncertain situation, with two main elements including the sample space containing the set of all possible outcomes of the experiment; and the probability law assigning probabilities to outcomes. There are two different ways to quantify the probabilities: the frequentist with maximum likelihood estimation (MLE) and Bayesian with maximum a posteriori (MAP). Frequentist methods are more traditional and focus on the long-term properties of procedures, while Bayesian methods explicitly incorporate prior knowledge and update beliefs based on new data.

KF uses a series of measurements that are observed over time (containing statis-

tical noise and other inaccuracies) to produce estimates of unknown variables. These estimates tend to be more accurate than those based on a single measurement alone. An output from the filter is obtained by estimating the joint probability distribution over the variables for each given time frame. KF is generally conceptualised as two distinct phases: "Predict" and "Update".

"Predict" includes a priori state estimation:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \mathbf{x}_{k-1|k-1} + \mathbf{B}_k \mathbf{u}_k \tag{1.13}$$

and a priori covariance esitmation:

$$\hat{\mathbf{P}}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^{\mathsf{T}} + \mathbf{Q}_k \tag{1.14}$$

"Update" includes Kalman gain calculation:

$$\mathbf{K}_k = \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^{\mathsf{T}} (\mathbf{H}_k \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^{\mathsf{T}} + \mathbf{R}_k)^{-1} \tag{1.15}$$

a posteriori state estimation:

$$\mathbf{x}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}) \tag{1.16}$$

a posteriori covariance estimation:

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \hat{\mathbf{P}}_{k|k-1} \tag{1.17}$$

where F_k is the state transition model, H_k is the observation model, Q_k is the covariance of the process noise, R_k is the covariance of the observation noise, B_k is the input control model and u_k is the control vector.

1.4.4.2 Particle filter

The particle filter (PF) is a non-parametric Bayesian filter that represents the entire state vector using a set of particles. Each particle is a sample from the entire state space, which carries a weight, and the collection of particles approximates the probability distribution. The algorithm starts with an initial set of particles, typically

drawn from an initial estimate of the state. In the prediction step, each particle is propagated forward in time according to the dynamic model of the system. This step accounts for the evolution of the system over time. The particles are updated on the basis of the measurements obtained from sensors. Following the update step, the likelihood of each particle given the measurements is computed. Particles that are consistent with the measurements are assigned higher weights, and those that are less consistent are assigned lower weights. To ensure that the weights sum to one and represent a valid probability distribution, the weights of the particles are then normalised. After a few iterations, some particles gain large weights, while weights of others will be almost zero. This degeneracy problem would then be fixed by re-sampling the particles with probabilities proportional to their weights. This step emphasises particles that are more consistent with both the system dynamics and the measurements. Resampling helps prevent particle depletion and maintains a diverse set of particles. Finally, the result of the system state estimation is obtained by combining the information from all particles, with more weight given to particles that are more consistent with the measurements.

It has the advantage for situations where the posterior distribution is complex and may not have a simple analytical form, such as non-linear or non-Gaussian systems, which makes it suitable for a wide range of real-world applications. However, in high-dimensional state spaces, particle filters can suffer from the "curse of dimensionality," requiring a large number of particles to maintain a representative sample of the state space, leading to computational challenges.

Although these model based methods may have accurate results under experimental environments, their performance can reduce quickly in real world scenarios with longer run times and more complex motions. Performance is also negatively affected if there is more variation between users. Factors such as age, ability, and lived experience are likely to lead to different behaviour patterns. Movement is not only affected by intrinsic factors. It is also affected by the environment [10]. This is further augmented if there is an increased unpredictability of the environment

or if the user activities themselves become less predictable. Because model-based pedestrian inertial navigation methods represent the functional relationships between inertial measurement inputs and estimated localisation information outputs with simple fitting equations, such as a simplified inverted pendulum model, and individual parameters for each person, which will be ineffective with more complex motions in real scenarios.

To solve this problem, mapping the hidden relations between measurements and real motions to a higher-dimensional space is one of the ways that scientists are searching for, which are mainly based on the Deep Learning (DL) techniques.

1.5 Deep learning based inertial positioning systems

1.5.1 Introduction of deep learning

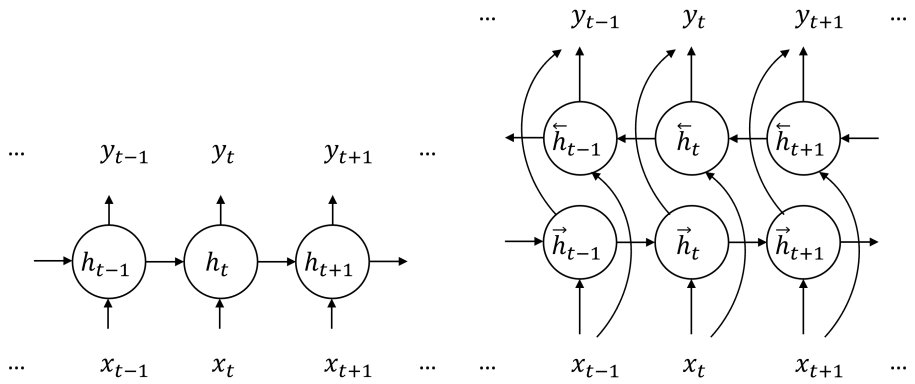
Deep learning is part of machine learning techniques that are based on artificial neural networks with multiple layers and parameters. With the rapid advancement of computing hardware, deep learning is developing quickly and has already shown its impressive potential in solving tasks in a variety of fields, including natural language processing [33], image processing [191], and healthcare [127] [72]. One of the key benefits consists of the ability to model complex non-linear relationships with large volumes of data. The precise identification of patterns or trends indicates its suitability in pedestrian inertial navigation.

Several different deep learning models have been used in this thesis. In the following, a basic explanation is given for each of the possible techniques.

1.5.1.1 Recurrent Neural Network (RNN)

RNNs are designed to process sequential or time-series data, like text, voice, sequences of stock prices or weather, for ordinal or temporal problems. These problems contain domains such as neural machine translation (NMT) [112], handwriting recognition

[38] and generation [52], speech synthesis [45], protein structure prediction [150], visual recognition and description [39].



(a) RNN (b) Bi-RNN

Figure 1.2: The architectures of (a) Recurrent Neural Network (RNN) and (b) Bidirectional Recurrent Neural Network (Bi-RNN).

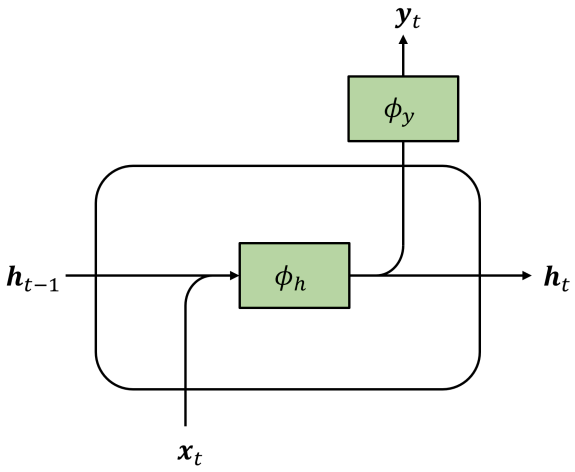
With a input sequence $x = (x_1, ..., x_t)$, a standard RNN model updates the hidden states $h = (h_1, ..., h_t)$ and produces the output vectors $y = (y_1, ..., y_t)$ sequentially, which is shown in Figure 1.2a. Each circle represents one RNN node, with the detailed structure illustrated in Figure 1.3a. In each cell, the current time step x_t and the hidden state of the last time step h_{t-1} are processed to produce the current hidden state h_t and output y_t via the Equations 1.18, where ϕ represents a full connected network with an activation function (sigmoid, tanh, ReLu, etc.)

$$\begin{aligned} h_t &= \phi(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \\ y_t &= \phi(W_{hy}h_t + b_y) \end{aligned} \tag{1.18}$$

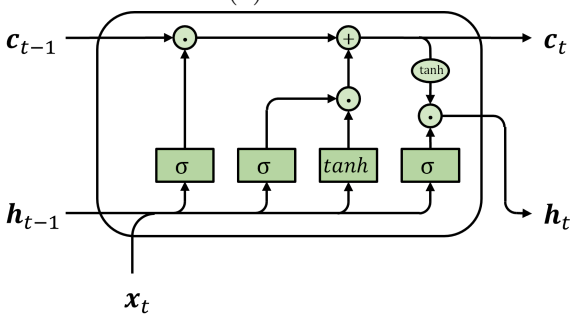
The output layer of standard RNN can only get information from past states. In order to get information from both past (backward) and future (forward) states simultaneously, Bidirectional Recurrent Neural Networks (Bi-RNNs) were introduced to increase the amount of input information by connecting two hidden layers in opposite directions [144], which is shown in Figure 1.2b. The hidden state h_t at each time step t is updated via Equation 1.19, where $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ denote the hidden states from the forward and backward layers. It is especially useful when the context of the

input is needed, such as predicting one word in the middle of a paragraph.

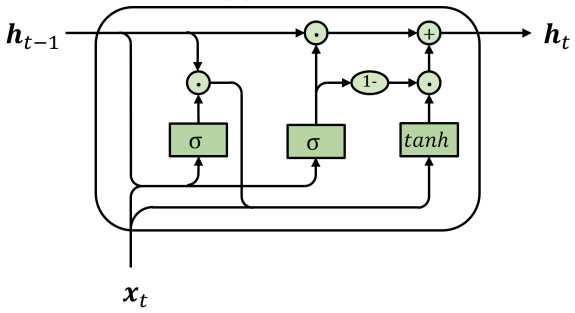
$$h_t = [\overrightarrow{h_t^T}; \overleftarrow{h_t^T}] \tag{1.19}$$



(a) RNN



(b) LSTM



(c) GRU

Figure 1.3: The Architectures of (a) RNN node (b) LSTM merory cell (c) GRU cell.

The vanilla RNNs face the problem of gradient vanishing in practice, which results in losing long-term dependencies. LSTM networks were then proposed which successfully solve the vanishing gradient problem partially [66]. LSTMs replace the

ordinary recurrent node by a memory cell, which contains an internal state besides the hidden state, to ensure that the gradient can pass across a lot of time steps without vanishing, thus keeping the information for a longer term. The memory cell of LSTM is illustrated in Figure 1.3b with the internal state and hidden state updated via Equations 1.20, where i, f, o represent the input gate, the forget gate and the output gate, c and h are the cell's internal state and hidden state, σ is the activation function, Ws and bs are weight parameter and bias parameter in the full connected layer, $\odot$ denotes the Hadamard (elementwise) product operator.

$$\begin{aligned} i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \\ f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \\ o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o) \end{aligned} \tag{1.20}$$

$$\begin{aligned} c_t &= f_t \odot c_{t-1} + i_t \odot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\ h_t &= o_t \odot \tanh(c_t) \end{aligned}$$

Gated Recurrent Units (GRU) [25] is a streamlined version of LSTM with a simplified architecture and faster computation speed, but it achieves comparable performance [28]. The hidden state in the GRU cell is updated by Equations 1.21 and illustrated in Figure 1.3c, where σ is the activation function, Ws and bs are the weight and bias parameter in the full connected layer, r and z are the reset and update gate.

$$\begin{aligned} r_t &= \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r) \\ z_t &= \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z) \\ h_t &= z_t \odot h_{t-1} + (1 - z_t) \odot \tanh(W_{xh}x_t + W_{hh}(r_t \odot h_{t-1}) + b_h) \end{aligned} \tag{1.21}$$

1.5.1.2 Convolutional Neural Network (CNN)

CNN [98] was initially designed for image processing as it was inspired by the receptive fields in the visual cortex of brain [79]. It has been the dominant machine learning approach in image processing tasks for over 20 years. However, with the introduction

of 1D convolutional layers, CNNs have become increasingly effective in processing sequential data such as audio [1], text [86], and time series analysis [97].

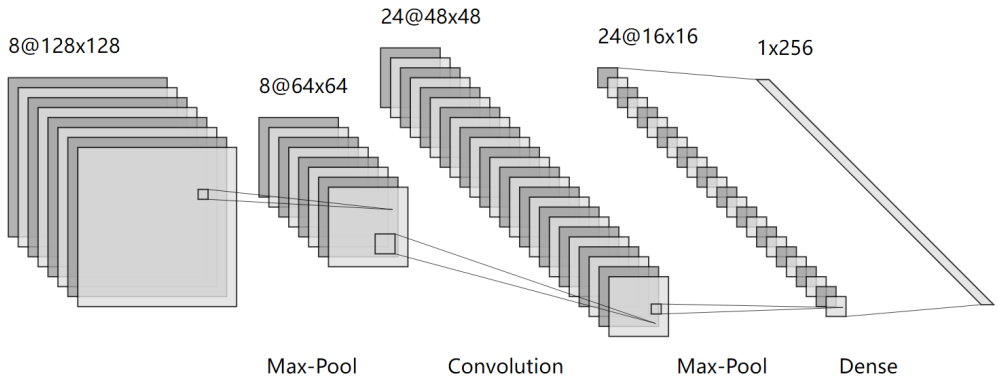


Figure 1.4: A structure of a Convolutional Neural Network (CNN).

An example of CNN structure is shown in Figure 1.4[1]. Convolutional layers are the fundamental blocks of CNNs, which are designed to extract features from input data by applying convolution operations with learnable filters sliding over it. A set of feature maps will be produced and pass through a non-linear activation function to introduce non-linearity into the model. A subsequent pooling layer will reduce the spatial dimension of the feature maps while preserving their essential information.

With the ongoing advancement of computer hardware, it has become possible to train CNNs with deeper structures that incorporate more layers for more complex models. As CNNs have been developed with more layers in deeper structures, an important issue that was mentioned previously has arisen, known as the "gradient vanishing" problem. This refers to the phenomenon in which the gradients propagated through the network during training become extremely small as they move backward through the layers, making it difficult to update the weights of the earlier layers and hindering the overall training process.

[1]Plotted with the software: LeNail, (2019). NN-SVG: Publication-Ready Neural Network Architecture Schematics. Journal of Open Source Software, 4(33), 747, https://doi.org/10.21105/joss.00747

A variety of different methods have been designed to solve this problem. ResNet [61] proposed residual blocks, which allows faster propagation of inputs through the residual connections across layers. DenseNet [76] is an extension of ResNet to some extent, where each layer is connected to all preceding layers by concatenations (rather than additions in ResNet) to maintain and reuse features from earlier layers. The structures are compared in Figure 1.5

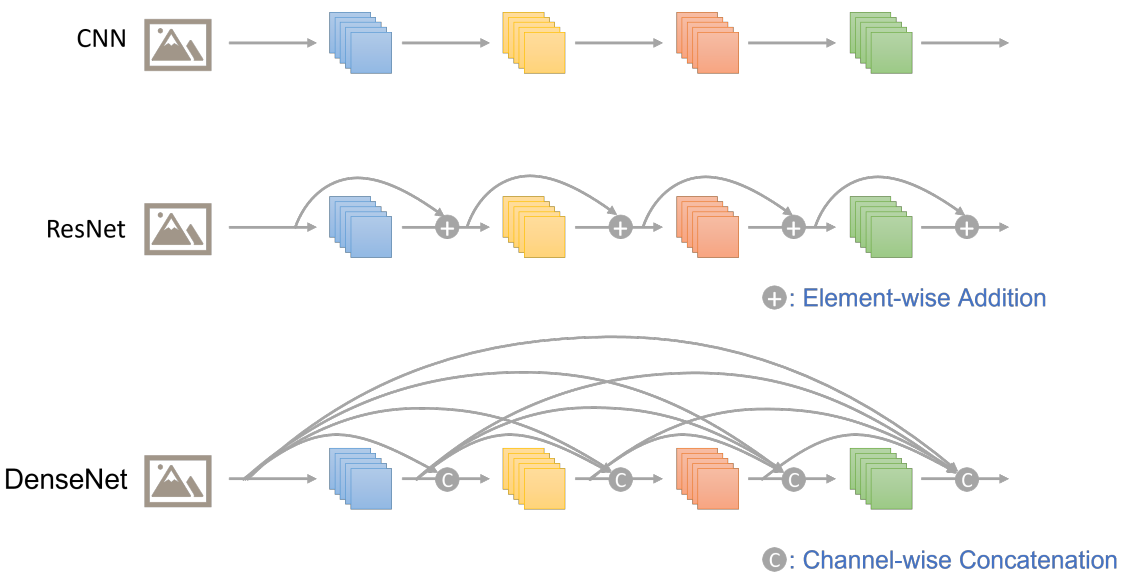


Figure 1.5: Structures of CNN, ResNet, DenseNet.

1.5.2 Inertial navigation systems with deep learning

IONET [22] was the first machine learning human inertial tracking method proposed in 2018. It adopted a 2-layer bi-directional LSTM to predict the rotation and distance in each $2s$ window. It was used in a study that applied it for positional tracking with a smartphone. A range of different phone placements were explored, such as keeping it in the hand, placing it in the pocket, or even carrying it in a handbag. Another inertial navigation method also used the IMU on a smartphone. RIDI [175] was designed to estimate walking velocities based on historical data of linear accelerations and angular velocities, and subsequently corrected the data for any low-frequency errors in the linear accelerations to ensure that their integration was consistent with predicted

velocities. Then a standard double integration was used to estimate the trajectory from these corrected linear accelerations. Finally, RONIN [62] used ResNet to regress 2D displacement vectors from normalised IMU data. It also focused on tracking with smartphones using different "placements" whilst measuring across a set of real day-to-day activities. Currently, almost all deep learning inertial tracking methods have been developed for smartphones. However, sensor placement will become even more important in the future, which will require new systems to be developed for other promising sensor locations.

1.6 Motivation

As mentioned previously, inertial tracking has a broader applicability in a wide range of scenarios, which comes from its cost-effectiveness, widespread availability, and self-contained nature. Promising applications include tracking athletes in sports for performance improvements or fatigue level estimations, recording trajectories of first responders or fire fighters in emergency rescues for safety, obtaining position information of medical staff and patients in hospitals for efficient allocations, etc.

Chapter 2 found that foot-mounted sensors are the most commonly utilised sensors in prior studies on human tracking with wearable devices, exhibiting the highest estimation accuracy as well, which is because the steps have direct relation with foot movements and it is easy to build the biomechanical model for them. But they are not convenient for daily use, as people need to specifically attach extra sensors to the shoes.The wear and tear of normal shoes is such that these objects often need to be replaced. The smartphone was concluded to be the second most common device in human tracking due to its popularity, but people do not always have it on them at home or can not take it during sports.

Sensors that could be worn in most of the scenarios without extra installations are needed for the more convenient, unobtrusive, discreet, and all-day tracking. Head-mounted devices provide a possible better solution to it. There are already many

head-mounted devices integrated with inertial sensors that can be leveraged, such as smart glasses, instrumented mouthguards, earphones, hearing aids, and smart caps. These locations are more convenient to use, while they also do not influence normal behaviour much since they are integrated into already widely used products. Current tracking methods with sensors at one position are hard to simply transfer to another position, as different parts of the human body show different moving patterns during activities, with a different set of signals from different biomechanical models. Special tracking methods are needed for head-mounted sensors. However, only 2 out of 145 previous wearable tracking studies adopted head-mounted sensors, which shows a significant research gap.

On the other hand, consumer-grade IMUs suffer from significant noise and bias, which brings considerable errors to inertial tracking, leading to even uncontrollable and unacceptable results in long-term tracking. Meanwhile, unpredictable human behaviours with the extra motions from the body part to which sensors are attached will lead to more complex patterns in the inertial data. This will influence position tracking, and any proposed algorithm needs to be able to adequately cope with these signals. Moreover, more realistic behaviour should be recorded to verify such systems. Most of the studies only focused on situations where participants were walking with at a constant pace, which does not reflect well on the more complex actions we create in real world scenarios.

These identified research gaps form the motivation for this thesis to *develop inertial tracking methods, for more conveniently placed wearable devices, and achieve accurate and robust long-term localisation.*

1.7 Research questions

In this thesis, head-mounted wearable devices are chosen as research subject. The main challenges faced by inertial tracking with head-mounted sensors are as follows:

- **Head motion confusion.** During human movement, the head may have its

own motions, such as turning to one side to look around. These kinds of motions would overlap with locomotion signals in the IMU data. For example, an increased angular velocity measured in the gyroscope may be caused by changing the walking direction or might just be a pure head rotation. If these signals cannot be correctly distinguished then this would lead to a heading estimation error.

- **Error accumulation in long-term tracking.** Tracking with only inertial sensors creates the problem that there is no simple way to determine the global (world) reference frame without additional external reference points. This implies that no "recalibration" can take place to help correct for estimation errors. Thus, the errors will accumulate as time goes on, resulting in unacceptable large differences between real and estimated positions when the aim is to monitor over a longer period of time.

- **Various locomotion.** Human locomotion not only consists of walking at a constant pace. In the real world, people can walk or run at various speeds, which increases the complexity of patterns in the IMU data.

1.8 Contributions

- **Chapter 2** is a systematic review of all available studies that use pedestrian dead reckoning to track human motion with wearable devices. It provides a summary of the sensor types, placements, algorithms, testing environments, and recorded findings. This work has been published in:

  > Hou, Xinyu, and Jeroen Bergmann. "Pedestrian dead reckoning with wearable sensors: A systematic review." IEEE Sensors Journal 21.1 (2020): 143-152.

- **Chapter 3** proposes a novel PDR method for head-mounted sensors. This has been published in:

Hou, Xinyu, and Jeroen Bergmann. "A pedestrian dead reckoning method for head-mounted sensors." Sensors 20.21 (2020): 6349.

- **Chapter 4** introduces a method to solve the error accumulation problem in long-term tracking. It has been published in:

  Hou, Xinyu, and Jeroen Bergmann. "HeadSLAM: pedestrian SLAM with head-mounted sensors." Sensors 22.4 (2022): 1593.

- **Chapter 5** proposes a machine learning based inertial tracking method that solves the problem of random head motions during localisation. This research has been published in:

  Hou, Xinyu, and Jeroen Bergmann. "HINNet: Inertial Navigation with Head-Mounted Sensors Using a Neural Network." Engineering Applications of Artificial Intelligence, Elsevier, 2023.

- **Chapter 6** describes an extension of the HINNet approach in Chapter 5 to allow for both long term tracking and free head motions. This has been submitted to:

  Hou, Xinyu, and Jeroen Bergmann. "HINNet + HeadSLAM: Robust inertial navigation with machine learning for long-term stable tracking", In submission to IEEE Sensors Letters.

- **Chapter 7** proposes a machine learning inertial tracking method based on a probabilistic DenseNet model, which could estimate both pose and uncertainty. It supports positional estimates during walking and running at various speeds, with free head motions and robust long-term tracking. It has been submitted to:

  Hou, Xinyu, and Jeroen Bergmann. "ROCIP: RObust Continuous Inertial Position tracking for complex actions emerging from the interaction of human actors and environment", In submission to 2023

International Conference on Intelligent Robots and Systems (IROS 2023).

Chapter 2

Pedestrian dead reckoning with wearable sensors: A systematic review

Pedestrian Dead Reckoning (PDR) plays an important role in modern life, including tracking the locations of people whenever GPS is not available. Self-contained PDR systems do not require an infrastructure, thus they can be used for rapid deployment in situations such as search and rescue, disaster relief, or medical emergencies. Wearable sensors are often applied in self-contained PDR, but implementation varies in terms of the number, type, and placement of sensors used. Many algorithms are designed for PDR in order to reduce the error or drift of the final estimate, with various levels of success. There is a lack of comparison between these different methods. This systematic review provides a comprehensive overview on sensor types, layouts, algorithms, and evaluations for all available PDR literature with a focus on wearable sensors. Further research directions are suggested on the basis of these results. This study also highlights the need for more standardised and robust assessment protocols to capture real-world tracking performance of PDR methods.

2.1 Introduction

Human position tracking technology, which provides information on the current location of an individual, has been rapidly developing over the last decades. It has

drastically changed modern life by offering information regarding the local position of things even during complex scenarios. LBS are supported by human positioning technologies, which provide help in a variety of contexts including health, entertainment, work, personal life, etc.

A wide range of sensors have been adopted in different pedestrian position tracking approaches. These approaches can be classified into two main categories: (i) dependent and (ii) self-contained methods. Dependent methods include the well known GNSS, and other methods relying on external aiding signals, information or infrastructure, such as using wireless local area network (WLAN), BLE beacons, etc. However, these methods are not applicable in environments where signals are affected or without specific infrastructures, which limits their using scenarios. On the contrary, self-contained position tracking does not rely on any infrastructure, which implies that it could be used in any kind of environment. These infrastructure-independent methods for human localisation are mainly based on Pedestrian Dead Reckoning (PDR), whilst utilising on-body sensors.

Most of the sensors used for PDR are small and portable. They include IMUs, accelerometers, gyroscopes, compasses, barometers, magnetometers, anemometers, and on-body cameras, to name some of the most popular types. These sensors are usually combined to create a multimodal system and are commonly mounted on the foot, wrist, waist, leg, or head. They are relatively low cost and have low energy consumption, making them more attractive for use. PDR does not depend on any infrastructure, so it can be readily deployed in almost any situation. It is a particularly useful application when conditions change rapidly. Thus, PDR is suitable during environmental emergencies, as the surroundings might be changing frequently. It also works well under smoky and dim lighting conditions, making it ideal for firefighting or rescue operations that could encounter similar environments.

Inevitably, measurement errors exist in sensor data, and they could accumulated over time to reach a critical value. Highly accurate sensors are used in aviation and marine applications with the aim to keep the error at an acceptable level, whilst track-

ing positions over several hours. However, these sensor systems are too expensive and bulky for wearable use. The key user preferences for body-worn sensor systems are often the compact nature, the ability to embed it in an everyday object, and the simplicity of operating / maintaining the system [11]. These specifications are captured perfectly in MEMS. However, the open-loop integration of MEMS inertial sensor data is only suitable for a few minutes before drift dominates the signal. Therefore, a lot of research has been done to develop effective algorithms that can reduce these errors. They range from ZUPT[47] to Extended Kalman Filter (EKF)[6] and Heuristic Drift Reduction (HDR)[14], etc.

Several review studies have already been published on PDR methods, but all of them have been limited in scope. For example, [58, 172] are surveys of the research status of PDR and only include inertial sensors; [35] only focused on step length estimation methods; [49] explored PDR only using inertial and magnetic sensors, and [29] reviewed PDR specifically for mass market applications. Despite these numerous reviews of PDR, none of them performed a comprehensive analysis of PDR across a range of wearable sensors. More importantly, none of them conducted a systematic search or included a quality assessment of the research.

In this chapter, a systematic review was conducted on all studies related to PDR with wearable sensors. This study provides an overview of the sensor types, sensor placements, algorithms, and evaluations applied in each study. And a quality assessment was conducted for all selected studies.

2.2 Methods

2.2.1 Search strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[117] was used to structure the content provided below. The literature search was conducted using these three databases: (i) PubMed, (ii) Web of Science, and (iii) IEEEXplore. A total of 18 keywords were selected and combined with Boolean operators

to obtain all papers related to PDR and wearable sensors. All papers published before February 2020 were included in this systematic search.

The search strategy applied consisted of the following keyword combination:

> ("dead reckoning" OR "position tracking" OR localisation OR "position monitoring") AND ("physical activity" OR sports OR walking OR running OR pedestrian) AND (accelerometer OR "Inertial Measurement Unit" OR Gyroscope OR Magnetometer OR "Body Sensor Network" OR "wearable sensor" OR "wearable sensors") NOT GPS NOT wifi

2.2.2 Study selection

A variety of sensors could be used for PDR, but only studies using wearable sensors were included. The wearable sensors consist of sensors, such as IMUs, gyroscopes, magnetometers, compasses, Electromyography (EMG) sensors, sonar, on-body cameras, barometers, and anemometers. Studies that used sensors that were not attached to the body or relied on other infrastructures (for example GPS, Wi-Fi, RFID, UWB, Bluetooth beacon, radio, etc.) were excluded. Studies that amend trajectories with previously known information or data were also excluded. Therefore, research that relied on information taken from an earlier established map, floor plan, magnetic fingerprints, light fingerprints, or similar constructs were not included.

The inclusion criteria for the articles were:

1. It must be written in the English language.
2. It must be peer reviewed.
3. It must contain an outcome measure - methodology alone does not suffice.
4. It must explore a PDR method using wearable sensors.
5. It must report a quantitative tracking error.

The characteristics and data extracted from each study consisted of (i) year of publication, (ii) algorithm used, (iii) type, amount, and position of sensors, (iv) error and total travel distance.

All eligible studies that were not accessible through library services were attempted to be obtained by contacting the corresponding author.

2.2.3 Quality assessment

The quality assessment of each paper followed the questions of the adapted Specialist Unit for Review Evidence (SURE) to assist with the critical appraisal of the studies [157]. The checklist was modified to better fit the field of this systematic review. The modified checklist included 8 questions, which can be found in Table 2.1.

Table 2.1: Modified SURE critical appraisal checklist for systematic review and 9 samples of quality scores

Item	Checklist	Wang [164]	Shin [148]	Shi [146]	Li [105]	Guo [54]	Park [122]	Lee [101]	Yang [176]	Cardarelli [20]
1	Does the study address a clearly focused question?	1	1	1	1	1	1	1	1	1
2	Do the authors discuss how they decided which method to use?	1	1	1	1	1	1	0	1	0
3	Is there sufficient detail regarding the methods used?	0	0	1	1	1	0	1	0	1
4	Is the data collection method well described?	0	1	1	0	1	0	0	1	1
5	Are the explanations for the results plausible and coherent?	0	1	1	0	1	0	0	1	1
6	Are the results of the study compared with those from other studies?	1	0	1	0	1	0	0	1	0
7	Did the authors identify any limitations?	0	0	1	0	1	0	1	0	0
8	Was ethical approval sought?	0	0	0	0	0	0	0	0	1

2.3 Results

A total of 877 records were identified by searching the aforementioned databases. A total of 265 duplicates were removed from the list. The titles and abstracts of 612 remaining papers were subsequently screened by both authors. The full manuscript was assessed for the 307 cases that remained after review of the title and abstract. The 307 articles were obtained using the University library system, online resources, and by contacting corresponding authors. This led to the full review of 301 papers, as not all articles were available. Finally, 145 studies met the inclusion criteria and were

further evaluated (see Figure 2.1). A sample summary of the study characteristic is presented in Table 2.2. The complete table for characteristic summary and quality assessment can be found in Appendix A.

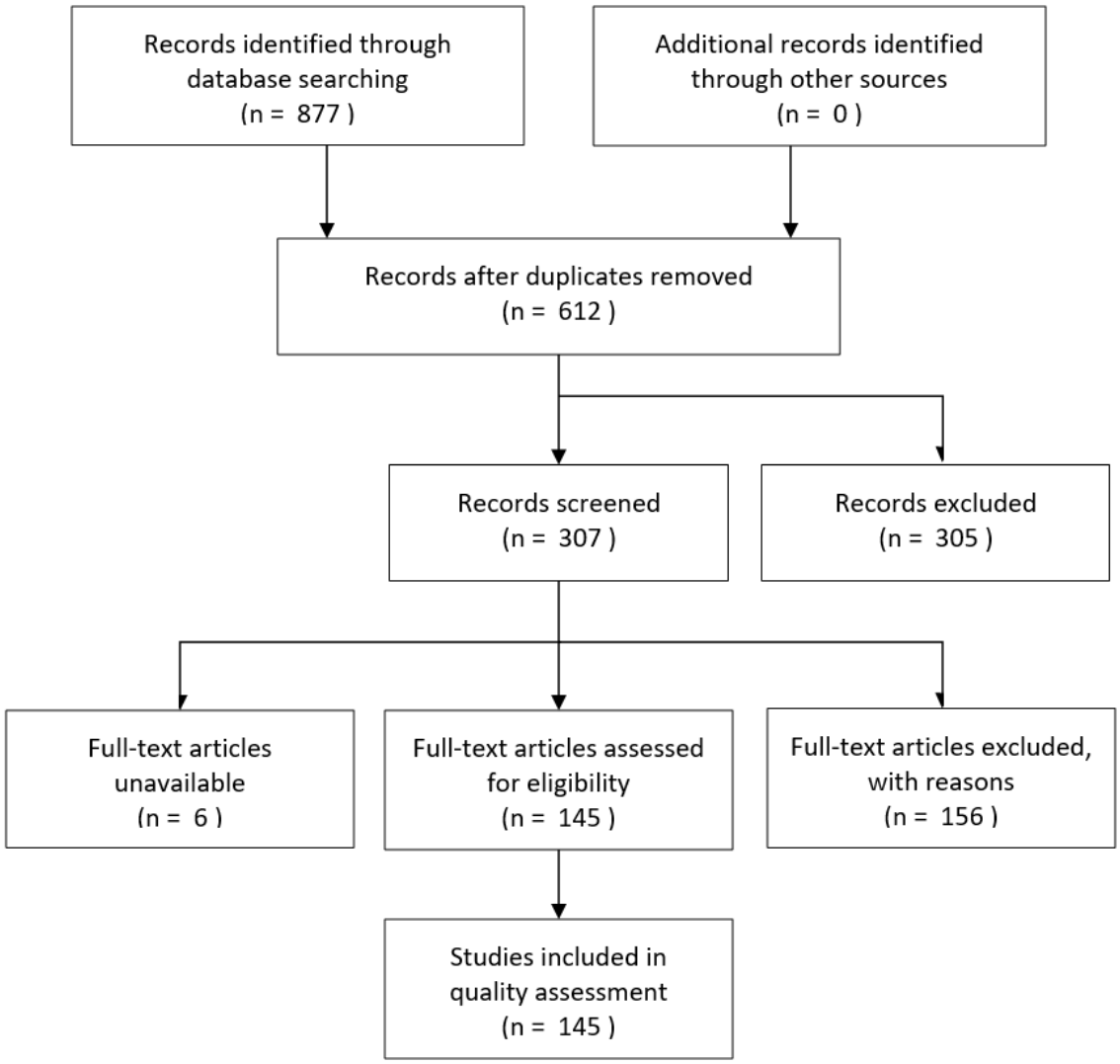


Figure 2.1: PRISMA flowchart of systematic review. The process consisted of identification, duplicate removing, screening and inclusion of relevant papers.

Table 2.2: Sample of study characteristics summary

First author, year of publication	Sensor layouts	Algorithms	No. of subjects	Mean error (m)	Total travel distance (m)	Accuracy	Environment	Dimension
Wang, 2016[164]	A smartphone, a foot-mounted IMU	zero velocity update, Indoor landmark detection, online magnetic trajectory matching	1	2.55	-	80 percentile localisation accuracy of 1.4m	Indoor	2D
Shin, 2014[148]	A smartphone	motion recognition, ZUPT, KF, SVM	3	2.25	133.2	5.00%	Indoor	3D
Shi, 2019[146]	An IMU mounted on foot	zero velocity update (ZVU), Gait Phase Detection Algorithm, KF	1	0.74indoor, 2.33outdoor	103.86indoor, 402.7outdoor	0.73%indoor, 0.58%outdoor	Indoor + Outdoor	2D
Li, 2017 [105]	A foot-mounted IMU, magnetometer	KF, peak detection	1	4.27	400	1.00%	Outdoor	2D
Guo, 2015[54]	A Foot-Mounted IMU	ZARU, ZUPT	1	0.88	90.48	1.00%	Indoor	2D
Park, 2012[122]	A waist-worn IMU	HDR	1	-	64.48	3.00%	Indoor	2D
Lee, 2002[101]	A biaxial accelerometer and the gyroscope located in one trouser pocket; a digital compass and a microcontroller attached to the middle of the user's waist.	fuzzy logic algorithm	8	-	90	8.20%	Indoor	3D
Yang, 2017[176]	A smartphone (camera)	Hybrid Orientation Filter	5	0.83	207	0.41%	Indoor	2D
Cardarelli, 2019[20]	A foot-mounted IMU	KF, ZUPT	1	0.96	43.16	2.22%	Indoor	2D

2.4 Descriptive summary of results

2.4.1 Sensor types and layouts

A variety of different sensors were applied in the studies that were identified. The difference in the type, the amount, and the placement on the body of the sensors also led to contrasting algorithm designs.

2.4.1.1 Types

The IMU was the most common used sensor (n=81) across all the studies. It consisted of a triaxial accelerometer, a triaxial gyroscope, and sometimes a triaxial magnetometer. It was reported that the data from the magnetometer was not always a reliable source for heading estimation, due to the existence of soft and hard magnetic disturbance in the environment. This was especially true for indoor environments. Only a few studies (n=32) adopted magnetometers for azimuth calculation.

Some studies also used other wearable sensors to increase the tracking accuracy,

beyond the application of IMUs. Cameras were used to extract features of the environment [176], while others used pressure sensors under the shoe to detect heel-strike and toe-off points that could aid in the step detection[23]. EMG sensors were also applied to get EMG signals of the Gastrocnemius for both step detection and step length estimation[24]. Other methods that were reported leveraged a compass to obtain geomagnetic information[186], used a barometer to estimate the vertical moving distance[5], used an anemometer to estimate walking speed[155], used an ultrasonic sensor to estimate step length[193] or used a laser scanner to calculate relative distances[73]. A large proportion of the studies included (n = 49) applied smartphones as their data collection and measuring device.

2.4.1.2 Layouts

The sensors were mounted on a variety of locations on the body. Studies with a single IMU, most often mounting the sensor on one of the shoes[173]. Another commonly selected placement was the waist[82]. Studies with multiple IMUs preferred to mount them on the lower limb and trunk[100]. This included placement on the pelvis, thighs, shanks, and feet. There were studies that had an even larger network of IMUs available and also placed sensors on the upper limb in addition to the aforementioned locations[26]. For the upper limb, sensor placement consisted of the shoulder, upper and lower arms. Studies relying on a smartphone usually resorted to having the sensor near the hand or in the pocket at the front/back of the trousers, as well as the pocket located on the shirt/jacket [116]. Holding the smartphone in the hand allowed for a further breakdown based on the specific task that was performed on the phone. Separation of phone tasks was used to help with data analysis, and activities consisted of texting, swinging, and calling[147]. The study with the most diverse type of sensors, placed a camera on the chest, one IMU on the head, one on the back and one on the shoe, whilst also relying on sonar with the user carrying a computer on the back[140].

2.4.2 Algorithms

To robustly estimate the trajectory of pedestrians, the designed algorithms should be adapted to operate correctly under a range of different conditions. Ideally, they should also be able to work for a range of sensors, placements, motions, people, and environments. The algorithms were subdivided to better group the results of the studies. The PDR systems are divided into categories that are described in [58]:

Inertial Navigation Systems

An INS is a system that tracks position by estimating the full 3D trajectory of the sensor at any given moment. INSs are based on the theory that the first integration of the acceleration signal measured by the accelerometer produces the real-time velocity and the successive integration of velocity then gives the real-time displacement. The real-time attitude is produced by integration of the angular velocity, which is measured by the gyroscope. The accumulation of displacement and attitude is subsequently used. However, the integration error due to noise, bias, and other disturbances in this measurement is not negligible and can increase quickly. There are various algorithms presented in the papers to correct or compensate for these errors.

Step-and-Heading Systems

An SHS is specific to pedestrians, estimating position by accruing distance, heading vectors representing either steps or strides. In certain cases, a 3D trajectory is not required for a specific tracking scenario, and the 2D navigation in the horizontal plane using step vectors, rather than complete limb trajectories, is sufficient. SHSs output a series of step vectors by detecting each step of the pedestrian, estimating the length and direction of it, and finally integrating every step to obtain a complete trajectory.

There are also other techniques available for data analysis and these will now be discussed.

2.4.2.1 Zero velocity update (ZUPT) and zero-angular rate update (ZARU)

This approach relies on the idea that the amount of drift in the signal could be reduced if the integration loop can be closed using external constraints. The ZUPT and ZARU both make use of PDR constraints[173, 100, 6]. ZUPT was first used by Foxlin et al. in the NavShoe project, and good results were reported for gait [47]. Human walking consists of two distinct phases (stance and swing phase). The foot is in contact with the ground during the stance phase, which provides a useful reference point as the angular and linear velocity of the foot is close to zero during large portions of the stance phase. This information can be detected using the aforementioned algorithms and allows for the separation of individual steps. The ability to label each step provides a method to reduce the long-term accumulation of errors. However, ZUPT and ZARU are only suitable for tracking systems consisting of IMUs mounted on the foot.

2.4.2.2 Magnetic information

Geomagnetic fields cover the Earth's surface and this information can be sensed using magnetometers. Magnetometers measure the direction of the local magnetic field and can help derive the absolute heading direction. This information can provide a useful correction for the inevitable heading drift, which arises from the integration of gyroscopic output. However, the geomagnetic field is not always uniform and can be easily disturbed by hard or soft magnetic interferences. This means that the heading information produced from magnetometers does not always provide a stable directionality. Only a few studies apply magnetometers and they have specially designed algorithms to deal with these magnetic disturbances. These algorithms aim to calibrate, compensate, or rely on the disturbances in order to create a more accurate position estimate. Magnetic measurements can also benefit from the application of a Kalman filter. As an example, Ashkar[6] estimated the magnetic field offset, which is then treated similar to the gyroscope and accelerometer biases within the measurement model. Another common correction algorithm depends on fusing gyroscope out-

put with the compass information. This works well because of their complementary error characteristics. The gyroscope has an accumulation of long-term orientation errors, while magnetometers are subject to short-term orientation issues [96]. Finally, a Quasi-Static magnetic Field (QSF) can also be used to detect particular QSF periods during movement. The changes in magnetic field components during these periods can be compared with the angular rates of the overall sensor unit. Therefore, providing more information can help compensate for any arising errors [2].

Some researches exploit the fact that the magnetic field is not uniform. This provides the opportunity to use the sensor output for "fingerprinting". They collect magnetic field maps before tracking and then apply a matching algorithm to estimate position during real-time data collection based on the previously obtained map [83, 75, 162]. The concept of matching to an established map does yield good results, but articles that employ this technique are excluded from this review since they require predetermined information. Interestingly enough, there is a study that does take advantage of the magnetic anomalies, but does not rely on collecting a historic map before data collection. This study uses a process that consists of collecting magnetic anomalies in real time to calibrate the PDR location without the need of deploying any additional infrastructure or performing offline site surveying [164]. Shin et al. use a support vector machine (SVM) for the heading estimation by recognising the consistent patterns in magnetic fields when orientated in a certain direction[148].

2.4.2.3 Gait cycle detection

Gait cycle detection was indirectly discussed in the ZUPT and ZARU section. However, step detection is more widely used and it can serve different purposes, including step counting. Wu et al. proposed that gait division strategies could fall into two categories: (i) abbreviated and (ii) detailed strategy[172]. The abbreviated strategy consists of simply dividing the gait cycle into only one or two events. For example, divide it into a stance and swing phase [146]. The detailed division strategy refers to the separation of the gait cycle into three or more phases. One possible option is

the division of gait into four main stages such as a contact phase, mid-stance phase, propulsive phase, and swing phase[105]. Other breakdowns of the gait cycle in the temporal domain are possible.

Detection of these phases can be facilitated by threshold-based algorithms, which are typically used for stance detection. Corresponding gait phases can also be determined by measuring whether a specific set of features is part of the threshold set. Guo et al. set three thresholds to split-up the gait cycle. They used the magnitude and variance of the acceleration, as well as the magnitude of the angular rate, to estimate the stance phase[54].

Another way of doing it is by relying on a peak detection method to recognise gait and steps by detecting the maximum peak of the acceleration output signal [105]. Other popular methods in the time domain include a zero crossing detection method[122], the application of fuzzy logic[101] or applying autocorrelations [131]. These methods are usually based on accelerometer signals, and the data is often preprocessed by using a low-pass filter to reduce noise. Approaches in the frequency domain have also been explored and include a Short-Time Fourier Transform (STFT)[125] or a continuous/discrete wavelet transforms (CWT/DWT)[21].

Feature clustering approaches that employ machine learning algorithms have started to become more popular. Hidden Markov Models (HMMs) can be trained by two-state HMM with Gaussian emissions in an unsupervised fashion using the Viterbi algorithm and K-means clustering to classify activities based on features extracted from inertial measurements in both the time and frequency domains. Zhang et al. used an HMM as a gait detector during dynamic and fast gait speeds in [186]. Edel and Koppe used Bidirectional Long Short-Term Memory Recurrent Neural Networks (BLSTM-RNNs) to accomplish a 3D step recognition approach [42].

Finally, step detection can also be accomplished using sensors that do not contain inertial sensing capabilities. Chen et al. used pressure sensors embedded in the sole of the shoe[23], while Zizzo et al. used two position-sensitive detectors (PSDs) mounted on the toe and heel of the shoe[193] for phase detection. Chen et al. investigated the

ability of two EMG sensors that were attached to calves[24].

2.4.2.4 Step length estimation

One of the simplest approaches for step length estimation is to set the step length as a constant value. An average value can be set when people walk at their own "natural" pace. The detection of the step frequency can then be combined with this pre-set step length value. Issues arise when subjects start to deviate from this value. This can occur when they change their activity to running, turning, or even when they start to walk in groups. A dynamic step length estimation method is more appropriate in these situations. One such model is the Weinberg model[166], which assumes that the vertical bounce is proportional to the step length. The Kim model[91] aims to associate the step length and the average acceleration that occurs during a step. Scarlet's approach[143] focuses on the spring variation in the steps of different people or in the steps of one person walking at different velocities. Methods have also been introduced that build a model using linear methods[136] or SC-based methods [163].

Direct measurement of step length can also be accomplished with non-inertial sensors. Zizzo used five ultrasound receivers on the left foot and two ultrasound transmitters on the right foot to measure the step displacement[193].

2.4.2.5 Heading estimation

The heading is usually estimated by gyroscopes, which output a relevant angular velocity. The integration of gyroscope signals provide the heading change and, as previously mentioned, the heading drifts can then be limited by e.g. ZUPT or ZARU. However, ZUPT and ZARU are not very suitable when the sensors are not placed on the foot. Thus, sometimes headings need to be corrected by using other processes. For instance, HDR can limit the constant drift based on the assumption that when roads are straight, the pedestrian heading should remain unchanged[14]. The suitability of this method is decreased when reality starts to deviate from this assumption.

2.4.2.6 Data fusion

Filtering algorithms that integrate multiple observations are important in PDR systems with multiple sensors. The most commonly used algorithm among the presented studies is the KF[21]. This is a recursive Bayesian filter, known to be an optimal filter for Gaussian linear systems. KF uses a series of measurements that are observed over time (containing statistical noise and other inaccuracies) to produce estimates of unknown variables. These estimates tend to be more accurate than those based on a single measurement alone. The output of the filter is obtained by estimating the joint probability distribution over the variables for each given time frame. Many studies did not use the original KF, but applied extensions or generalisations of it. The following filters were reported; the Unscented Kalman Filter (UKF)[132], Cascaded Kalman Filter (CKF)[115], EKF[6], quaternion-based extended Kalman Filter[74], Adaptive Kalman Filter (AKF)[81] or Robust Adaptive Kalman Filter (RAKF)[168]. These were some of the more common adaptations, but other filter methods were also found. A Complementary Filter (CF) is such a filter. It is applied to make the accelerometer and gyroscope data complement each other to obtain the initial attitude[44]. The PF is another one. This filter is a numerical approximation to a Bayesian filter[64] and is often used in the estimation of walking length and direction. PF consists of many "particles" that represent possible positions and headings with a weighted value that indicates the probabilities of each. Adjusting these values is an iterative process consisting of three steps (update, correct, and resample) [57].

Wu et al. divided filtering methods into direct and indirect estimates. The direct estimation is a method where the filter directly outputs the final tracking data, whilst the indirect estimation refers to the fact that the filter estimates the errors of the tracking data before compensating for errors in the integration results.

2.4.2.7 Mode Classification

It is essential to keep in mind that pedestrians do not always walk at a regular pace. Gait itself can be affected by environmental or behavioural situations. Inderst et

al. classified patterns as walking, standing, ascending, and descending stairs before estimating any positions[82]. Hussain also classified the recorded signals into different motion states that included walking, running, and stopping[80].

The classification of modes can go beyond mobility classification, whenever a smartphone is used as the primary data collection system. Different classes can be defined in terms of how the phone is carried or held. Errors can occur if this is not taken into account, since there are various "holding" locations and carrying modes for smartphones [147]. The phone might be placed in the hand, located in the front trousers pocket, front shirt / jacket pocket, or back trousers pocket[116]. The ability to know where the phone is located creates an additional relevant layer in the process of estimating the position of the user.

Adopting these diverse tracking strategies for different modes aims to increase the overall tracking accuracy. Appropriate models are needed to correctly classify these modes, and most papers make use of machine learning algorithms to do so. The classification models encountered in this review are LSTM[80], Support Vector Machine (SVM)[183], Muti-layer perceptrons (MLP)[173], Probabilistic neural network classifiers (PNN)[152] and Random Forest models[55].

2.4.2.8 Simultaneous Localisation and Mapping (SLAM)

The SLAM proposition combines the localisation problem with a map estimation. It iteratively builds a map of the environment and localises the user within this map. Kaiser and Diaz investigated PocketSLAM by combining PocketNav with FootSLAM[85]. Hardegger et al. proposed ActionSLAM, which used location-related actions as landmarks[56].

2.4.3 Evaluation

The results show that no standard procedure was implemented in all the papers to evaluate the precision of PDR. The process by which the algorithm was evaluated differed greatly between the studies. Moreover, the underlying ground truth regarding

the positions that an individual passed over time also varied between the experiments. This lack of consistency makes it harder to compare the performance of each algorithm. Ideally, the exact same procedure and setup is used to create a condition in which only the selected algorithms differ between measurements.

The testing environments of the studies were heterogeneous, and thus a one-to-one mapping of accuracies between techniques was not possible. One area of divergence in terms of environment related to the experiments taking place either in- or outdoors. A total of 20 studies tested their systems in an outdoor environment, 108 papers stayed indoors, and 17 studies tested both in- and outdoor. For indoor evaluations, the ground truths of the tracking trajectories often came from optical motion capture systems[193], multiple cameras[56], labels along a certain route with a downwards-facing hand-held camera[42], centre of the route in a floor plan[130], laser distance meters[36], commercial robot localisation system including an infrared camera and a set of passive landmarks placed on the ceiling[43] or a general motion tracking system (MTS)[158]. All of these data collection systems come with their own levels of accuracy. They also limit the kind of activities that could be performed, as well as the distances and patterns that can be verified with these devices. For outdoor evaluations, ground truths came mainly from a GPS receiver[135] or simply by referencing a predetermined trajectory[146]. The same considerations need to be taken into account for gold reference systems used in an outdoor environment. Determining whether a certain accuracy of a PDR technique is generalisable to other scenarios is limited by the initial experimental setup. Creating results in terms of accuracy that translate easier to real-world deployments involves the inclusion of more realistic testing conditions. It is suggested that more work is done to reflect specific scenarios of interest during the testing of PDR approaches.

The tracking was performed in 2D or 3D. The 2D PDR estimated the trajectory of a person in the horizontal plane, while the 3D PDR provided an extra vertical dimension to calculate the current position. Most studies ($n = 105$) tracked people in 2D. However, the 3D methodology was still explored in 40 studies. Dimensionality

was not the only thing that differed in terms of the paths that were set in each study. There was also a great dissimilarity between the total travel distances in the experimental setups. The longest walking distance for one particular test was 4,029m[135], while the shortest was found to be 6m[151].

Some studies were conducted to evaluate the robustness of the system, for example, by testing trajectories with different shapes[179] or by testing with smartphones of different brands[94]. This is a suitable notion to investigate, as it is well connected to the need for more real-world evidence for PDR. Including variability into scientific research allows for a broader interpretation. The same benefit can be claimed for using multiple volunteers to assess PDR performance. People do produce changed gaits under different conditions, and capturing more representative groups of users will contribute to a potential greater uptake of this technology. Currently, the total number of test subjects varied considerably between studies. The majority of studies had only one participant (n=96), but 15 studies had more than 10 subjects and the maximum number of subjects recruited for a single paper was 30.

2.4.4 Errors

No standard method was adopted to calculate the error in all studies, which echoes the issue that was highlighted for the experimental setup. Researchers showed their tracking accuracy through a range of error representations. They consisted of calculating the end-to-end error, determining a position/distance error, defining the position error rate, giving a return position error, as well as reporting deviations on three axes and calculating the root mean square error or standard deviation. In this review, a percentage error was calculated to provide a better comparison between research studies. The smallest error found was 0.07%, while the largest error is 10.07%. The median of all errors reported for PDR was around 1%.

2.5 Discussion

This chapter performed a systematic review and quality assessment of studies that discussed PDR with wearable sensors. The selected papers were analysed across several factors to provide an easy comparison for those interested in this topic. The sensor type, placement, applied algorithm and evaluation methods were reported across all included studies. It is difficult to objectively compare the different PDR systems, as the experimental conditions varied along with the error metric that was used to quantify the performance of the technique. The field could benefit from adopting a testing method that can be applied more widely whenever systems and techniques need to be compared. The testing setup should aim to represent real-world scenarios to be able to generalise the results to potential functional applications.

Figure 2.2 presents an overview of the sensor placements and the average error for each position. It shows that the attachment of sensors to the foot may yield a better result, which is likely mainly due to the adoption of ZUPT or ZARU. These are two algorithms that are often used and can only be used in PDR with foot-mounted sensors. Filters like KF are also essential to PDR because of the commonly applied multi-sensor methods. Regarding the type of sensors, almost all the studies used IMUs, which is recognised as the dominant type of sensor for detecting movements. It is expected that IMUS will remain one of the core technologies for PDR in the future.

As stated above, many studies relied on IMUs to perform PDR. In [189], Zhao, et al. designed an approach to reduce the heading drift for foot-mounted sensors by implementing dual-gait analysis and multi-sensor fusion. The setup consisted of a dual-sensor configuration with one IMU on each foot. This study yielded a relatively small position error of 0.76%. It introduced a novel implementation of dual-gait analysis for multi-sensor fusion, relying on a Kalman-type filter. The advantage is that it does not require a personalised parametric adjustment or calibration for each individual user. It also primarily depends on the gyroscope signal, which has a lower

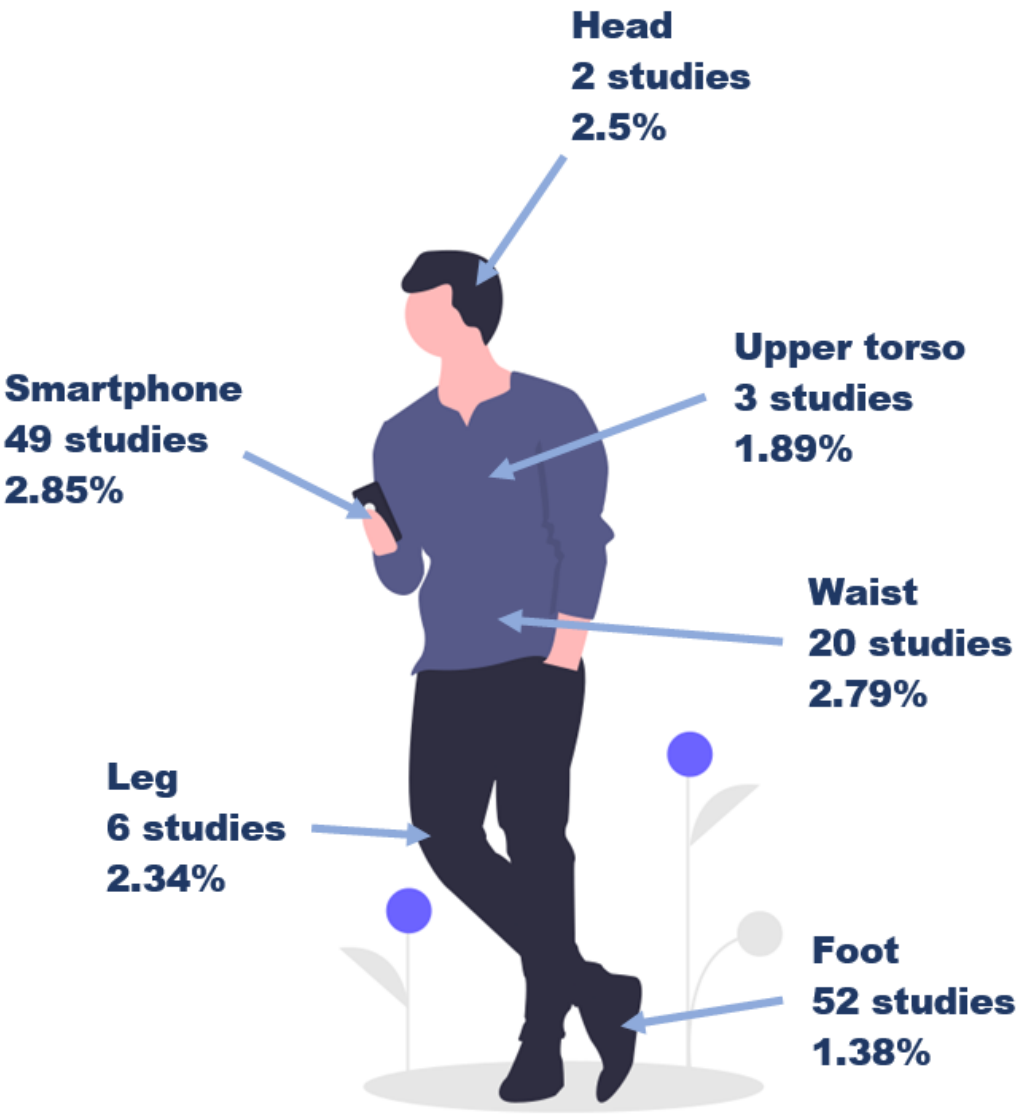


Figure 2.2: Average error of different sensor mounted positions.

signal-to-noise ratio (SNR) than an accelerometer, due to the specificity of the relevant foot movement. However, a foot-mounted system might not always be functional for prospective users. Another study [128] designed an indoor position estimation algorithm using an IMU located in a smartphone. They created a pitch-based step detection method that combined the accelerometer and gyroscope signals, with an additional sensor fusion heading estimation based on the gyroscope and magnetometer. In the study a displacement error of 0. 61% was obtained, which shows that it could outperform a conventional method in terms of error by a factor of 2. This described method is easy to implement and could be used in real- time navigation. However, the results might not generalise well, as it was only tested once on a single subject. The described experiment was carried out indoors under ideal conditions that required the subject to walk on an even floor. It remains unclear how these results translate to surroundings that are less ideal. One study that did perform tests under more diverse scenarios showed some promising results. In [134], Qiu, et al. proposed a multi-sensor fusion PDR method using EKF, zero velocity updates (ZVU) and a clustering algorithm for stance phases distinction. This study completed the most diverse set of experiments out of all papers that were included in this review. It reached rather competitive errors of less than 1%, despite the fact that the method was tested under more challenging conditions than most other studies. Indoor and outdoor scenarios were tested and included walking on level ground, as well as climbing stairs. This outcome provided some insights into the robustness of this method.

Interestingly enough, among all 145 papers, only one of them had ethical approval in place. This poses a serious question about how some of these data was collected and stored. It is strongly recommended that reviewers and editors of journals only accept human-based studies that have obtained the relevant ethical approval.

The summarised results show that in all papers a relatively high level of precision was reached (1%). However, it is hard to say whether this result had external validity. Real-world scenarios are very different from most of the testing conditions reported

in the articles and it is very probable that the error significantly increases when these systems move outside the lab. People may run, jump, turn, slide, or stumble in daily activities, and the ground we walk on could be uneven, rugged, or inclined. Position tracking in these environments is much more complex than walking on a smooth surface at a constant pace. Pushing system testing to include more challenging conditions will help create the next generation of PDR systems.

2.6 Conclusion

This chapter concluded the current research status on human tracking with wearable sensors in the aspects of sensor type, placement, algorithm, and evaluation.

IMU is the most commonly used sensor due to its self-contained property and wide adoption in electrical devices. It is so far the best option for low-cost, unobtrusive, discreet, everyday, anywhere wearable human tracking solution. In this thesis, the IMU has been adopted for the following studies.

The foot and smartphone were concluded to be the most utilised sensor positions for human tracking, since the movements of feet are directly related to steps, the swing and stance phases of feet provide perfect reference for calibrations, and the smartphone is widely used. However, there does not exist a widely used electric device on the feet, so extra installations and periodically replacements of a special device on the feet or shoes are needed, which brings unnecessary inconvenience to potential users. However, smartphones are hard to carry during sports or do not need to be carried with people all the time in environments such as home or office. The head-mounted sensor is a better choice for continuous daily tracking, as there are already many smart objects integrated with IMUs like smart glasses, smart mouthguards, earphones, etc. More detailed reasons are summarised in the next chapter. In this thesis, human tracking with head-mounted sensors was studied.

Chapter 3

A pedestrian dead reckoning method for head-mounted sensors

In the last chapter, it was found that most previous PDR methods adopted foot-mounted sensors. However, humans have evolved to keep the head steady in space when the body is moving in order to stabilise the visual field. This indicates that sensors placed on the head might provide a more suitable alternative for real-world tracking. Emerging wearable technologies that connect to the head also make this a growing field of interest. Head-mounted equipments, such as glasses, are already ubiquitous in everyday life. Other wearable gear, such as helmets, masks, or mouthguards, are becoming increasingly common. Therefore, an accurate PDR method specifically designed for head-mounted sensors is needed. It could have various applications in sports, emergency rescue, smart home, etc. In this chapter, a new PDR method for head mounted sensors is introduced, compared to two established methods. Data were collected using sensors placed on glasses and embedded into a mouthguard. The results show that the newly proposed method outperforms the other two techniques in terms of accuracy, with the new method producing an average end-to-end error of $0.88m$ and a total distance error of 2.10%.

3.1 Introduction

As mentioned, in the last chapter, the majority of PDR papers with wearable devices apply sensors on the foot, which is closely followed by publications that use

smartphone-based sensors. The waist is the third most researched option, after which the leg and upper torso seem to follow. Only 2 published papers identified looked at PDR for head-mounted sensing systems. Most studies have adopted foot-mounted sensors, as this location makes it easier to detect specific gait features that reoccur during walking. This gait information can be utilised by subsequently applying ZUPT or ZARU techniques. ZUPT and ZARU allow for a reduction in the long-term accumulation of errors. However, when potential users within a healthcare setting were asked where they would like to wear sensor technologies, placement on the foot was rarely mentioned (only 2% of the time this location was mentioned) [9]. It was also shown that small, discreet, and unobtrusive systems were preferred with many people referring back to everyday objects. The head provides an interesting location to attach sensors to, as there are several everyday worn objects that could easily be adapted for monitoring purposes. Potential objects that the sensor system could be integrated into are glasses, mouthguards, face masks, helmets, earrings, earphones, hearing aids, or even caps. Glasses might be more acceptable for monitoring in an everyday environment, whilst within the (contact) sports community the applicability of smart mouthguards or helmets might be more appropriate.

Glasses are already a requisite to people who need to correct for certain visual impairments, and it is estimated that there are 1,406 million people with near-sightedness globally (22.9% of the world population). This number is predicted to increase to 4,758 million (49.8% of the world population) by 2050[67]. Besides vision correction, people also wear glasses to protect themselves from ultraviolet or blue light or just to accessorise. There are already several smart glasses products on the market, such as Google Glasses, Vuzix Blade, Epson Moverio BT-300, Solos, and Everysight Raptor. The field of smart glasses also links well with the growing interest in providing tracking in virtual reality (VR) environments without the need of any other technology, accepting what is integrated into the VR headset.

Protective gear is often located on the head. Helmets, masks, and mouthguards are some of the obvious devices that are used both in industry and sports. It was

estimated that 40 million mouthguards are sold in the United States each year[15]. With increasing participation in contact sports, the consumption of mouthguards will continue to increase. Smart mouthguards are already being equipped with sensors for a variety of purposes [32] [68], which show the feasibility of sensor-embedded mouthguards. For contact sports, watches or other wearable objects are not allowed, as they could cause injuries and therefore form a risk to players. However, the mouthguard is often required to be worn to protect the teeth, which makes it a very suitable piece of equipment for sensor integration. This integration would provide an unobtrusive, head-mounted sensor system that can be worn on-field. Smart mouthguards may be the future of the mouthguard industry, as they combine physical protection with relevant information for the sport community [30].

Moreover, in humans, the stability of the visual field is essential for efficient motor control. The ability to keep the head steady in space allows control of movement during locomotor tasks [129]. Therefore, placing sensors on the head would provide a very suitable location, as the whole body is working on stabilising that particular segment of the system. The head is also where the vestibular system is located, which acts as an inertial guidance system in vertebrates. Placing the artificial positional tracking system near the biological one seems a suitable approach.

For head-mounted solutions, Hasan and Mishuk proposed an IMU sensor fusion algorithm, which uses data collected by a 3-axis accelerometer and a 3-axis gyroscope embedded in smart glasses [59]. The glasses were worn by a subject and the steps were detected by applying a peak detection of the accelerometer norm. The linear model in 1.8 is used to obtain a step length estimation. An EKF is used to then fuse the accelerometer and gyroscope data to determine the heading direction. The EKF is the non-linear version of the KF, where the state transition and observation matrices are Jacobian matrices of non-linear measurement model equations. In this comparing study, the state vector x_t, the measurement vector y_t, and the estimated measurement h are shown in Equation 3.1, where θ represents the Euler angle ($^\circ$), ω is the angular rate ($^\circ/s$), and a is acceleration (ms^{-2}) on each axis.

$$x_t = \begin{bmatrix} \theta_x \\ \theta_y \\ \theta_z \\ \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} ; y_t = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \\ a_x \\ a_y \\ a_z \end{bmatrix} ; h(x_t, 0) = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \\ \sin\theta_z \sin\theta_x \\ \cos\theta_z \sin\theta_x \\ \cos\theta_x \end{bmatrix} \tag{3.1}$$

The result found by [59] showed that the sensor fusion positioning technique achieved an average error between the estimated and real position of 2 m in the most complex case, when the monitoring time frame was less than 2 min.

Zhu et al.[192] utilised a hybrid step length model and a new azimuth estimation method for PDR. Steps were detected by the peak detection algorithm or the positive zero crossing detection algorithm. The hybrid step length model is given in Equation 3.2,

$$l = (a \cdot (r_{lf} \cdot f + r_{lv} \cdot v) + b) \cdot \sqrt[4]{a_{max} - a_{min}} \tag{3.2}$$

where a and b are coefficients, f is the walking frequency (Hz), v is the accelerations' variance. a_{max} and a_{min} (ms^{-2}) are the maximum and minimum acceleration in one step, r_{lf} is the Pearson correlation coefficient between step length and walking frequency, and r_{lv} is the Pearson correlation coefficient between the step length and acceleration variance.

$$r_{lf} = \frac{\sum_{i=1}^{N}(l_i - \bar{l})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^{N}(l_i - \bar{l})^2} \cdot \sqrt{\sum_{i=1}^{N}(f_i - \bar{f})^2}} \tag{3.3}$$

$$r_{lv} = \frac{\sum_{i=1}^{N}(l_i - \bar{l})(v_i - \bar{v})}{\sqrt{\sum_{i=1}^{N}(l_i - \bar{l})^2} \cdot \sqrt{\sum_{i=1}^{N}(v_i - \bar{v})^2}} \tag{3.4}$$

The heading information was generated from the hybrid value of the azimuth estimated by gyroscopes and magnetometers:

$$\varphi_{hyb} = k_{gyr} * \varphi_{gyr} + k_{mag} * \varphi_{mag} \tag{3.5}$$

where φ_{gyr} and φ_{mag} ($^\circ$) are the heading angles estimated by gyroscope and magnetometer, k_{gyr} and k_{mag} are coefficients, and φ_{hyb} ($^\circ$) is the hybrid heading estimation. In their study, they found a maximum error of 1.44m and a mean error of 0.62m.

We propose exploring the performance of these two PDR methods (which are specifically for head-mounted sensors) [59][192] and a novel algorithm based on Peak detection, Mahony algorithm and Weinberg step length model. These three methods will be compared, and the average end-to-end error, as well as total distance error, will be recorded. The proposed PDR method has a higher accuracy than the two currently available methods. The performance is consistent across different sensors and placements, which verifies its robustness. This method provides a suitable approach across various hardware platforms and thus supports the potential creation of truly innovative smart head-mounted systems in the long term.

3.2 Methods

The proposed techniques belong to the field of SHSs. SHSs output a series of step vectors by detecting each step of the pedestrian, estimating the length and direction of it, and finally integrating every step to obtain a complete trajectory. The next position of the pedestrian could then be estimated in Equation3.6 when the current position after k_{th} step (x_k, y_k) is known, where l_{k+1} (in m) and φ_{k+1} (in $^\circ$) represent the step length and the forward direction of the next step.

$$\begin{aligned} x_{k+1} &= x_k + l_{k+1} * \sin(\varphi_{k+1}) \\ y_{k+1} &= y_k + l_{k+1} * \cos(\varphi_{k+1}) \end{aligned} \tag{3.6}$$

3.2.1 Step detection

Accurate step detection is a primary requirement for precise position estimation with PDR. In this study, a peak detection method was used to detect a step at a heel strike. First, the norm of the accelerometer signal was calculated by Equation3.7, where A_x, A_y, A_z (ms^{-2}) represent accelerations on three axes.

$$Accel_{norm} = \sqrt{A_x^2 + A_y^2 + A_z^2} \tag{3.7}$$

The accelerometer norm was then filtered by a first-order low pass filter (LPF) with a cut-off frequency set to $2Hz$ to eliminate any high-frequency noise, which

should be sufficient to capture the walking motion of most users. The filtered signal peaks that cross the minimum threshold are detected as a step. The period between two adjacent peaks represents the process of the centre of gravity moving from the lowest point to the highest point and back again, which corresponds to a single step when walking. Therefore, peak detection method can be applied to detect each step (see Figure 3.1).

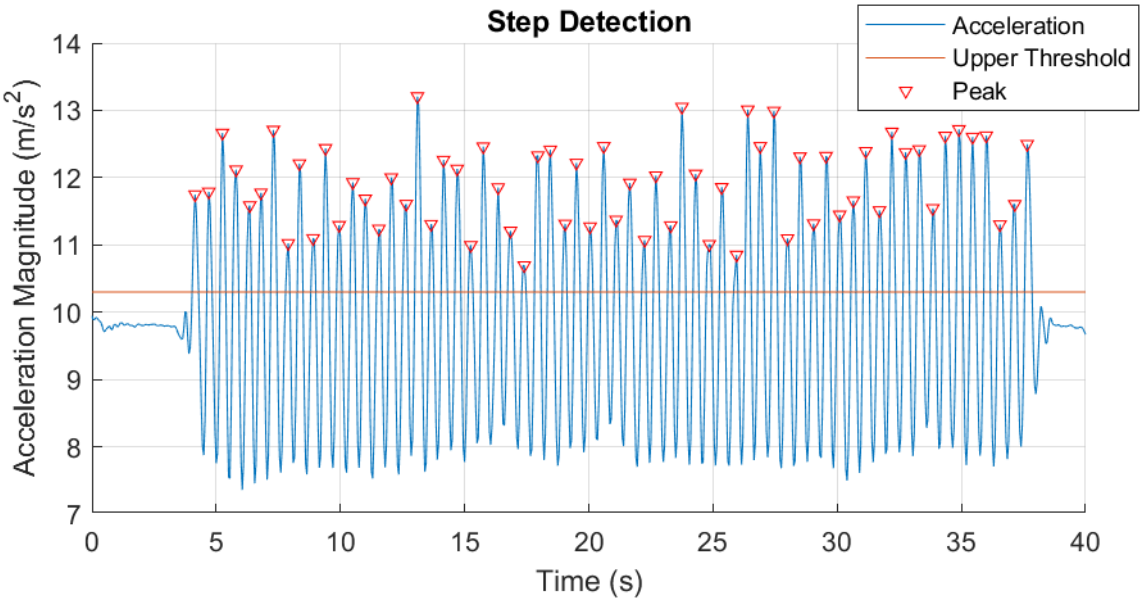


Figure 3.1: Step detection result.

3.2.2 Step length estimation

The most commonly used models consist of the linear model [103], Weinberg model [166], Kim model [91], Scarlett model [143] and Shin model [149]. These methods were tested on single-user gait data (walking in a straight path). The preliminary results in Table 3.1 show that Weinberg, Kim, and Shin models have the best performance with head-mounted sensors in this study. Considering the speed of calculation, Weinberg model was finally chosen as the step length estimator in this study. The application of a single step length estimator across the PDR algorithms allowed for a more unbiased assessment of the performance within this study.

Table 3.1: Comparison of step length estimation algorithms. (Real distance: 11.28m.)

Methods	**Linear**	**Weinberg**	**Kim**	**Scarlett**	**Shin**
Estimated distance (m)	8.86	11.11	11.23	10.08	11.12
Error (m)	2.41	0.16	0.05	1.20	0.15
Error rate	21.42%	1.48%	0.45%	10.67%	1.41%

3.2.3 Heading estimation

Mahony et al.[113] proposed an explicit complementary filter, which considers the problem of obtaining good attitude estimates from measurements obtained from typical low-cost inertial measurement units. This algorithm requires only accelerometer and gyro outputs to estimate heading direction. In this study a quaternion-based derivation of the explicit complementary filter was adopted as the heading estimator.

The explicit complementary filter in terms of unit quaternion representation can be expressed in Equation3.8, where Ω_y is the biased measure of angular velocity by gyroscope, k_P is the proportional gain and k_I is the integral gain, $\mathbf{p}(\Omega) = (0, \Omega)$ is a pure quaternion, $\hat{b}$ denotes gyro bias, ω_{mes} is a correction term and can be thought of as a non-linear approximation of the error.

$$\begin{aligned} \omega_{\text{mes}} &= -\text{vex}\left(\sum_{i=1}^{n} \frac{k_i}{2}\left(v_i \hat{v_i}^T - \hat{v}_i v_i^T\right)\right) \\ \dot{\hat{q}} &= \frac{1}{2}\hat{q} \otimes \mathbf{p}(\Omega_y - \hat{b} + k_P \omega_{\text{mes}}) \\ \dot{\hat{b}} &= -k_I \omega_{\text{mes}} \end{aligned} \tag{3.8}$$

When implementing it, the accelerometer data (showing the direction of gravity) is used as a reference. First, the estimated direction of gravity is calculated by the unit quaternion. The orthogonal matrix corresponding to a rotation by the unit quaternion $\mathbf{q} = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ (with $|\mathbf{q}| = 1$), when post-multiplying with a column vector, is given by:

$$R = \begin{pmatrix} a^2 + b^2 - c^2 - d^2 & 2bc - 2ad & 2bd + 2ac \\ 2bc + 2ad & a^2 - b^2 + c^2 - d^2 & 2cd - 2ab \\ 2bd - 2ac & 2cd + 2ab & a^2 - b^2 - c^2 + d^2 \end{pmatrix} \tag{3.9}$$

Then the estimated gravity direction is:

$$\begin{aligned}\hat{\mathbf{v}}_i &= \hat{R}^T \begin{bmatrix}0 & 0 & 1\end{bmatrix}^T \\ &= \begin{bmatrix}2bd - 2ac & 2cd + 2ab & a^2 - b^2 - c^2 + d^2\end{bmatrix}^T\end{aligned} \tag{3.10}$$

The measured gravity direction is:

$$\mathbf{v}_i = \frac{\mathbf{a}}{|\mathbf{a}|} \tag{3.11}$$

The error is given as the cross product between estimated gravity direction and measured direction:

$$\mathbf{e}_i = \mathbf{v}_i \times \hat{\mathbf{v}}_i \tag{3.12}$$

In this sample period, the quaternion rate of change is:

$$\dot{\hat{q}}_i = \frac{1}{2}\hat{q}_{i-1} \otimes \mathbf{p}(\Omega_i + k_P \cdot e_i + k_I \cdot \sum_{i=1}^{N} e_i) \tag{3.13}$$

The estimated quaternion in this sample period is:

$$\hat{q}'_i = \hat{q}_{i-1} + \dot{\hat{q}}_i \cdot t \tag{3.14}$$

where t is the length of sample period. Finally, the normalised quaternion will be outputted:

$$\hat{q}_i = \frac{\hat{q}'_i}{|\hat{q}'_i|} = \begin{bmatrix}q_0 & q_1 & q_2 & q_3\end{bmatrix}^T \tag{3.15}$$

After calculating this for each sample period, a series of estimated quaternions is generated. To obtain the heading angle, each quaternion is subsequent transformed into an Euler angle:

$$\begin{bmatrix}\phi \\ \theta \\ \psi\end{bmatrix} = \begin{bmatrix}\text{atan2}(2(q_0q_1 + q_2q_3), 1 - 2(q_1^2 + q_2^2)) \\ \text{asin}(2(q_0q_2 - q_1q_3)) \\ \text{atan2}(2(q_0q_3 + q_1q_2), 1 - 2(q_2^2 + q_3^2))\end{bmatrix} \tag{3.16}$$

The Mahony algorithm can also fuse magnetometer data into the calculation. However, because of the unpredictable magnetic interference and noise in the low-cost sensors, our proposed method did not include any magnetometer data in the final estimation.

3.3 Experimental conditions and results

3.3.1 Hardware description

The sensor that was placed on the glasses and the mouthguard is the SensorTile Microcontroller Unit (MCU) module (STEVAL-STLCS01V1), which includes a low-power 3D accelerometer and 3D gyroscope (LSM6DSM), ultra-low power 3D magnetometer (LSM303AGR), Bluetooth low energy network processor (BlueNRG-MS), 32-bit ultra-low-power MCU with Cortex®M4F (STM32L476JG), and a 100 mAh lithium-Ion polymer battery. Data were collected at 20Hz and transferred to a mobile phone by Bluetooth.

The MCU was attached to the glasses and embedded into a mouthguard. The mouthguard was made out of two Ethylene-vinyl acetate (EVA) layers with 0.6 mm thickness, which were thermoformed using the upper teeth cast of the subject. The MCU was embedded between the EVA sheets and placed near the hard palate of the oral cavity (Figure 3.2a). The MCU was also attached to the legs of a pair of glasses with a sensor located near the left temple (Figure 3.2b). Both placements provided a rigid arrangement that ensured that the sensors precisely followed the head motions.

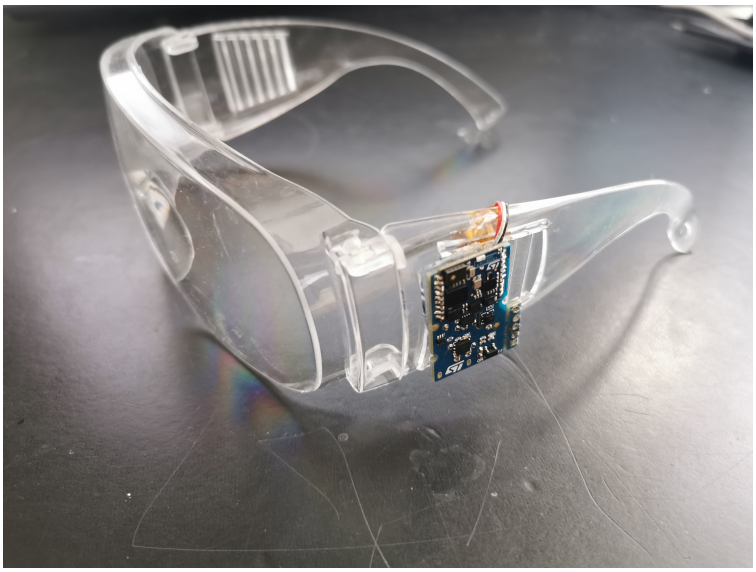
3.3.2 Data collection

Three healthy adults voluntarily participated in this study, including one female volunteer and two male volunteers. In [69], the majority of PDR studies (66.2%) had only one participant, which made the results somewhat limited. This study had 3 subjects and data were collected 4 times with each device.

The tests were carried out on a hard outdoor surface on a clear day. The red dotted line shown in Figure 3.3 is the walking trajectory used in the tests. The set path formed a rectangle with an approximate length of 25.5m and a width of 8.5m. The satellite image shown in Figure 3.3 is from Google Map.

Subjects were asked to face the walking direction and stand still for 5 seconds at the beginning of each data collection trial. Then they were asked to walk at a normal

(a) Smart Mouthguard



(b) Glasses with MCU

Figure 3.2: Devices used in this study.

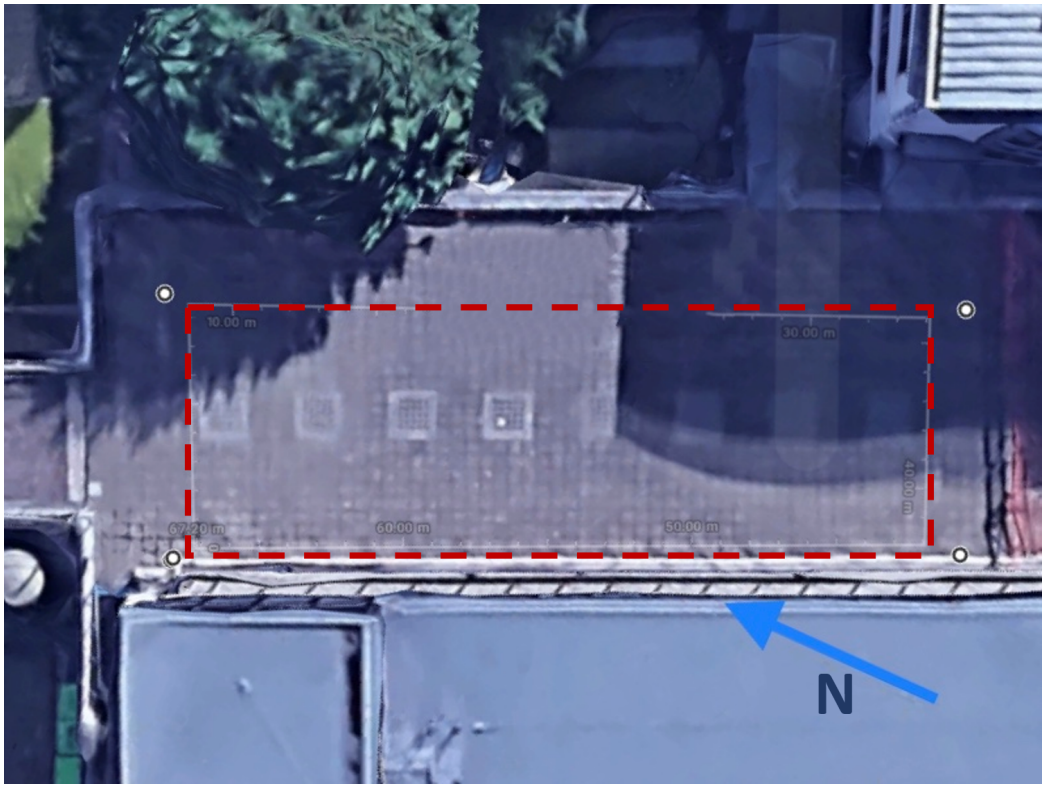


Figure 3.3: Top view of the data collection trajectory that was set out for each subject. The red dotted line shows the trajectory subjects were asked to walk.

and constant speed while keeping their head facing the walking direction. When the subject reached the end of the trajectory, they were asked to stop walking and stand still again for another 5 seconds. Specific information on walking speed was provided, as it is known that small changes in the instructions given could change the result of the test [8]. Ethical approval was obtained from the University and this experiment was part of a larger study (R43470/RE001).

All data analysis was conducted in MATLAB (R2020a, Mathworks, United States).

3.3.3 Results

Figures in Figure 3.4 show the estimated trajectories with glasses and mouthguard by Hasan's, Zhu's and the newly proposed algorithm. The proposed method generated trajectories that are closer to the ground-truth trajectory compared to the other two methods.

The end-to-end error represents the distance between the start point and the end point of the estimated trajectory. The total distance error is the absolute difference between the real and estimated trajectory lengths. The mean end-to-end error of Hasan's method is 2.62m, and Zhu reached 5.41m, while the new method reached a mean end-to-end error of 0.88m. Figure 3.5a shows that Zhu's method has a larger variance of the error with different sensor layouts, while the proposed method has the smallest mean error and the smallest variance for both sensor locations.

The novel method has a mean total walking distance error of 1.43m, which is 2.10% of the total length. However, Hasan's and Zhu's methods have a relatively larger error rate of 25.97% and 36.36%, respectively. Total distance errors are shown in Figure 3.5b.

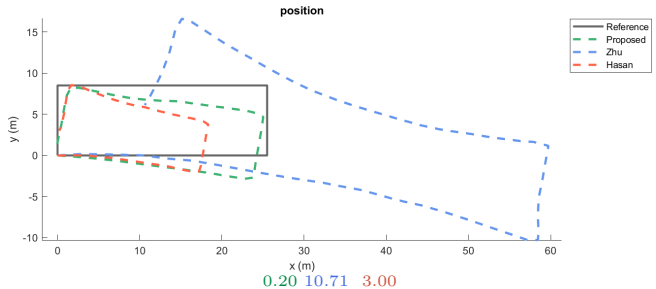
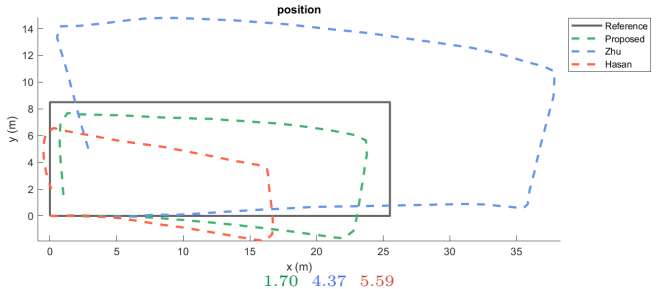
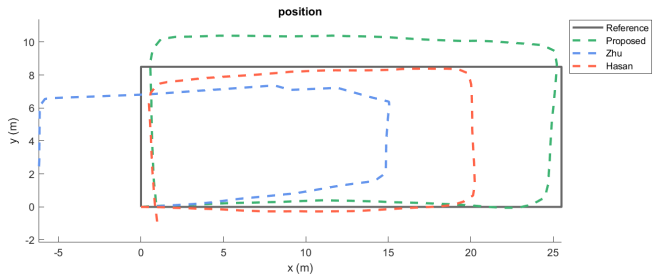
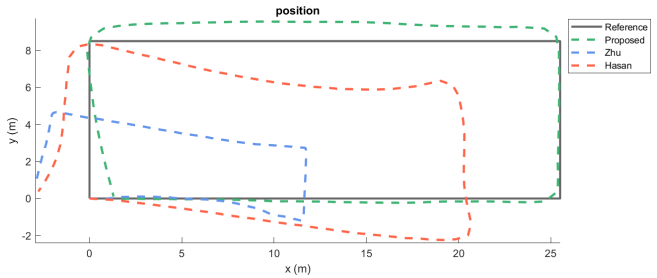
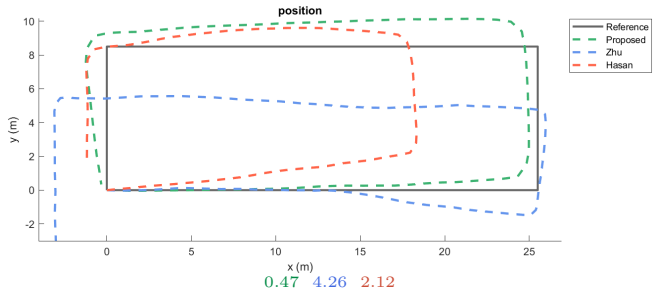
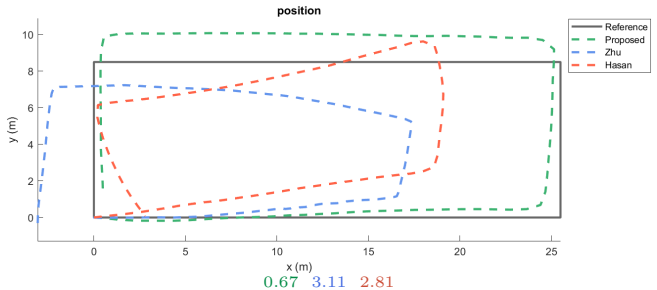
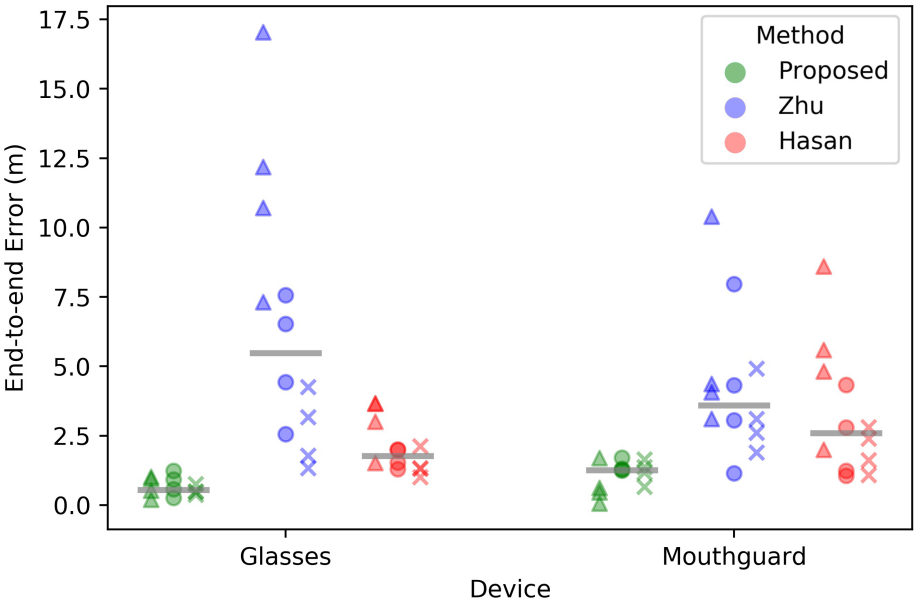
	Glasses	Mouthguard
Subject 1		
End-to-end error (m)	0.20 10.71 3.00	1.70 4.37 5.59
Total distance error (m)	1.90 62.49 21.60	1.16 25.83 24.99
Subject 2		
End-to-end error (m)	0.92 6.53 1.54	1.31 3.06 2.80
Total distance error (m)	0.15 22.92 12.53	0.23 34.63 10.08
Subject 3		
End-to-end error (m)	0.47 4.26 2.12	0.67 3.11 2.81
Total distance error (m)	0.32 0.70 17.73	1.37 19.81 16.94

Figure 3.4: Estimated trajectories for the three methods plotted against the ground-truth. One of the measurements (randomly selected) of each subject is shown for each placement. The end-to-end error and total distance error (m) are shown in the sequence of proposed, Zhu's, and Hasan's algorithm.

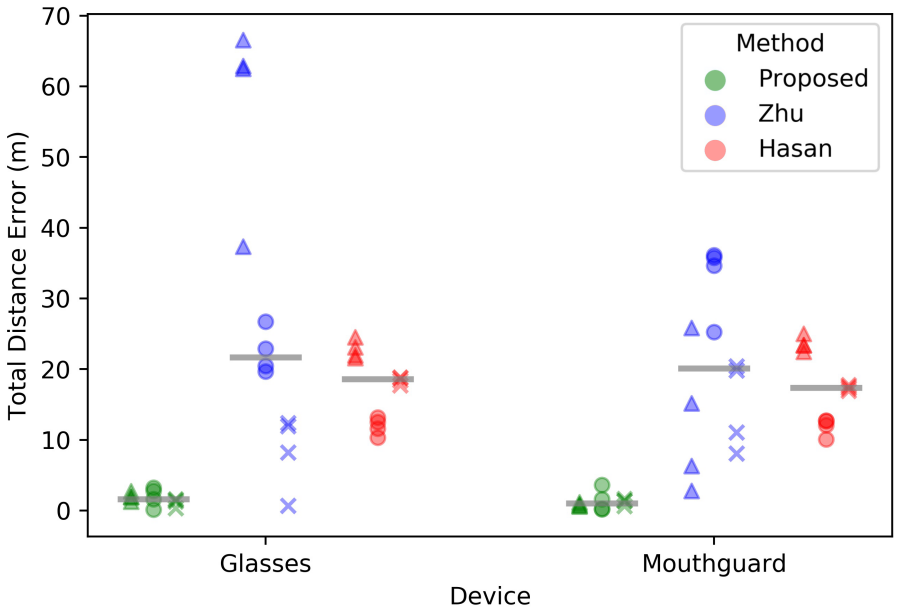
3.4 Discussion

The results show that the end-to-end and total distance error was the lowest in terms of central tendency for the newly proposed method. Extra tests were carried out with a different set of sensor platform (smartphone) to provide a further preliminary cross-platform comparison. Data were collected using a Huawei P30 smartphone, which consists of a 3-axis accelerometer and gyroscope (ICM-20690) produced by TDK InvenSense and a 3-axis magnetometer (AK09918) from Asahi Kasei Microdevices. The smartphone was placed on the cranial part of the head and was held securely in place using adaptable straps. Only one subject was tested under this head-mounted smartphone condition. The three methods all performed comparatively well for heading estimation under smartphone conditions (see Figure 3.6), likely due to the higher accuracy of the IMU in the smartphone compared to the MCU. However, when a cheaper IMU is used (such as the one attached to the glasses and mouthguard), the results of Hasan's method and Zhu's method quickly perform with greater errors (Figure 3.4). The newly proposed method in this article seems rather robust (in terms of errors) even when the sensor platform changes, indicating that a more broader inclusion of hardware is possible and that the system is likely to perform relatively stable even under less optimum operational conditions.

For the heading estimation, the error of Zhu's method was heavily influenced by sensor type and layout. The main reason for the lack of robustness in this algorithm is probably due to the fact that the noise from the sensors is not well compensated for. The hybrid heading estimation is simply the proportional sum of the raw integration of gyroscope and magnetic direction, with the proportional gain set manually. Although it sets a threshold for the gyroscope to start integration only when the Euclidean norm of the gyroscope data is above this threshold, the bias in gyroscopic data still remains and thus the integration error remains. Heading estimation from the magnetometer is also still a problem, because the geomagnetic field is not always uniform [19] and it can easily be disturbed by hard or soft magnetic interferences. In addition, the



(a) End-to-end error



(b) Total distance error

Figure 3.5: Errors across all subjects for each condition. Horizontal lines represent median values. A triangle is used to represent data from subject 1, a circle is given for subject 2 and data from subject 3 is shown as a cross.

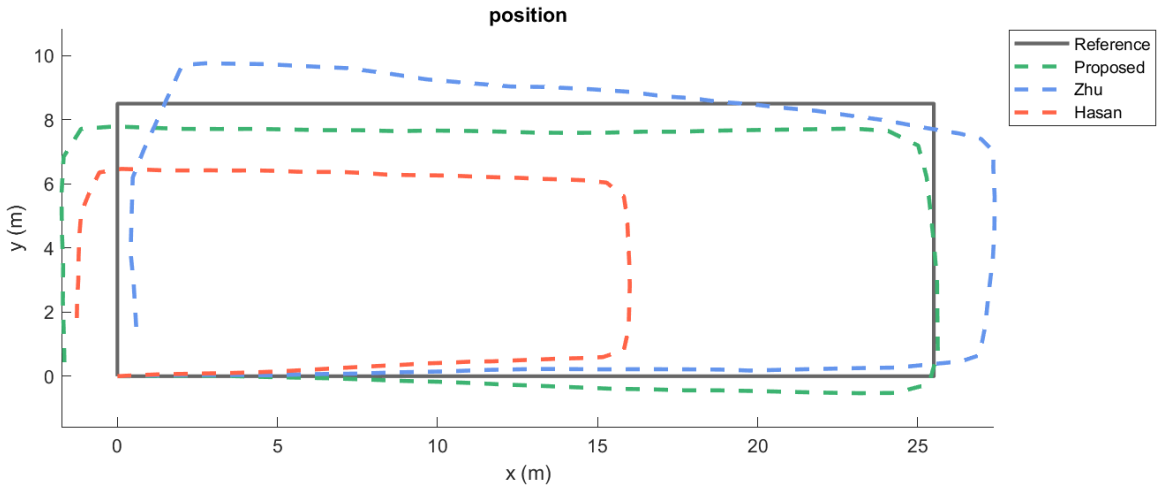


Figure 3.6: Estimated trajectories for the three methods plotted against the ground-truth for a head-mounted smart phone. The measurement is from one subject.

oscillation during walking creates non-negligible high-frequency noise in the heading estimation.

Regarding the step length estimation, Hasan's and Zhu's methods yield larger errors for a range of reasons. Hasan's algorithm used a linear model with the coefficient values adopted from another study, which obtained parameters from 4000 steps of 23 different people[103]. The linear regression based on these data would provide a general model, and this model does not necessarily scale well across individuals. Especially if the conditions, subjects, or instructions differ from the original database. Generating specific parameters for each individual is therefore a key factor in obtaining more accurate step length estimates. As for Zhu's method, the step length estimation model combines information of step frequency, acceleration variance, as well as maximum and minimum accelerations in one step. This is a more scalable solution that should yield more accurate estimates. Nevertheless, this method only performs well when the conditions and sensor placement are exactly the same between the step length estimation session and the subsequent tracking sessions. If changes occur between these sessions then the step length estimate might no longer be accurate enough under the new conditions. Even when the temperature or sensor changes, acceleration variance, maximum, and minimum acceleration can start to

vary, because of the changing noise or the different sensor properties[107]. This can lead to contrasting estimations even when all other conditions remain constant.

The attitude representation in Hasan's method and in the proposed method of this study is notably different. The former uses Euler angles in the EKF, which consists of three rotation angles around three axes. The latter adopted quaternions within the Mahony algorithm. Euler angles are more human understandable but have several disadvantages. These include the ambiguity in the rotation sequence of axes and the possible occurrence of a gimbal lock, which leads to the loss of one degree of freedom. In contrast, expressing rotations as unit quaternions has some advantages, such as: concatenating rotations is computationally faster and numerically more stable; extracting the angle and axis of rotation is simpler; interpolation is more straightforward and quaternions do not suffer from gimbal lock as Euler angles do [34]. Although the EKF with Euler angle in Hasan's method performs well, it would be more stable if quaternions were used.

Hasan's method and the newly proposed method showed high accuracy in heading estimations, proving the effectiveness of EKF and Mahony's algorithm in the attitude estimation. This is reflected in the fact that they are already one of the most popular algorithms in this area. In EKF, usually there are two parameters that need to be set and tuned, which are the process noise covariance matrix Q and measurement noise covariance matrix R [110]. In Mahony's algorithm, there are also two parameters that need to be adjusted: proportional gain k_P and integral gain k_I. In both of these algorithms, the parameters are found by tuning until the best results are achieved. This is, of course, time consuming and there is no guarantee of optimality. More importantly, when the sensor position changes, for example, from glasses to the mouthguard, the parameters need retuning. More research on automatic parameter tuning will be useful, as it can help with the creation of the next generation of algorithms.

The processing of the data recorded in this study was done off-line. However, many applications will need online PDR to provide real-time navigation or localisa-

tion. PDR algorithms should be assessed for online implementation feasibility, and preliminary estimates of running time provide some insight into this. The running time of each method is shown in Table 3.2. Hasan's method takes much more time than the other two methods because the Jacobian calculation in EKF is complex and time-consuming. The mean running time of one sample is 0.0661s, which indicates that the sampling frequency will be lower than 15Hz. Empirical run-time estimates are limited as algorithms are platform-independent and for that reason more theoretical assessment needs to take place to compare the complexity of these algorithms, which is beyond the scope of this chapter.

Table 3.2: Running time of methods.

Methods	**Hasan**	**Zhu**	**Proposed**
Number of samples	701	701	701
Total time (s)	46.34	0.23	0.95
Mean time (s)	0.0661	0.0003	0.0014

Subjects were instructed to follow the reference path shown in Figures 4 and 6. However, it is likely that there is some deviation from the ideal path, especially during turning, as subjects were asked to walk in their own preferred manner. However, all subjects started and stopped always at the same location. Caution needs to be taken regarding the generalisability of the results. Different environmental conditions, such as pathways taken or surrounding objects, can yield different outcomes. In this study, a simple walking pattern (rectangle) was selected to minimise any variability in the trial between subjects and sensor placement. It should be noted that most studies use only a single person in their testing [69] and that this study included multiple subjects. It will be useful for the field to further expand the number of subjects, as well as environmental conditions (e.g. walking surface) during experimental testing to create ever increasing levels of external validity.

The method presented in this chapter was the algorithm that performed the best in all locations of the sensors. This new technique offers an approach that can be generalised across sensors and smart head-mounted equipment, such as helmets, hats,

glasses, mouthguard, dentures, or earphones. The ability to have a high-performance general algorithm for head-based systems can provide new opportunities for a range of industries, ranging from healthcare to entertainment.

The data collected by mouthguard was expected to have higher accuracy because it is closer to the longitudinal axis of the body. But the results show that glasses got lower values in both end-to-end error and total distance error. The temperature affect was the first factor considered. The normal oral temperature is 37 °C which has a 12°C difference to the normal testing environment temperature 25°C. But according to the IMU hardware datasheet, sensitivity change of the linear acceleration vs. temperature is $\pm 0.01\%/^{\circ}C$, of the angular rate is $\pm 0.007\%/^{\circ}C$; the linear acceleration zero-g level change vs. temperature is $\pm 0.1mg/^{\circ}C$; the angular rate typical zero-rate level change vs. temperature is $\pm 0.015dps/^{\circ}C$. They all change very little over temperature and time. Another potential reason is that the board in the mouthguard is not rigidly fitted inside. There is extra space in the cavity of the mouthguard that can cause board movements inside during walking with unconscious mouth or tongue muscle movements. More importantly, the size of the dataset is limited, which may not be representative of the overall population, leading to bias in the estimation.

The current experiment did not collect synchronised data from the different locations. Thus, any assessment of the relationship between these locations needs to take into account, as well as the subject variability between trials. Nonetheless, the task at hand was relatively simple to complete, and variability is likely to be limited within a subject. Information about how strongly these placements are correlated allows for a further insight in terms of potential reciprocal nature of these systems. Spearman's rank correlation coefficient (rs) was calculated to assess this, as a Pearson Correlation is susceptible to outliners in small sample sizes. Coefficients were determined between the glasses and mouthguard condition for the proposed (rs= 0.4261, p= 0.0390), Zhu's (rs= 0.6435, p=0.0009) and Hasan's (rs= 0.9920, p= 3.0338e-21) method. It showed a strong correlation across all methods between the locations. This indicates that these systems can act as potentially interchangeable. More work

is needed to further confirm these preliminary findings regarding the association between different head-mounted sensor positions.

3.5 Summary

In this chapter, a PDR method for head-mounted equipment was proposed. Peak detection was used to detect steps and heading estimation was carried out with the Mahony algorithm. The Weinberg model was implemented to obtain step length estimation. This proposed method was executed in MATLAB with another two methods for performance comparison. To evaluate the accuracy of three methods, we collected walking data on a rectangle trajectory using low-cost IMUs on glasses and in a mouthguard, respectively. The result shows that the proposed method reached an average end-to-end error of $0.88m$ and a total walking distance error of 2.10%. A consistent performance was obtained for the novel algorithm across conditions indicating the potential robustness of the proposed method. This algorithm provides potential implementation into various hardware platforms, which can be translated into truly innovative smart head-mounted systems in the long term.

Chapter 4

HeadSLAM: pedestrian SLAM with head-mounted sensors

PDR with IMUs is one of the most promising solutions within the human position tracking domain, as it does not rely on any additional infrastructure, whilst also being suitable for use in a diverse set of scenarios. However, PDR is only accurate for a limited period of time before unbounded accumulated errors, which will affect the position estimate. Error correction can be difficult, as inertial tracking has no way to calibrate with the real world without other sensors. HeadSLAM, a method specifically designed for head-mounted IMUs, is proposed to improve the accuracy during longer tracking times (>10 mins). Research participants (n=7) were asked to walk in both indoor and outdoor environments with head mounted sensors, and the accuracy obtained with HeadSLAM was subsequently compared to that of PDR method. A significant difference (p <.001) in the average root mean square error and absolute error was found between the two methods. HeadSLAM had a consistent lower error across all scenarios and subjects in a 20-hour walking dataset. The findings of this study show how HeadSLAM algorithm can provide a more accurate long-term human tracking for low-cost head-mounted sensors. The improved performance can support affordable applications for infrastructureless navigation.

4.1 Introduction

In recent years, human position tracking technology has drastically changed modern life by offering location information for a variety of scenarios. In some scenarios, long-term stable tracking is required for continual LBS. For example, physical activity through sports participation has become an essential part of ensuring healthy living in today's world and key to this is continual tracking of workload [156] for monitoring or optimising their physical performance during sports. A robust long-term position tracking method provides a suitable method for determining external load during sport activities. Subsequently, this can be combined with objective internal load measurements to accurately provide a complete picture of an athlete's workload [41]. Tracking time during normal sports would be around an hour. Long-term positioning can also facilitate new ways of human-environment interaction, in which positional information could be leveraged to create responsive systems. For example, smart homes can use user position information to make better decisions on how to support those living in these spaces[124]. The tracking time needed for smart homes may last for hours. These responsive interactions could be particularly interesting for healthcare settings, such as care homes. For example, services can be deployed to provide assistance to the elderly by identifying their daily routines and establishing care plans that are specifically developed around the patient, instead of the other way around. It could allow for more personalised healthcare, by integration of user mobility patterns and mapping of their physical behavioural routines. The tracking time in care homes would last for a whole day.

Inertial tracking is still a competitive and user-friendly method in long-term position tracking for a lot of reasons: IMUs are energy efficient and could last for a longer time after charging (i.e., lower charging frequency); IMUs are small, discreet, portable, and easy to implement in everyday objects like glasses, mouthguard, earphones,etc., which reduce the affect on people's normal daily life to the lowest level; Inertial tracking is self-contained and could be used in any scenarios without extra

cost on infrastructures or privacy issue; IMUs are affordable, and inertial tracking does not need complex implementations, which allows it to be generalised to more users and more scenarios even in the middle-and low-income countries.

A lot of inertial tracking methods have been proposed by scientists and declared to have high accuracy, which have been summarised in Chapter 2. However, 91% of the summarised studies tested in a tracking of less than $10min$ (i.e., $<840m$). These techniques are only accurate across a "short" time period, and they become prone to drifting errors when measurement time increases. The accumulated error and the lack of a reliable re-calibration way subsequently limits the utility of this method. Thus, it is essential to find a pedestrian tracking method that has a lower error when tracking is applied over longer periods of time.

SLAM is a computational problem of constructing or updating a map of an unknown environment, while tracking the location of the object at the same time[40]. Current SLAM methods can generate accurate location tracking trajectories and environment maps by using reference landmarks to reduce errors. These landmarks can be observed with exteroceptive sensors, such as cameras[119], laser rangefinders[120], LiDAR[63] or sonar[137]. However, these extra sensors also come at an added cost, and image capturing devices face the problem of privacy or security threats when used in a private environment. Michael Angermann and Patrick Robertson previously proposed a FootSLAM algorithm, which only uses accelerometers and gyroscopes embedded in a foot-mounted IMU. FootSLAM applies a Rao-Blackwellised particle filter to build a probabilistic transition map. This approach was able to prevent unbounded error growth, and they presented two subsequent extensions of this (PlaceSLAM and FeetSLAM)[4]. Susanna Kaiser and Estefania Munoz Diaz subsequently created the PocketSLAM, which is a combination of a pocket navigation system with a Foot-SLAM method [85]. They showed that it was possible to reduce error accumulation with this technique.

In the last chapter, a PDR method specially designed for head mounted IMUs (HeadPDR) was proposed, which could generate accurate results for short trajectories,

during brief recording periods (< 1 min). In this chapter, SLAM is combined with the previously proposed HeadPDR method to determine if this would yield lower position errors compared to application of the PDR method on its own. It has been tested for indoor and outdoor environments, which simulates a sports and healthcare setting. The contribution of this chapter consists of proposing a SLAM method for head-mounted sensors, which only requires low-cost portable IMU data and remains accurate over longer time periods.

4.2 Methods

HeadSLAM method consists of two consecutive stages; (i) PDR and (ii) SLAM. The PDR estimates the heading direction and the length of each step based on raw accelerometer and gyroscope data. This is then fed into the SLAM to generate a suitable map. The whole process is shown in Figure 4.1.

4.2.1 PDR

A PDR algorithm designed for head-mounted sensors was proposed in the last chapter [68] and is adopted in this study. It uses SHSs. The SHSs output a series of step vectors by detecting each step of the user and estimating the length and direction of it. This information is then integrated throughout each step to obtain a complete trajectory. The next position of the pedestrian could then be estimated with (4.1). The current position after the k_{th} step (x_k, y_k) is entered into the equation. The l_{k+1} and φ_{k+1} represent the step length in metres (m) and the forward direction of the next step given in degrees $(^\circ)$. The iterative process provides an updated estimate of the x and y position, which reflects the 2D space in which the person is moving.

$$\begin{aligned} x_{k+1} &= x_k + l_{k+1} * \sin(\varphi_{k+1}) \\ y_{k+1} &= y_k + l_{k+1} * \cos(\varphi_{k+1}) \end{aligned} \tag{4.1}$$

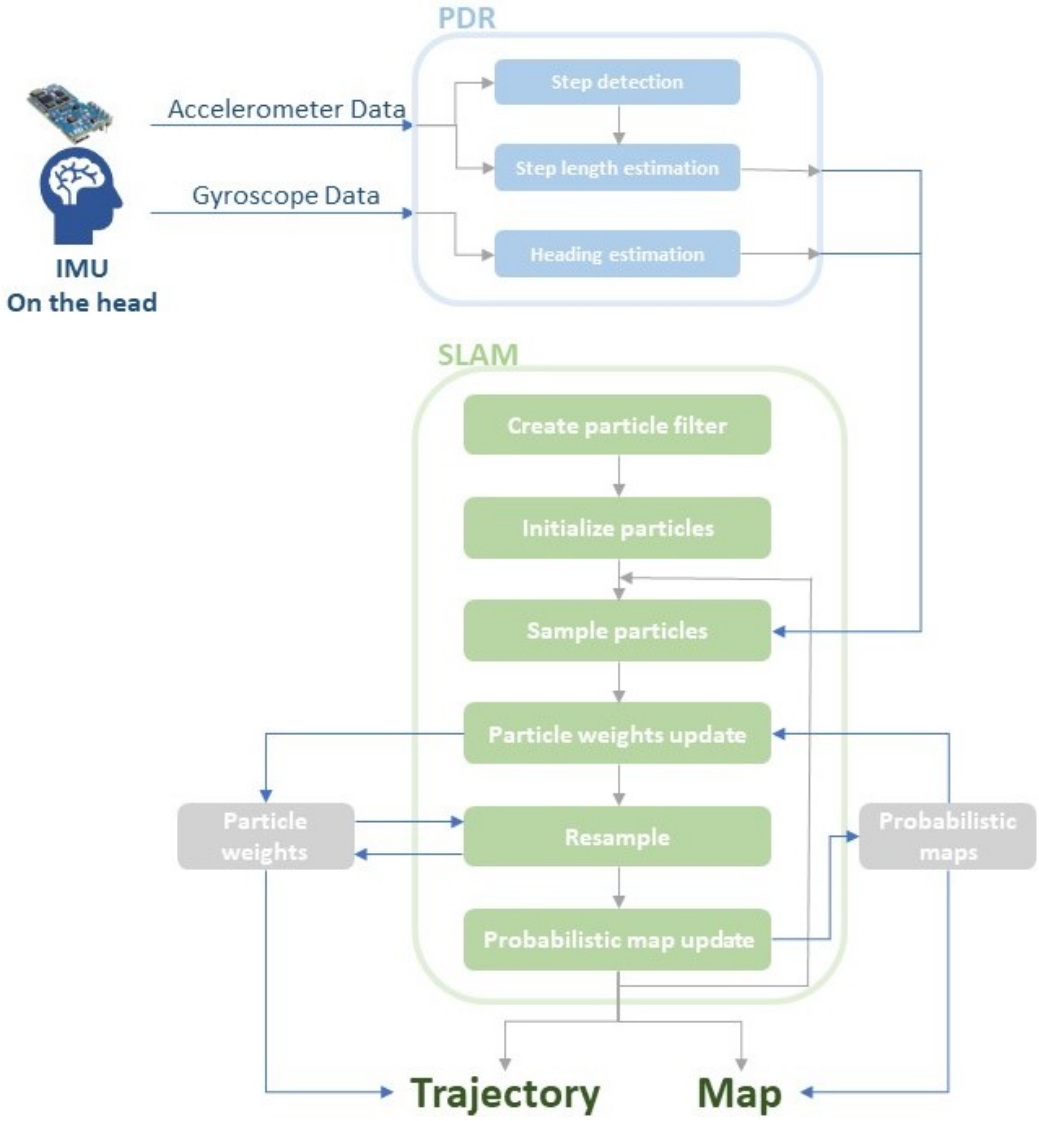


Figure 4.1: Overview of HeadSLAM method. The Simultaneous localisation and mapping (SLAM) step is performed after the Pedestrian Dead Reckoning (PDR) in order to create a map based on the data collected from head-mounted inertial measurement units. The gray arrows show the sequence of the flowchart. The dark blue arrows represents the data or variables usage or update.

4.2.1.1 Step detection

Step detection is updated in a manner similar to the method used in [68], which also adopts a peak detection approach to identify a step in a heel strike. The vertical acceleration (z-axis) is first filtered by applying a first-order low pass filter (LPF) with a cut-off frequency set to 2Hz to eliminate any high frequency noise. This is sufficient to capture the general walking motion of most users. Peaks and valleys in the filtered signal are then identified. If the value difference between a peak and its subsequent valley exceeds the predetermined threshold (0.5 m/s^2), the time interval between the previous and subsequent valley will be recognised as a step.

4.2.1.2 Step length estimation

Weinberg's model[166] was used as step estimator:

$$step\_length = k \cdot \sqrt[4]{a_{max} - a_{min}} \tag{4.2}$$

where k is a constant coefficient for unit conversion, while a_{max} and a_{min} (ms^{-2}) indicate the maximum and minimum acceleration measured in the z direction for a single step. It was proved to perform best for personalised sets of constants compared to 12 representative step length estimation models.[160]

4.2.1.3 Heading estimation

A quaternion-based derivation of the explicit complementary filter [114] was adopted as the heading estimator, which considers the problem of obtaining good attitude estimates from measurements obtained from typical low-cost inertial measurement units.

The complementary filter can also fuse the magnetometer data into the calculation. However, this study did not include any magnetometer data in the final estimate, due to unpredictable magnetic interference in indoor spaces, the noise in these low-cost sensors, and the expectation for longer battery life of the wearable devices[102, 99, 154].

4.2.2 HeadSLAM

The length and heading direction of each step (estimated by the PDR) are fed into the SLAM for calibration. The output from this is subsequently used to build the map.

In the SLAM algorithm, the two-dimensional space is first divided into a grid of adjacent hexagons of a given radius. The Rao Blackwellised Particle filter (RBPF), which is applied in the FastSLAM algorithm[118], was then used.

Rao-Blackwellised Particle filter

Rao-Blackwellisation is a statistical technique that aims to improve the efficiency of estimating certain parameters in a probabilistic model. The basic idea is to exploit the conditional independence structure of the problem to obtain a more efficient estimator for some subset of the parameters.

The Rao-Blackwell theorem states that: Suppose T is a sufficient statistic for a parameter θ, $\hat{\theta}$ is an estimator of θ, $\theta^* = \mathbb{E}(\hat{\theta}|T)$ which is the conditional expectation of θ given T. θ^* is a superior estimator in terms of mean squared error and it is the Rao-Blackwellised estimator of $\hat{\theta}$:

$$\mathbb{E}(\theta^* - \theta)^2 \leq \mathbb{E}(\hat{\theta} - \theta)^2$$

The RBPF is an extension of the basic PF that incorporates the principles of Rao-Blackwellisation to provide more efficient and accurate estimates. RBPF can overcome some of the issues associated with the curse of dimensionality by analytically marginalising over parts of the state space. RBPF decomposes the state space into two parts: one for which an analytical solution is possible (typically a lower-dimensional subset) and the other for which particles are used. The example of a decomposition equation is Equation 4.3, where the entire state space is decomposed into map estimation and pose estimation. This can lead to more accurate and computationally efficient state estimates compared to a regular particle filter.

FastSLAM

FastSLAM adopted RBPF for localisation. The differences between FastSLAM and ours are as follows.

- Environment representation: FastSLAM used a set of landmarks. HeadSLAM used grids.
- Map estimation: FastSLAM used EKF for each particle. HeadSLAM used multinomial states in grids.

HeadSLAM

The SLAM problem was decomposed into a pedestrian localisation problem and a mapping problem conditioned on the pedestrian's position (pose). The posterior can be simplified as :

$$p(\mathbf{P}_{0:k}, \mathbf{M}|\mathbf{Z}_{1:k}) = p(\mathbf{M}|\mathbf{P}_{0:k}) \cdot p(\mathbf{P}_{0:k}|\mathbf{Z}_{1:k}) \tag{4.3}$$

where $\mathbf{P}$ and $\mathbf{M}$ represent the pose and the map, $\mathbf{Z}_k$ is a noisy measurement of the difference between $\mathbf{P}_{k-1}$ and $\mathbf{P}_k$, which is the step vector estimated from the previous PDR layer. The pose could be estimated recursively:

$$p(\mathbf{P}_{0:k}|\mathbf{Z}_{1:k}) \propto p(\mathbf{Z}_k|\mathbf{P}_{k-1:k}) \cdot p(\mathbf{P}_k|\mathbf{P}_{0:k-1}) \cdot p(\mathbf{P}_{0:k-1}|\mathbf{Z}_{1:k-1}) \tag{4.4}$$

$p(\mathbf{Z}_k|\mathbf{P}_{k-1:k})$ is the likelihood function, which adopts a Gaussian distribution to draw possible poses after each step. The errors of the estimations of the step length and rotation are treated as Gaussian distributed. So in each step, the particles are sampled and updated following a Gaussian distribution with a mean of the estimated value and a variance defined by the noise in measurements and errors in PDR model. The pose transition function $p(\mathbf{P}_k|\mathbf{P}_{0:k-1})$ is computed by marginalising over the map. Integrating it yields:

$$I^i \propto \frac{N_{\tilde{h}}^{\tilde{e}} + \alpha_{\tilde{h}}^{\tilde{e}}}{N_{\tilde{h}} + \alpha_{\tilde{h}}} \tag{4.5}$$

where $N_{\tilde{h}}^{\tilde{e}}$ is the number of times the i-th particle crossed edge $\tilde{e}$, $N_{\tilde{h}}$ is the sum of the crossed times of all edges of the hexagon in this particle's map counters. $\alpha_{\tilde{h}}^{\tilde{e}}$ and $\alpha_{\tilde{h}} = \sum_{e=0}^{5} \alpha_{\tilde{h}}^{\tilde{e}}$ are the prior counts. The result is used in the particle weight update:

$$w_k^i \propto w_{k-1}^i \cdot I^i \tag{4.6}$$

where w_k^i denotes the weight of the i-th particle at step k. If a particle crossed an edge which has been crossed more frequently than the other edges of the previous hexagon, it tends to have more weight. Thus a consistent walking pattern would be generated.

Each particle contains information about the previous track and the probability of transitions from each hexagon to its adjacent hexagons, which is represented by a probabilistic map. The final result is the best map based on the particle with the highest weight.

4.3 Experimental conditions

4.3.1 Data collection site

The data collection was conducted in two environments, consisting of an indoor and an outdoor setting. Indoor tests were carried out in a building with a known floor plan, while outdoor experiments were carried out using a basketball court, to have exact measurements for the reference map. These maps were used as ground truths to allow a direct comparison with the PDR and HeadSLAM outcomes.

4.3.2 Participants

There were 5 volunteers for the indoor data collection session and 5 volunteers for the outdoor data session with 2 people participating in both. Demographic information of the participants is given in Table 4.1. All participants signed a consent form before data collection started, and they were given the opportunity to ask any questions before deciding to participate in this study. Ethical approval was obtained

from the university ethics committee and this experiment was part of a larger study (R70833/RE001).

Table 4.1: Demographics of participants.

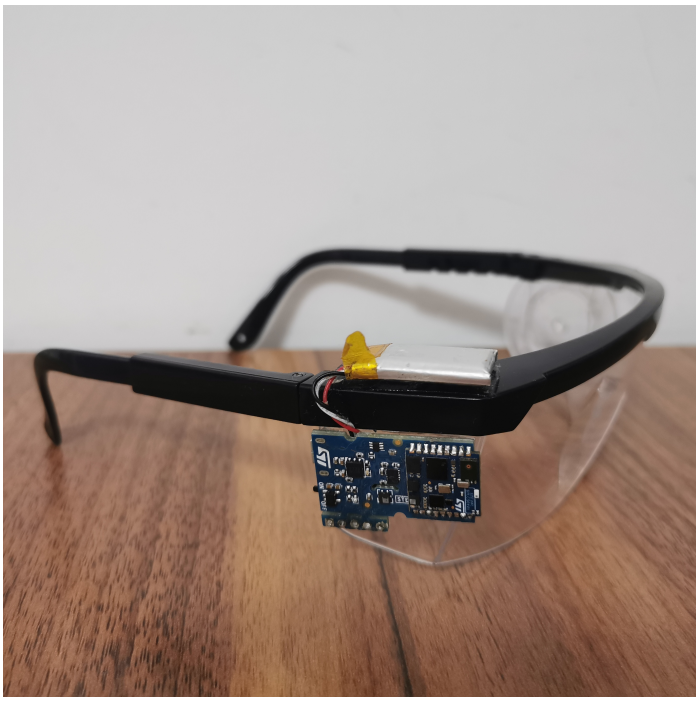
Subjects	1	2	3	4	5	6	7
Age	25	20	23	24	49	46	47
Height (m)	1.80	1.81	1.77	1.66	1.60	1.75	1.60
Weight (kg)	80	74	61	58	55	89	62
Gender	M	M	F	F	F	M	F
Indoor tests	✓	✓	✓		✓		✓
Outdoor tests		✓	✓	✓		✓	✓

4.3.3 Devices

The sensor adopted in the experiments is the SensorTile Microcontroller Unit (MCU) module (STEVAL-STLCS01V1), which includes a low-power 3D accelerometer and 3D gyroscope (LSM6DSM), ultra-low power 3D magnetometer (LSM303AGR), Bluetooth low energy network processor (BlueNRG-MS), 32-bit ultra-low-power MCU with Cortex®M4F (STM32L476JG), and a 100 mAh lithium-Ion polymer battery. Data were collected at 20Hz and transferred to a mobile phone by Bluetooth. Two modules were used for each session with different placements. One was firmly attached to a pair of glasses, whilst the other was connected to a cap. The placement is shown in Figure 4.2.

4.3.4 Experimental setup

Before experiments, participants were asked to put on the cap and wear glasses with the sensors on them. They were requested to place these on their head in such a way that they remained comfortably in contact with the head during walking. Each volunteer was asked to complete the test, for a given environment, six times. Each test took around 10 minutes. In the first three tests, subjects were requested to walk the same predetermined trajectories. The last three tests consisted of volunteers walking randomly in the indoor space or randomly on the lines of the basketball



(a) Glasses with IMU

(b) Cap with IMU

Figure 4.2: Placement of sensor modules on glasses and cap, as used in this study.

Figure 4.3: A subject wearing the devices.

court (outdoor). In all tests, participants were instructed to walk at a normal and constant speed while keeping their head facing the walking direction. A total of 60 datasets were collected in all experiments. Each dataset contained both the data from the instrumented glasses, as well as from the instrumented cap.

4.3.5 Statistical analysis

A total of 3 error measurements were computed by comparing the positional results of the algorithms with the known reference map or path. These error measurements consisted of the Root Mean Square Error (RMSE), average absolute error, and maximum absolute error. All errors were computed across the entire trajectory and in metres. A Kolmogorov-Smirnov (K-S) test was applied to determine if the data were normally distributed. The K-S test showed that they were not normally distributed, which was further confirmed by visual inspection of the histograms. The errors obtained by the PDR were therefore compared with those of headSLAM using the Wilcoxon signed rank test. A p value of less than 0.05 was considered significant. All data analysis was conducted in MATLAB (R2020a, Mathworks, Natick, MA, USA).

4.4 Results

An example of the indoor test results is shown in Figure 4.4 and for the outdoor test results an example is shown in Figure 4.5. The light blue shape is the passable area extracted from the floor plan and acts as a reference for the results generated by the PDR and HeadSLAM. The red lines in (a) represent the trajectory estimated by PDR method. The blue lines in (b) show the trajectory of the particle with the highest weight in SLAM. The red hexagons in (c) represent the trajectory map generated by HeadSLAM, with the heavier shades representing those with a higher visiting frequency. The results for all subjects on all 120 tests can be found here[1].

The obtained errors are shown in Table 4.2 for the two different environments. The ground truths consist of the floor plan for the indoor environment, and the lines on the basketball court for the outdoor environment. The data are split further between PDR and HeadSLAM methods. The K-S test showed that there was a significant difference between PDR and HeadSLAM method (in all 3 error measurements) for both the indoor and outdoor environment.

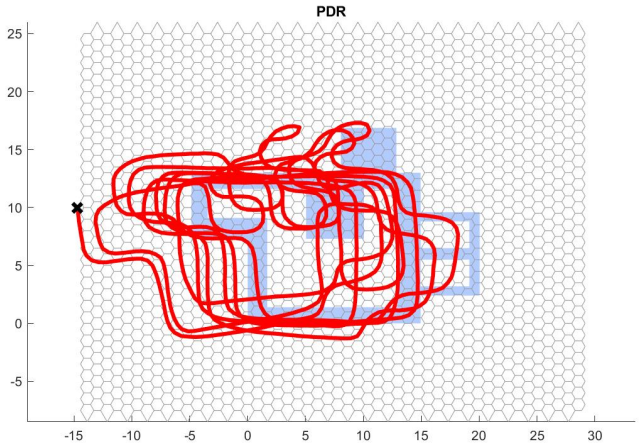
Table 4.2: Errors in meters (m) are given in- and outdoor environments for both PDR and HeadSLam methods. RMSE is the Root Mean Square Error. A significant difference based on Wilcoson signed-rank test (p value $1.63 \cdot 10^{-11}$) was found for all six direct comparisons between the PDR and HeadSLAM outcomes.

Environment	Algorithm	RMSE	Average absolute error	Max absolute error
Indoor	PDR	2.2943	1.4108	8.2473
	HeadSLAM	0.3399	0.1610	1.6597
Outdoor	PDR	2.4358	1.7712	7.6218
	HeadSLAM	0.8343	0.6400	2.8368

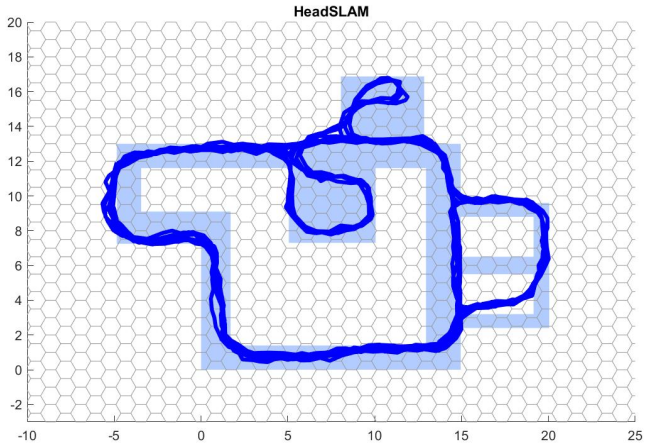
4.5 Discussion

The results showed that HeadSLAM performed better than the PDR in all volunteers and environments. This increase in performance is likely due to the efficient calibration approach of HeadSLAM. However, the effectiveness of HeadSLAM only exists

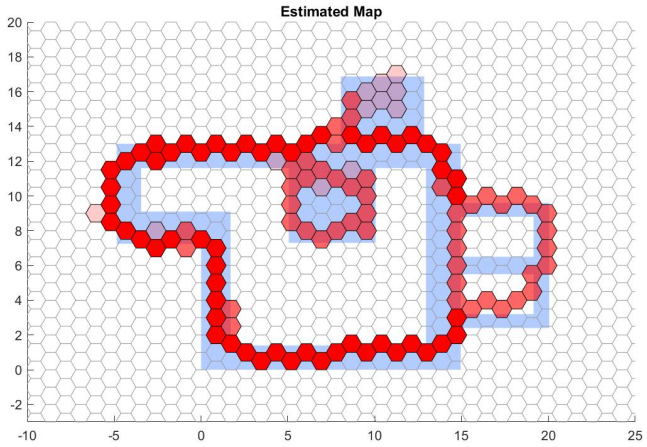
[1]https://zenodo.org/record/5562364#.ZBI7WXbP3tU



(a) PDR

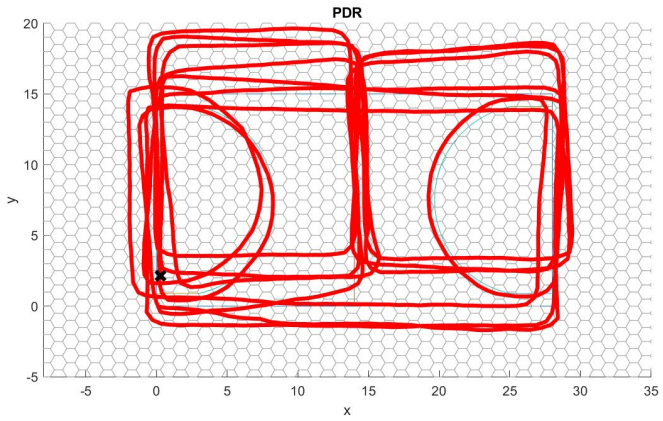


(b) HeadSLAM

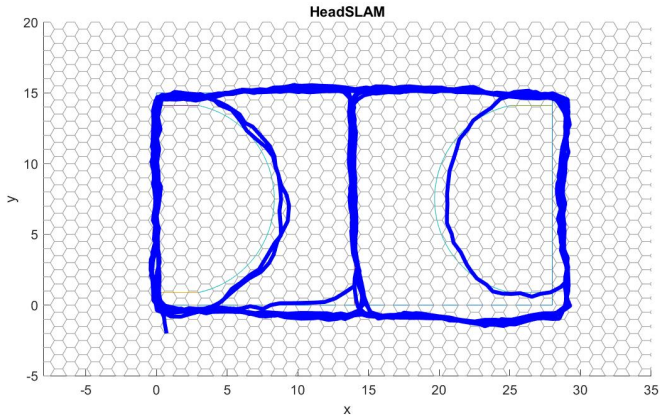


(c) Map

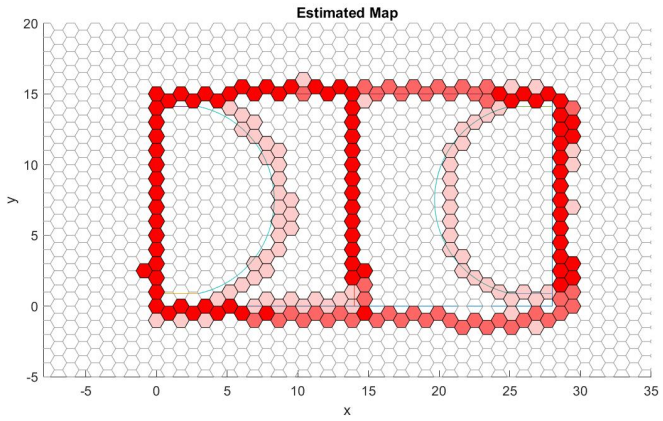
Figure 4.4: Indoor test results example.



(a) PDR



(b) HeadSLAM



(c) Map

Figure 4.5: Example of outdoor test results.

when trajectories on a confined path are repeated and overlap, such as walking in the corridor for several laps. The overlaps allow for re-calibration and provide updates to the probability map that is being generated. HeadSLAM is thus applicable for scenarios in which people cover the same path multiple times.

Our dataset, containing around 20-hour walking data (2 devices · 5 subjects · 2 scenarios · 6 repeats · 10 minutes walking), is much larger than previous studies. [85] used 1 subject to test indoor and outdoor once with 17 minutes of walking in total. [4] used 1 subject to test three times with 30 minutes walking in total. A larger dataset is essential to validate the robustness of these systems and provides a better way to ensure that it performs across different movement behaviours. The situations encompass an indoor setting within the office and an outdoor setting on a basketball court, to increase the complexity of the trajectory, which now includes straight lines, curves, and turns. However, its representation remains inadequate to capture all intricate situations in real-world scenarios, as most corners of the map adhere to right angles. This problem has been addressed in the next chapter with more complex maps.

Although the size of the data set is large, there are other aspects that can lead to potential errors during experiments. For example, although subjects have a balanced gender ratio, the height, weight, and age distribution of the subjects did not follow the real distribution of the entire population, which can lead to bias in the model. Another aspect that may lead to unforeseen errors is the obvious head motion. The head was assumed to be relatively still throughout the experiments. However, various subjects exhibit distinct motion patterns, including both larger and smaller head movements. The affect of head motions in the tracking will be solved in the next chapter.

It should be noted that the current approach still requires parameter optimisation. This is currently not conducted in a automated manner. Automatically tuning the parameter is one of the key issues that need to be solved before practical application can be considered. Since there are internal drift errors in low-cost sensors and because of external factors which influence the sensors such as temperature, the

optimal parameters in Mahony algorithms and the particle filter will be different in each test. FootSLAM could solve this problem by calculating them based on the data collected in the beginning when the foot was kept still. For example, [161] used the adaptive threshold in ZUPT. However, it is impossible to leverage this technique in HeadSLAM, because the head cannot be kept completely still when the participant is in standing pose. Fixing the head during a calibration period does not provide a minimally obtrusive method of tracking. This is something that should be explored in further research.

It should also be noted that other parameters also vary between people. The k in the step length estimation will differ between individuals. In this study, these parameters were adjusted manually. If a plug-and-play system is required for a better user experience, then all these parameters should be set automatically. Itzik Klein and Omri Asraf proposed a method that uses deep learning to estimate the Weinberg gain[92]. This could be an interesting way to solve some of these issues.

HeadSLAM approach presented here can be useful for human tracking at scale. It only requires low-cost IMUs, whilst other infrastructure solutions for positional tracking can be expensive to setup and maintain. It is also not prone to privacy issues, which could arise when cameras are used. More importantly, the system can be fully self-contained creating possibilities for very secure tracking. Although HeadSLAM could work without external infrastructures or previously known information, there is also a possibility to combine it with other methods to create a more accurate or robust system. Switching between HeadSLAM and approaches that require infrastructure can also solve loss of position estimation whenever, for example, there is a temporarily weak or loss of WiFi/GPS signal.

The K-S test is a non-parametric statistical test used to determine whether a sample comes from a specific probability distribution or comparing if two samples follow the same distribution. It has been used in this chapter twice to check whether one sample follows the normal distribution and difference of two lists of errors from two methods. They both rejected the null hypothesis that the distributions are the same.

In practice, the K-S test can be very sensitive to differences between distributions, especially more sensitive near the centre, potentially leading to a rejection of the null hypothesis even if the bulk of the distributions is similar.

4.6 Summary

HeadSLAM could reach an average RMSE of $0.34m$ indoor and $0.83m$ outdoors during a 10-minute walk, which shows a significant improvement compared to a PDR method. It shows the potential of longer-term stable inertial tracking based on head-mounted low-cost sensors, which allows for possible cheap applications in healthcare, sports, or emergency services. The particle filter in HeadSLAM approach smooths out small errors that are due to head motions. However, further research will need to conducted to deal with the problem of unexpected (larger) head movements during walking. These kinds of unique problems need to be addressed to generate real-world impact for head-mounted sensors.

Chapter 5

HINNet: Inertial navigation with head-mounted sensors using a neural network.

Previous chapters proposed a PDR method for head-mounted sensors based on the traditional model-based method, and its extension HeadSLAM allowing long-term stable tracking. However, there exists a non-negligible problem that has not been solved in inertial tracking with a head-mounted sensor, which is the affect of head motions during tracking. HeadSLAM could smooth out the small errors caused by head motions, but it still needs an effective way specifically for this problem.

To solve this problem, HINNet is presented in this chapter, which is the first deep neural network (DNN) pedestrian inertial navigation system that allows free head movements with head-mounted IMUs. It deploys a 2-layer bi-directional LSTM. A new "peak ratio" feature is introduced and utilised as part of the input to the neural network. This feature unveiled latent information in the frequency domain of the time series data. This information can be leveraged to solve the issue of differentiating between changes in movements related to the head and those associated with the walking pattern. A dataset with 8 subjects totalling 528 minutes has been collected on three different tracks for training and verification. HINNet could effectively distinguish head rotations and changes in walking direction with a distance percentage error of 0.46%, a relative trajectory error of 3.88 m, and an absolute trajectory error

of 5.98 m, which outperforms the current best head-mounted PDR method.

5.1 Introduction

The unpredictability of human behaviour in dynamic environments makes it a complex task for position tracking, as various complicated motions overlap locomotion. Tracking with GPS, WiFi or other infrastructure based methods could ignore this issue for that complex motions would not affect as long as the signal receiver is still with the subject. But these infrastructure based methods have limitations which were summarised in Chapter 1. More than that, in some of the use cases, such as those that can be found in healthcare, sports, or emergency services, they also come with a potential additional level of behavioural complexity, which could be measured and used for further studies on behaviours. Although these signal-receiving methods are not affected by behaviours, they lose the ability to measure the behaviours. The inertial navigation is more sensitive to variations in behavioural patterns, as it builds-up the positional estimation from the information that is extracted from motions, which allows it to also extract behaviour information for further usage.

Traditional pedestrian inertial navigation methods utilise integration of measurements from accelerometers and gyroscopes, or apply PDR. For example, Chapter 3[68] adopted peak detection to detect steps, combined with a complementary filter [114] to determine orientations and a Weinberg model to subsequently detect step lengths with head-mounted sensors. [87] and [123] used IMUs in smartphones to track users also with a step and heading system, while [57] used a single foot-mounted IMU and a hip-worn smartphone for tracking with a zero velocity update, action recognition algorithm and particle filter. [51] used EKF and Weinberg model with a waist-worn IMU. Foot-mounted IMUs can also be used, as shown by [44], indicating that a range of possible sensor placements are available for inertial navigation.

Although traditional model-based inertial tracking methods may have accurate results in experimental environments, their performance can be reduced quickly in

real-world scenarios. Performance would be negatively affected if there are more varied users, if there is an increased unpredictability of the environment, or if the user activities themselves become less predictable. Because model-based pedestrian inertial navigation methods represent the functional relationships between inertial measurement inputs and estimated localisation information, which use simple manually selected fitting equations with different parameters for each individual. This means that the output suffers from sensor drifts, accumulated errors, and fitting errors from the simple models. For example, Weinberg step length model estimates the step length only with a fourth root of the difference between the maximum and minimum vertical accelerations in one step, which multiplies with a parameter k. This will be different for each participant [166].

Recently, machine learning has shown its impressive potential in solving tasks in a variety of fields such as natural language processing [33], image processing [191], and healthcare [127] [72] [133]. One of the key benefits consists of the ability to model complex non-linear relationships with large volumes of data. The precise identification of patterns or trends indicates the suitability of it to be utilised in pedestrian inertial navigation. [22] adopted a 2-layer Bi-LSTM to learn human odometry with IMU that were either hand-held, in a pocket, in a handbag, or on a trolley. [62] estimated horizontal positions and heading direction of a moving subject from a sequence of IMU sensor measurements using a phone with ResNet, LSTM and Temporal Convolutional Network (TCN). The feasibility and performance of human odometry estimation with machine learning have been demonstrated in these previous published studies. It also showcases the ability to collect this data with different sensor placements. The selection of placement should be based on the applicability of that sensor for use in complex real-world activities.

The advantages of using head-mounted sensors have been discussed in previous chapters. And according to the systematic review in Chapter 2, there are only a few papers [59, 192] that explore inertial navigation with head-mounted sensors [69]. That can be attributed to the behaviour issue that rotational head motions (used to

change the visual field) could be mistaken with body rotations (related to a change in walking direction). This is not more or less a problem for sensor attachments to the feet, waist, or chest. The "heading" direction of the head may not always align with the body orientation, which can lead to an obvious position error in pedestrian localisation. The three papers with head-mounted sensors mentioned above are all based on traditional pedestrian inertial navigation methods with step-and-heading systems, which lack generalisability in more complex real-world scenarios. None of them solved the problem of misalignment of head and body orientation, as they simply ignored head rotations during walking.

To solve this interesting problem, this chapter proposed the first DNN framework for pedestrian inertial navigation with head-mounted sensors allowing for free head movements: HINNet. Previous studies on pedestrian navigation with head-mounted sensors all depended on traditional pedestrian inertial navigation methods with step-and-heading models, and only work when the head is facing forward. This is the first study to propose deep learning methods in pedestrian navigation with head-mounted sensors, and allow for free head rotations during walking. In this study, the relative trajectory error (RTE), absolute trajectory error (ATE) and the percentage error of distance for HINNet and PDR are determined against a ground truth measurement during long-period walking trials. This chapter introduces a DNN with a set of new features to address the issue of separating head rotation signals from walking data for head-mounted IMUs. This information is then used for positional tracking.

5.2 Methods

HINNet system is summarised in Figure 5.1. It provides an overview of how data from the head-mounted sensor are processed to estimate trajectories.

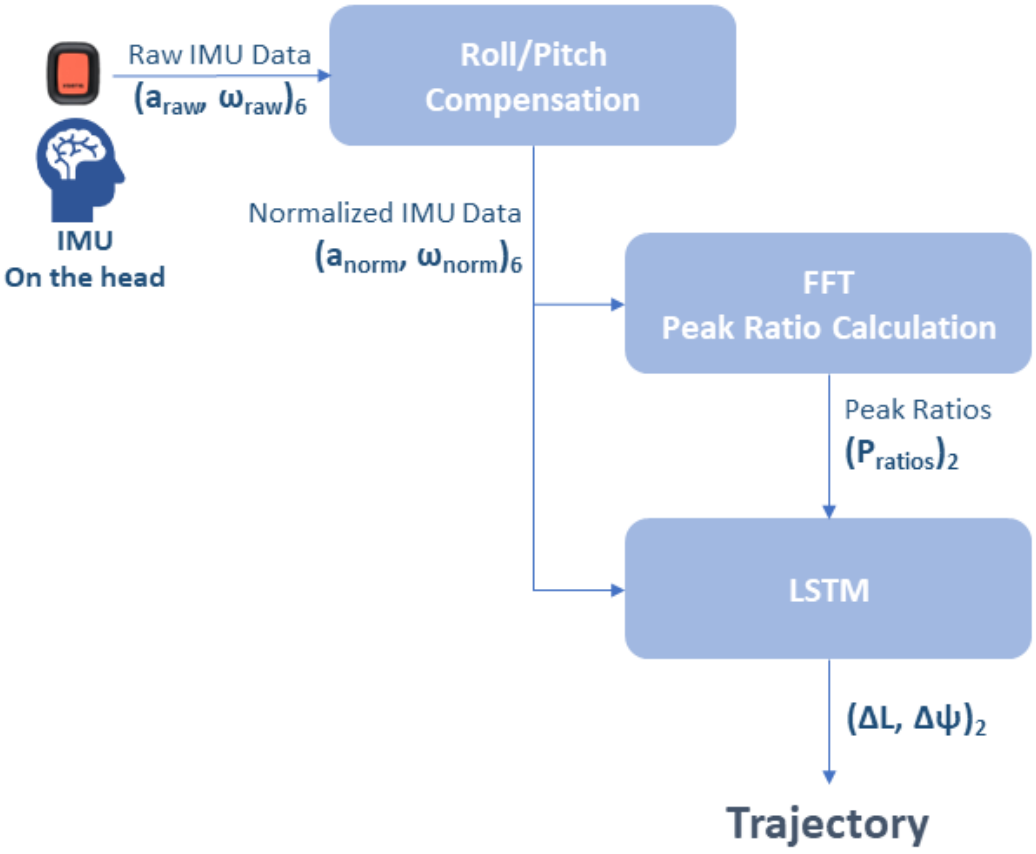


Figure 5.1: Overview of HINNet, a deep neural network for pedestrian inertial navigation with head-mounted sensors that allows for free head movements. The Inertial Measurement Unit (IMU) placed on the head is shown in the top left corner. After normalisation, the IMU Data is first transformed into frequency domain with a fast fourier transform (FFT) to generate the peak ratio feature. The peak ratios and normalised IMU data form the input into the Long Short-Term Memory (LSTM) model.

5.2.1 Roll and pitch compensation

The raw IMU sensor data was first transformed into a normalised coordinate system in which the z axis is perpendicular to the horizontal plane, which was accomplished by Equation 5.1 and Equation 5.3, where a and ω are $3 * length_{data}$ vectors representing accelerometer data and gyroscope data.

$$a_{norm} = R_a^{-1} \cdot a_{raw} \tag{5.1}$$

$$R_a = R_x(\phi)R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ \sin\phi\sin\theta & \cos\phi & \sin\phi\cos\theta \\ \cos\phi\sin\theta & -\sin\phi & \cos\phi\cos\theta \end{bmatrix} \tag{5.2}$$

$$\omega_{norm} = R_\omega^{-1} \cdot \omega_{raw} \tag{5.3}$$

$$R_\omega = \begin{bmatrix} 1 & 0 & -\sin\theta \\ 0 & \cos\phi & \sin\phi\cos\theta \\ 0 & -\sin\phi & \cos\phi\cos\theta \end{bmatrix} \tag{5.4}$$

After compensation, the normalised z-axis of the IMU will align with the gravity direction. The heading direction (yaw angle ψ) is considered in reference to the angular velocity ω_z and there is no expected gravity component on the normalised x-axis and y-axis, which will be used later for peak ratio calculation in Section 5.2.2.

In this part, the input $(a_{raw}, \omega_{raw})_6$ is the 6-dimensional raw IMU data, the output $(a_{norm}, \omega_{norm})_6$ is the 6-dimensional normalised IMU data.

5.2.2 Feature of peak ratio

The key is to be able to distinguish the head rotation from the whole body rotation. This can be done by finding the difference in the walking patterns, which are recorded by the head-mounted IMU. There are two obvious kinds of body movements during walking that can generate regular acceleration waves. They are generated by (i) the body swinging from left to right and (ii) the variation in linear acceleration in the front-to-back direction caused by stepping. The change in signals that are recorded on

the x and y axes when the head is rotating is shown in Figure 5.2. Only stepping will be recorded on x axis when the head is held "straight" during walking. Alternatively, both stepping and side swing will be measured on x axis when walking with the head facing sideways. This concept is used to help determine whether (i) a rotation occurred due to a head movement without any change in walking direction, or (ii) if the head relative to body remained aligned and a change in walking direction took place [167].

One full wave of a swing requires one stride (two steps) and a full stepping wave requires one step. The frequency difference between stepping and side swings gives us the possibility of visualising these two motions in the frequency domain. The spectrums of accelerometer sensor data in the horizontal plane (x axis and y axis) are shown in Figure 5.3.

The relation of the two peaks in Figure 5.3 can be represented by the peak ratio P_{ratio} in Equation 5.5 [167].

$$P_{ratio} = \frac{P_{swing}}{P_{stepping}} \tag{5.5}$$

The method applied for this paper consists of the following steps: for each time point t, the accelerations on the x axis in a $2s$ window $[t - 2s, t)$ prior to t are transformed into the frequency domain using an FFT. Then the two peaks P_{swing} and $P_{stepping}$ are extracted from the generated frequency spectrum. These are subsequently used to calculate a P_{ratio}^{before}, which represents the 2 seconds prior to t. The same operation is conducted on the $2s$ window $(t, t + 2s]$ after the time point t, for which we get a P_{ratio}^{after}. Gyroscope data are used to determine if a rotation took place at the time point t. If P_{ratio}^{before} and P_{ratio}^{after} are very similar, then it is likely that the walking direction has changed. Essentially, only the "stepping" information should be projected on the x axis before and after rotation (see the top image in Figure 5.2). In the case of pure head rotation, the projection of the "side" swing on the x axis will alter P_{ratio}^{after}, creating a different outcome compared to P_{ratio}^{before} (see the bottom image in Figure 5.2).

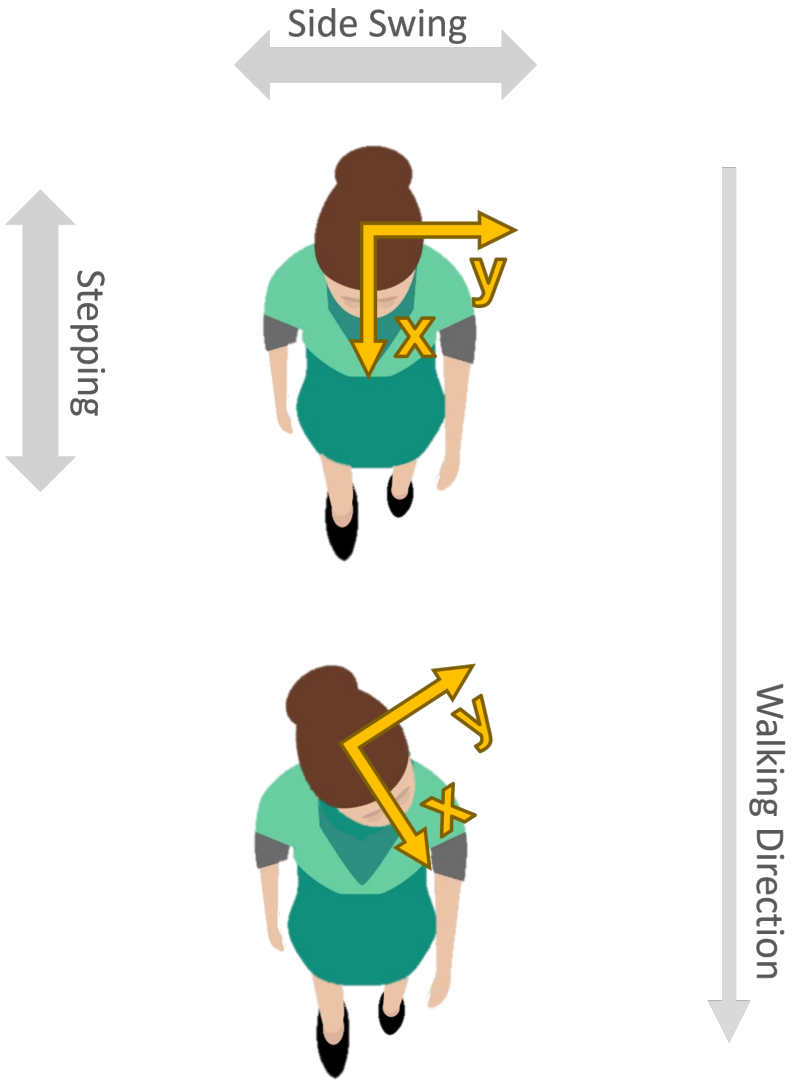


Figure 5.2: The images show a person walking forward at two different arbitrary time points. The top image shows them with the head facing in the same direction as they are walking. The bottom image still depicts a forward walking direction, but with the head rotated to the left. Different magnitudes of side swing and stepping are projected on the x and y axes of the head-mounted IMU depending on the orientation of the head.

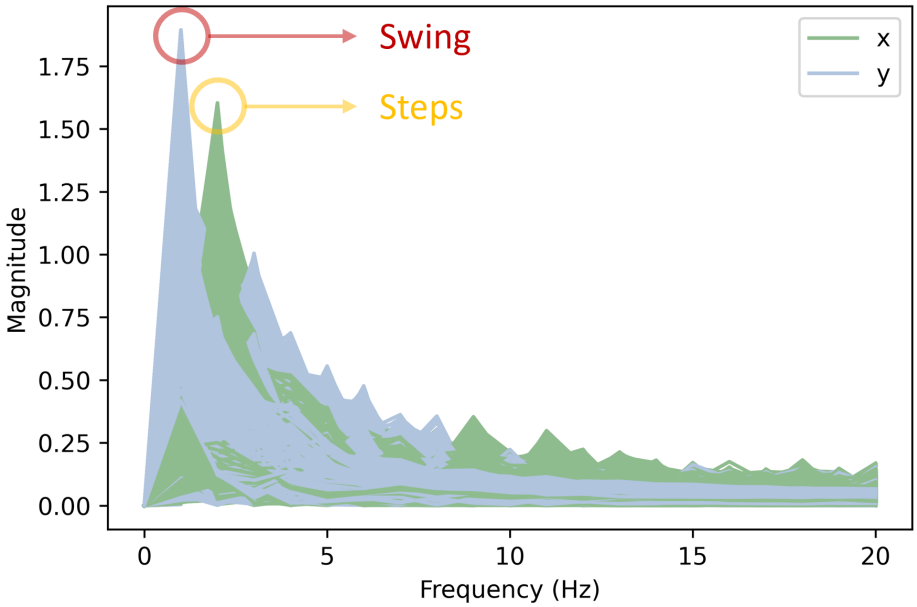


Figure 5.3: Frequency spectrum of accelerations obtained from the x and y axes. This information was produced by a Fast Fourier Transform (FFT) taken from the normalised accelerations. The peak in the red circle is generated by swings and the peak in the yellow circle is produced by steps.

The $(P_{ratios})_2$ (see Figure 5.1) consists of both the P_{ratio}^{after} as well as the P_{ratio}^{before}. This information is then used in the LSTM.

5.2.3 Deep neural network framework

After pre-processing and feature generation, a sequence of normalised IMU data and peak ratio features were fed into the neural network as input. The RNN would be the best choice for this task, since it was designed to process sequential data or time series data, and has already been successfully applied in e.g. speech recognition [142] and machine translation [171]. The LSTM was leveraged in this research instead of vanilla RNN to prevent the vanishing gradient problem. The LTSM is a special kind of RNN, capable of learning long-term dependencies [66]. The LSTM works as a function f_θ mapping normalised accelerometer data a, gyroscope data ω and peak ratio features P_{ratios} to walking distance Δl and rotation $\Delta\psi$ over a window,

$$(a_{norm}, \omega_{norm}, P_{ratios})_{8*120} \xrightarrow{F_\theta} (\Delta l, \Delta\psi)_{2*1} \tag{5.6}$$

where a window length of 120 frames ($2s$) was used. To leverage each sample's past and future contexts, a bi-directional LSTM network was adopted in the study, with 2-layers stacked to add levels of abstraction of input observations over time. Each layer had 128 nodes. The framework overview is shown in Figure 5.4.

After getting $(\Delta l, \Delta\psi)_2$ in every window, the whole trajectory could be generated by connecting them sequentially.

5.3 Experiment conditions

5.3.1 Data collection site

The data collection was conducted in three different environments with different sizes and paths to ensure a broader ability to generalise. Figure 5.5 shows the shapes and dimensions of trajectories of the three scenarios. The trajectories include straight routes, curves, and turns with different angles, increasing the complexity and applicability of the tests.

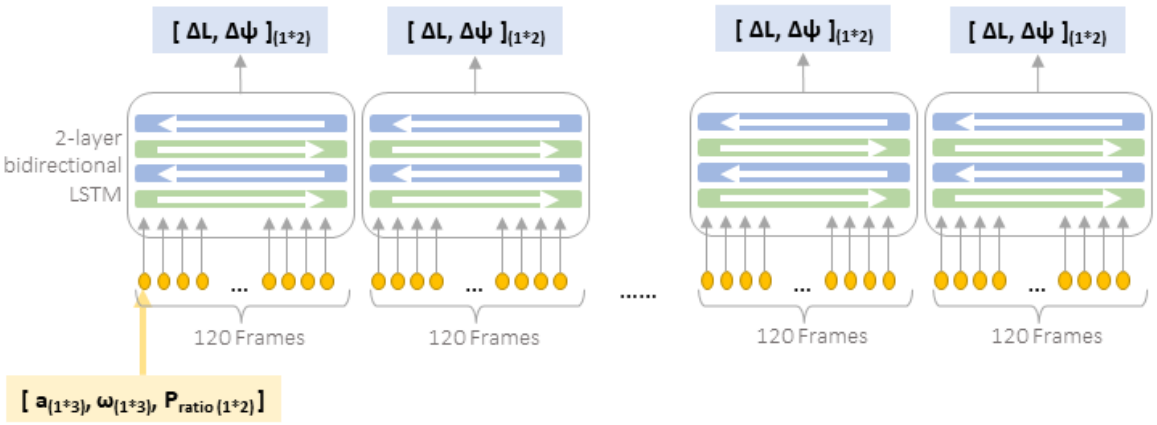


Figure 5.4: Overview of HINNet neural network. In each window, 120 frames of normalised IMU data and peak ratios are sequentially inputted into a 2-layer bidirectional LSTM, which outputs the displacement and orientation variation.

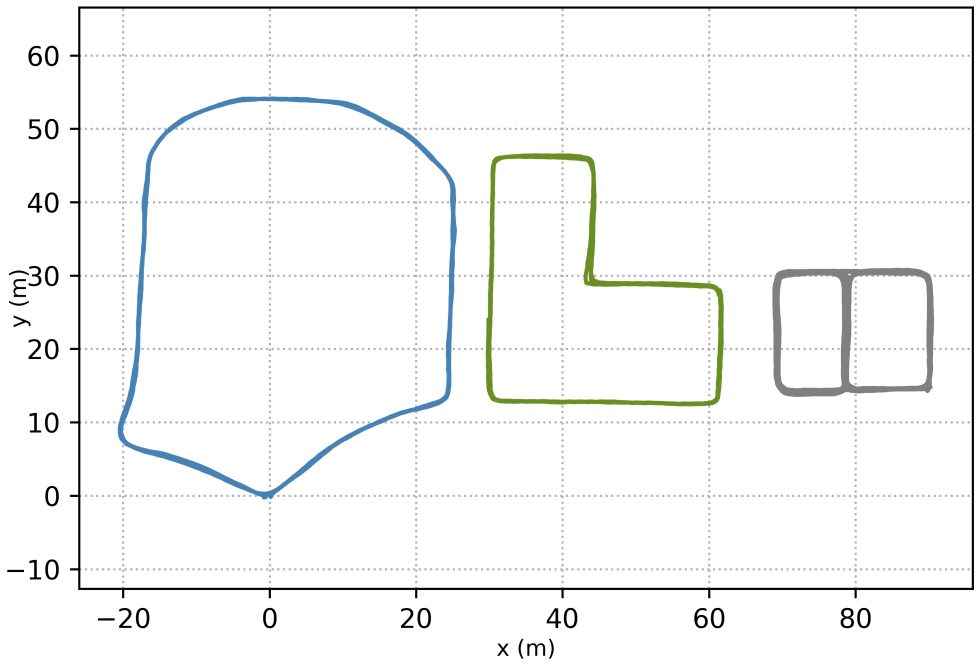


Figure 5.5: The exact map of each of the walked paths is shown in this figure. The volunteers were asked to follow these paths during data collection.

5.3.2 Participants

There were 8 volunteers participating in the experiments, with the demographic information of the participants given in Table 5.1. According to the result of a recent systematic review [69], 128 out of 145 papers about human inertial navigation used less than 8 subjects. The amount of subject in this study exceeds 88% papers in this domain. All participants signed a consent form before data collection started and were given the opportunity to ask any questions before deciding to participate in this study. Ethical approval was obtained from the University ethics committee and this experiment was part of a larger study (R70833/RE001).

Table 5.1: Demographics of participants. (M=male, F=female)

Subject	Age (years)	Height (m)	Weight (kg)	Gender
1	42	1.79	73	M
2	27	1.65	60	F
3	25	1.79	77	M
4	24	1.77	62	F
5	48	1.75	85	M
6	49	1.60	58	F
7	30	1.70	66	M
8	21	1.65	65	F

5.3.3 Devices

Xsens Dot sensors (Xsens Technologies BV, Enschede, Netherlands) were used in the experiments. They logged the 3D angular velocity from a gyroscope, 3D acceleration from an accelerometer, and 3D (local) magnetic field from a magnetometer. Data was collected at 60Hz and transferred to a mobile phone by Bluetooth. An Xsens Dot was firmly attached to the forehead with a strap and another one was placed on the chest for reference. The placement is shown in Figure 5.6. A phone was also firmly attached on the chest with straps.

An application developed in Unity 3D game engine (Unity Technologies, San Francisco, CA, USA) based on Google ARCore was installed in the phone for ground truth

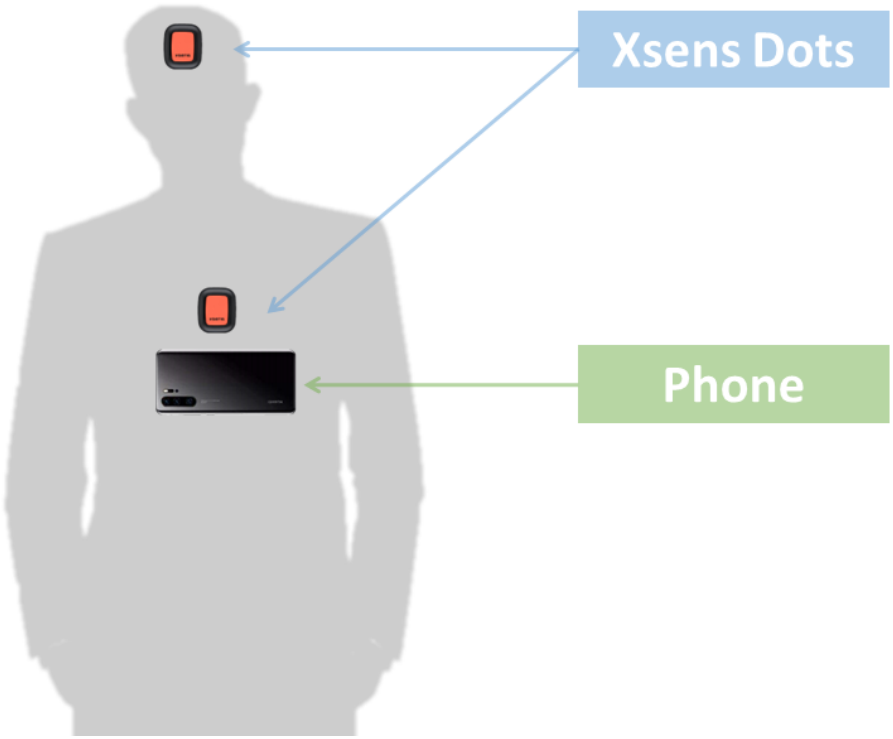


Figure 5.6: Schematic of the sensor placement. A IMU (Xsens Dot sensor) was attached on the head using a strap, while another IMU is attached by applying a chest harness. This harness also holds a phone to collect positional reference data.

generation. The ground truth trajectories in Figure 5.5 were generated using this information.

5.3.4 Experimental setup

Before the experiment started, participants were asked to put on the straps with sensors. They were requested to place one Xsens Dot on their head in such a way that it remained comfortably in contact with the head during walking. In each test, subjects were requested to walk the predetermined trajectories for 5 to 10 minutes. Each volunteer was tested between 4 to 16 times. In all the tests participants were instructed to walk at a normal and constant speed whilst rotating their head in a random manner. A total of 79 datasets were collected across all experiments, with a total time of around 528 minutes. Each dataset contained both the data from the Xsens Dots, as well as from the phone.

5.3.5 Training setup

We trained HINNet for 150 epochs on the single NVIDIA TITAN Xp GPU with 12 GB memory. Adam optimiser was used to train. The learning rate is set to $1e^{-4}$ with a dropout rate of 0.25.

5.4 Results

A recently published PDR system [68] was selected as the current baseline method for comparison, which is based on traditional step-and-heading systems, with peak detection to detect steps, complementary filter [114] to determine orientations, and Weinberg model to subsequently detect step lengths with head-mounted sensors. This particular method is the most accurate inertial tracking method available for head-mounted sensors, as the algorithm has been adapted to this specific sensor location, as shown in [68]. Current PDR algorithms are not agnostic to sensor placement, since the signals can drastically change depending on where the sensor is attached to the body. The percentage errors of total distances, RTE, and ATE were computed

for PDR and HINNet. RTE and ATE are standard position evaluation metrics in navigation [187]. RTE is defined as the average RMSE over a fixed time interval of 1 minute. ATE is the RMSE between the whole ground truth trajectory and the estimated trajectory. These particular metrics have been widely utilised in other tracking studies, thus deemed to be a suitable output for this research.

Figure 5.7 shows the estimated and real walking distance (Delta Distance) and variation of orientation (Delta Orientation) in each $2s$ from one test dataset.

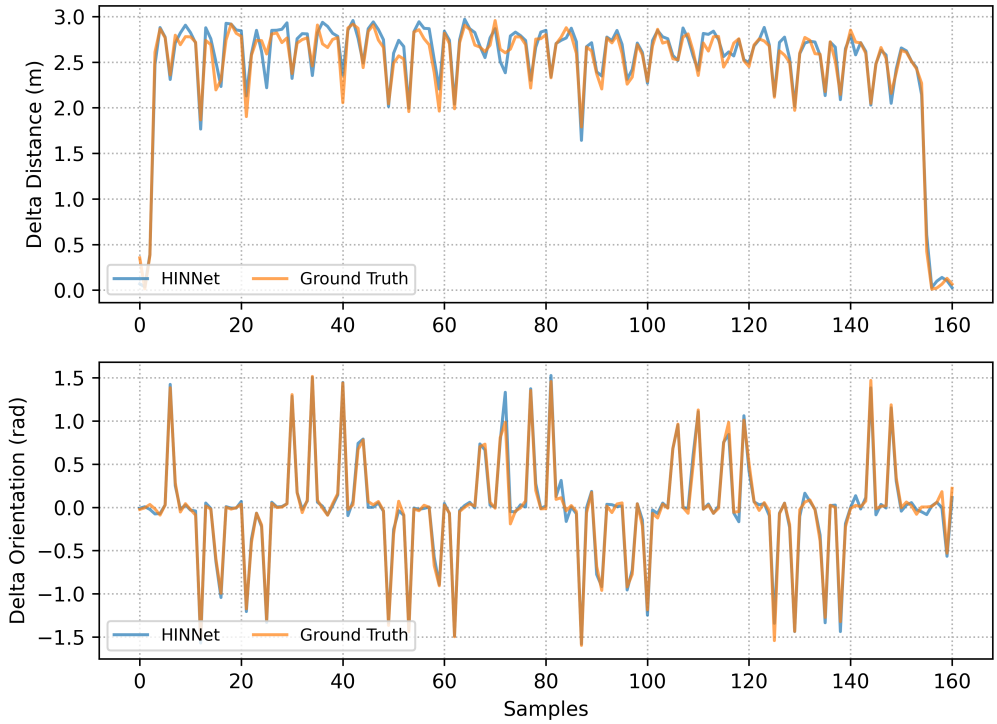


Figure 5.7: The walking distance in each 2 second window (Delta Distance) and the variation of orientation in each 2 second window (Delta Orientation) from one of the test dataset. Orange lines represent the values from ground truth. Blue lines show the values estimated by HINNet.

The estimated trajectories of different methods under the three scenarios are shown in Figure 5.8.

RTE, ATE and the percentage error of the total distance of each method are

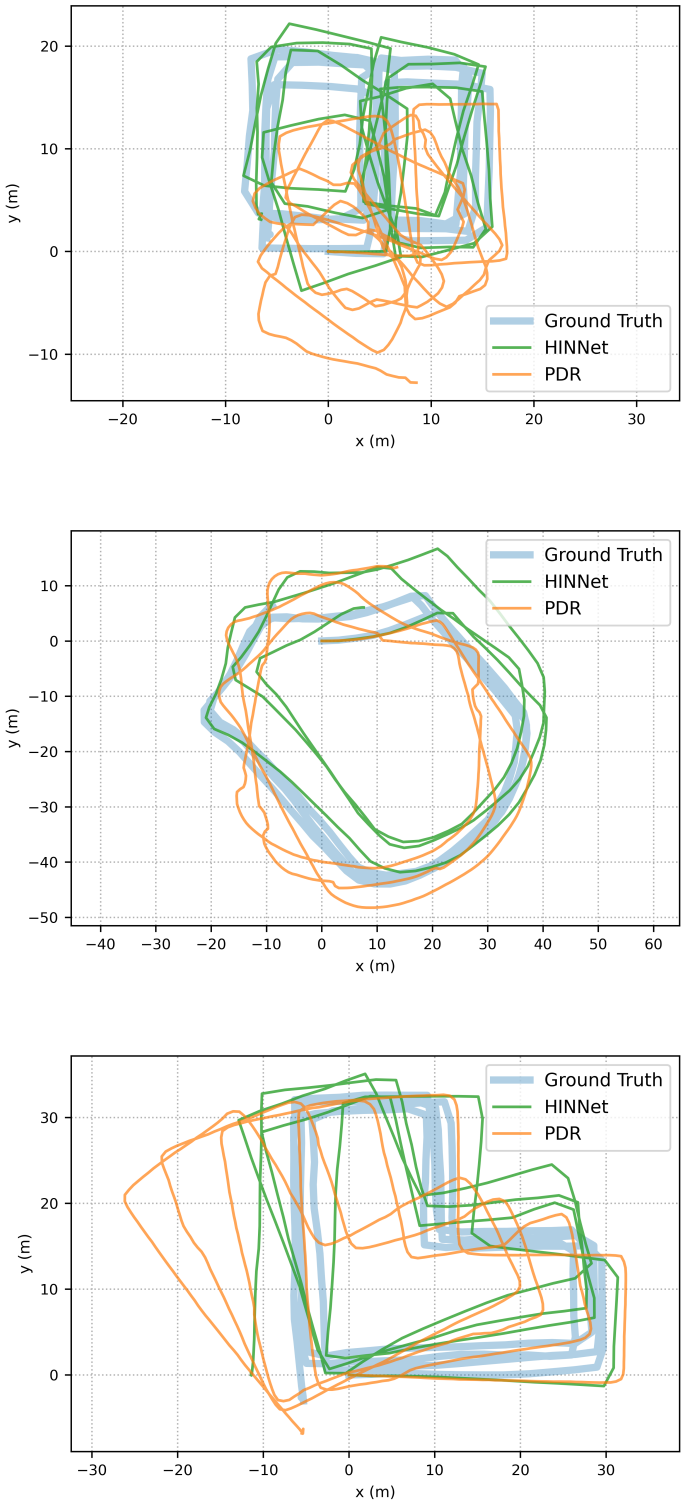


Figure 5.8: Estimated trajectories of different methods. The light blue lines represent the ground truths. Green lines show trajectories estimated by HINNet. Orange lines are the trajectories generated by the PDR.

summarised in Table 5.2.

Table 5.2: Relative trajectory error (RTE), absolute trajectory error (ATE), and percentage error of total distances of HINNet and PDR.

Methods	RTE(m)	ATE(m)	Distance Error (%)
HINNet	3.88	5.98	0.46
PDR	5.76	9.89	6.05

An ablation study was conducted to investigate the contribution of peak ratio features in the neural network. A model without peak ratio feature input was trained with the same parameters as the one with the features. Results are shown in Table 5.3.

Table 5.3: Ablation study on the peak ratio feature. Relative trajectory error (RTE), absolute trajectory error (ATE), and percentage error of total distances of HINNet with and without peak ratio feature.

Methods	RTE(m)	ATE(m)	Distance Error (%)
With peak ratio	3.88	5.98	0.46
Without peak ratio	4.47	7.94	0.61

5.5 Discussion

We presented the first pedestrian inertial navigation system for a head-mounted IMU that allows for free head-movements by applying a DNN. HINNet performs much better than a dedicated PDR on walking distance estimation, with a distance error of 0.46%. The PDR requires tuning of specific parameters for each individual, while HINNet does not need human input to tune any parameters manually. The difference in performance can also be observed in the lower RTE and ATE of HINNet compared to that of the PDR. However, the biggest improvement accomplished by HINNet is the possibility to eliminate head motion that might be due to "sightseeing" while walking. In Figure 5.8, there are a lot of small curves in the PDR trajectories. These represent the head motions that are not related to the walking, and thus they lead to false additional trajectories. While the trajectories estimated by HINNet are a lot

more similar to the shape of the ground truth, which implies its ability to distinguish head rotations that are independent of walking. These promising results need to be considered carefully due to the relatively small sample size.

The curves presented in the PDR trajectories have different sizes and frequencies, which indicates the variation in head rotation behaviours between different volunteers. Interestingly, the data showed that when people were specifically told to rotate their head, head rotations were more frequent with larger angles compared to previous tests in which volunteers were only asked to walk as usual. It indicates that the instructions given to subjects should be composed carefully to avoid implications leading to biased data. Nonetheless, for this study the additional head rotation was requested to better asses the performance between the algorithms. It should also be noted that these increased head motions might very likely when people are, e.g., sightseeing or navigating busy streets.

Although HINNet could effectively differentiate head motions, limitations still exist when head and body rotate separately compared to synchronised motion. Overlaps of head and body movements lead to confusion in estimation and can benefit from further information produced by other sensors (e.g. magnetometer, GPS) to maintain a correct heading direction.

It should also be noted that the position error is still accumulating over time, although the head rotation error has been eliminated. To improve this, additional information or complementary sensors could be leveraged, such as Bluetooth RFID, GPS, etc. Another option would consist of using self calibration methods like Head-SLAM [70].

HINNet uses the peak ratio feature in the input, which is a hand-crafted feature and seems redundant for deep learning, as one of the advantages of deep learning is that representations can be learnt without feature engineering. In fact, peak ratio features represent the latent information in the frequency domain, while LSTM is more good at featuring time domain information. Therefore, it is reasonable to add frequency domain features in the network like the peak ratio feature. However, it is

not only limited to a hand-crafted peak ratio feature, rawer data in the frequency domain as input probably also work. There is also research proposed LSTM with both time and frequency domain input [109] that could be leveraged in the future studies.

Considering the potential implementation in real-time positioning, HINNet would have a lag of $2s$, since the bi-directional LSTM requires information from both past and future in a $2s$ window, and P_{ratio} is also calculated in a $2s$ window. As the preferred walking speed of human is 1.42 ms^{-1} [18], less than 3 m lag will be contained in the application, which is acceptable in most tracking scenarios. However, if the accurate real-time position is strictly required, algorithms may need to stop using future information. Model size also needs to be considered for a trade-off between accuracy and running time in the real-time applications.

Besides the above research directions, future studies could also focus on developing robust tracking methods to include more complex motions like running, jumping, sitting, etc. Furthermore, extending the tracking system to a three-dimensional space would also provide real-world monitoring benefits, as it would allow for incorporation of information related to, e.g. stair climbing or walking on slopes.

5.6 Summary

This study highlighted the performance increase that can be obtained by considering the sensors placement and designing suitable algorithms that can overcome identified biases in localisation that result from behaviours at these placements. It showed that HINNet could cope with head rotations for head-worn sensors outperforming the current best head-mounted PDR method.

Chapter 6

HINNet + HeadSLAM: Robust inertial navigation with machine learning for long-term stable tracking

Human inertial tracking with head-mounted sensors had some problems before. Two studies proposed in Chapter 4,5 solved part of the problems separately: HINNet is able to track people with free head motions; HeadSLAM allows long-term tracking with stable errors. In this chapter, to allow free head rotations meanwhile support long-term tracking, HINNet is combined with HeadSLAM and tested. The result shows that the combination could effectively distinguish head rotations and maintain a low and stable position error in long-term tracking, with an ATE of 2.69 m and an RTE of 3.52 m.

6.1 Introduction

There has been a rapid development in technology and algorithms that allow real-time human position tracking. The maturation of this technology has brought many possibilities that could substantially change our modern way of life. However, the applicability of monitoring in the real world is governed by the performance of these systems. However, certain scenarios require robust and accurate information, even when complex environmental constraints are in place. The environment can lead

to a range of behavioural responses that influence our motor outcomes [10]. This indicates that people can move in unpredictable ways, as they navigate and interact within their environment. This is particularly important to consider when we are exploring solutions for areas such as security, first responders, or healthcare. The location of people might need to be tracked accurately, as their safety and lives might depend on it. They can e.g. move in and out of buildings with unknown layouts and the monitoring system will need to be able to deal with that. Furthermore, in environmentally complex environments, it is unlikely that any infrastructure is either available or operational under those conditions (disaster areas are a good example of this). Normally, additional infrastructure (such as Wi-Fi) can be leveraged for positional tracking, but it should be clear that there is no certainty of this in the aforementioned situation. Therefore, the solution should be infrastructure-agnostic.

Inertial tracking has been discussed to be one of the competitive solutions. And for inertial tracking with head-mounted sensors, two systems were proposed in Chapters 4 and 5: HeadSLAM[70] and HINNet [71].

HeadSLAM was proposed to improve tracking accuracy during a longer tracking duration. Traditional PDR methods suffer from error accumulations, because of the lack of calibration methods. HeadSLAM uses estimated trajectories at the earlier stage, which was proved to be more reliable, to calibrate estimated trajectories in the later stage that tended to have larger accumulation errors. HeadSLAM could reach an average RMSE of 0.34 m indoors and 0.83 m outdoors during 10 min walks in a 20 hour dataset. This showed a significant improvement compared to PDR method. However, HeadSLAM in the original study still used odometry from a basic PDR method [68], which has two drawbacks; (i) parameters need to be optimised for each individual, and (ii) it does not allow free head rotations during walking.

HINNet is a pedestrian inertial navigation system that allows free head movements with head-mounted IMUs by applying a DNN[71]. It could effectively distinguish head rotations and changes in walking direction. It was shown to have a relative trajectory error of 5.57m. Although it solved the problem of head rotations, the estimation

errors got larger as the testing time got longer. The underlying reason for this was that there is no efficient calibration that is being performed.

In this chapter, the above two methods are neatly combined to solve both the head rotation problem and the long-term estimation error accumulation. ATE and RTE were used as performance measurements. A performance comparison with just HINNet will be given.

6.2 Methods

The whole system is summarised in Figure 6.1.

HINNet was proposed in Chapter 5. It estimates the odometry with deep learning; meanwhile, it solves the problem of differentiating between head rotations and changes in the walking direction.

HeadSLAM was proposed in Chapter 4. The odometry generated in HINNet is subsequently fed into HeadSLAM for trajectory calibration to eliminate the accumulated drift.

6.3 Results

This chapter uses data from HINNet [71], in which the IMU data was collected from a head-mounted XSens Dot and the ground truth was collected from a chest-mounted phone based on visual inertial odometry (VIO).

Figure 6.2 shows the results from HINNet with HeadSLAM compared to the original HINNet on three different tracks.

Three metrics were used for the quantitative analysis:

ATE (m): Absolute trajectory error. ATE is the RMSE between the whole ground truth trajectory and the estimated trajectory.

RTE (m in Δt): Relative trajectory error. RTE is defined as the average RMSE over a fixed time interval (1 minute in this study) with alignments of the initial states.

RTE and ATE are standard position evaluation metrics in navigation [187].

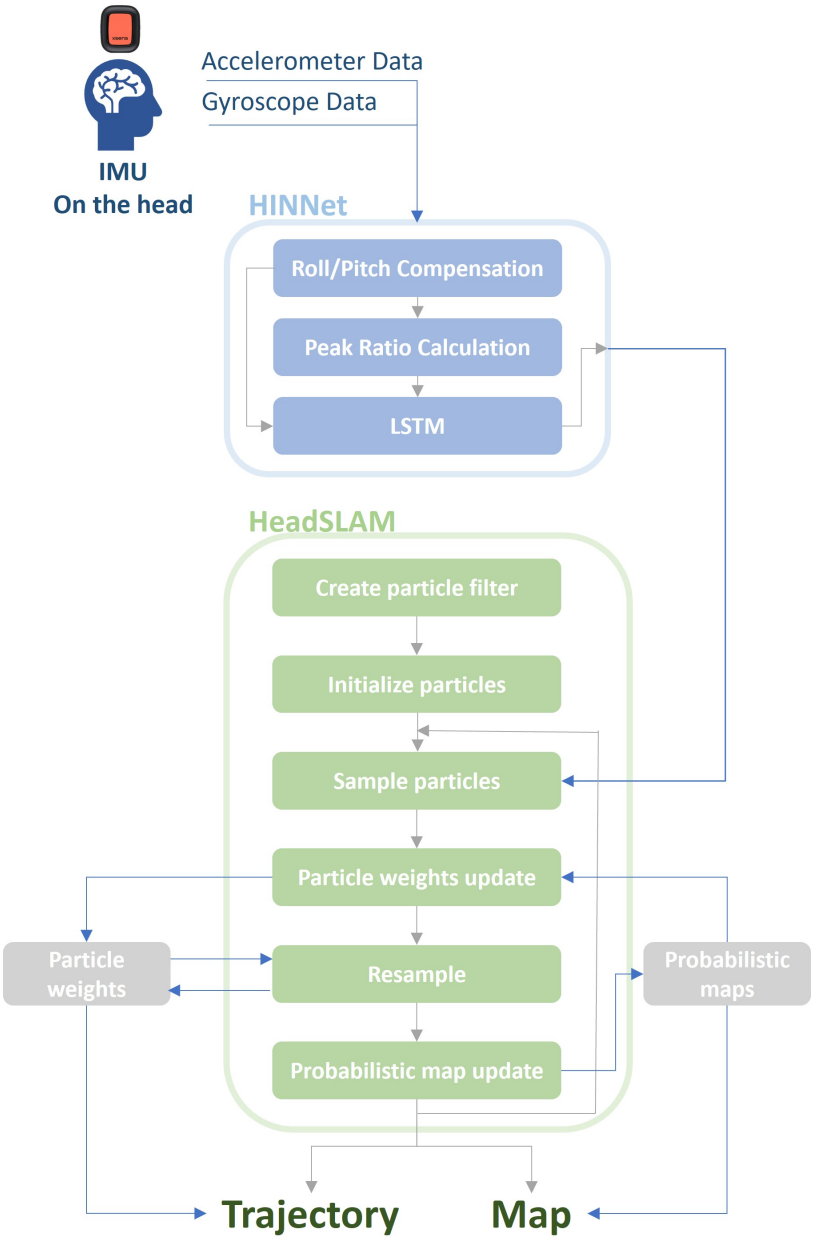
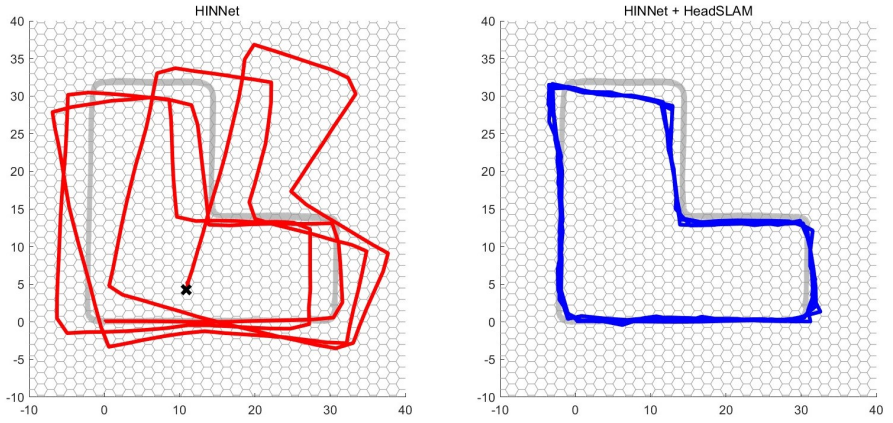
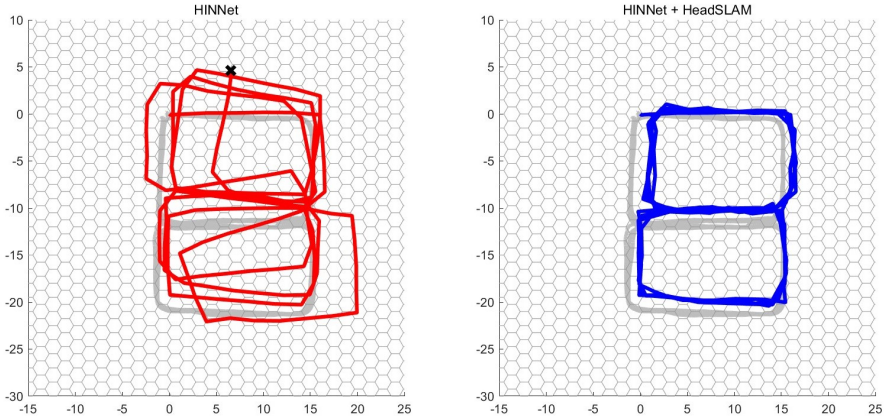


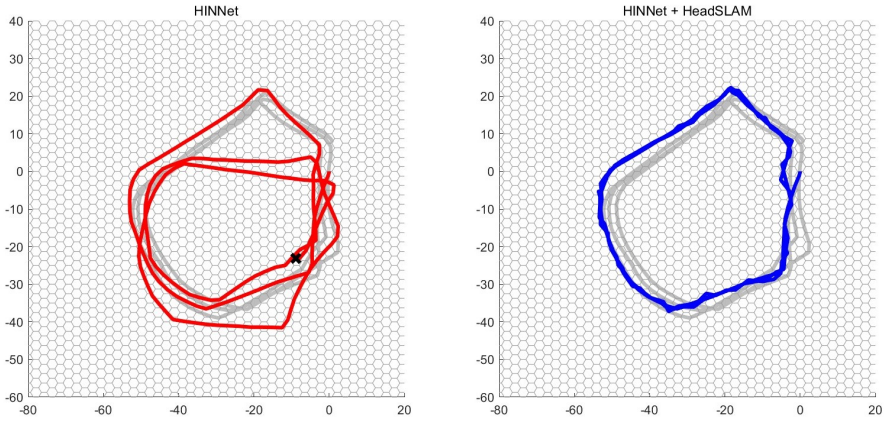
Figure 6.1: Overview of HINNet-HeadSLAM system. HINNet receives raw accelerometer and gyroscope data from the IMU on the head, and output the odometry to HeadSLAM. HeadSLAM calibrates and estimates the final trajectory. Long short-term memory is abbreviated by LSTM.



(a)



(b)



(c)

Figure 6.2: Estimated trajectories in metres. Light grey lines are ground truth. The red lines represent the trajectories generated from HINNet only. The blue lines show the results from the combination of HINNet and HeadSLAM.

Distance error rate (%): Drift of the estimated total distance.

RTE, ATE and the percentage error of the total distance of each method are summarised in Table 6.1.

Table 6.1: Relative trajectory error (RTE), absolute trajectory error (ATE), and percentage error of total distances of HINNet and HINNet+HeadSLAM.

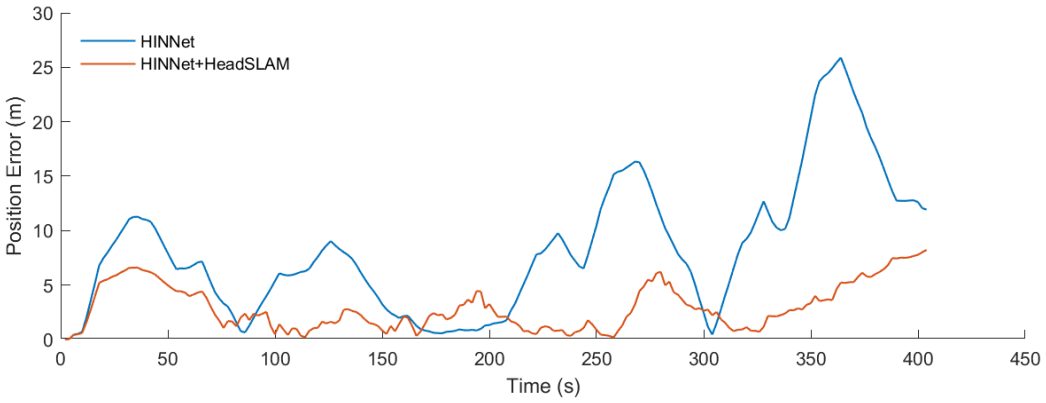
Methods	RTE (m)	ATE (m)	Distance Error(%)
HINNet	3.69	7.13	1.15
HINNet + HeadSLAM	3.52	2.69	2.19

The position error over time is presented in Table 6.3.

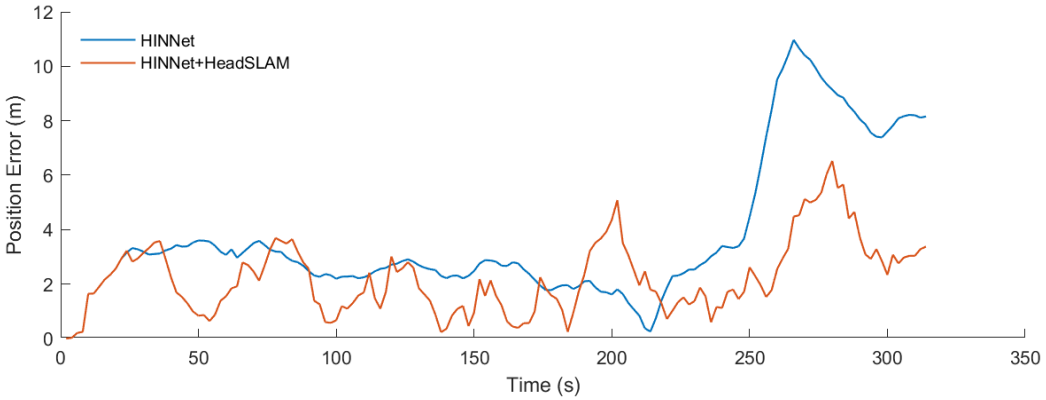
6.4 Discussion and conclusion

HINNet solved the problem of confusion between pure head rotations and walking direction change when using head-mounted IMUs. HeadSLAM solved the problem of error accumulations in long-term tracking. The combination of them could solve both problems: allowing free head rotations and meanwhile supporting long-term tracking. ATE is the RMSE of the whole trajectory (8 - 12 minutes in this study). RTE could be recognised as ATE in one minute. Both methods have similar RTEs. But the ATE of the combined method is significantly lower than that of original HINNet, which proved that HINNet + HeadSLAM is able to maintain a stable and consistent error in long-term tracking. It has also been shown in Figure 6.2, that the trajectories generated by HINNet are gradually getting farther and farther away from the ground truth due to the error accumulations and the lack of calibration. However, the trajectories estimated by HINNet + HeadSLAM remain close to the ground truth as time goes on. Figure 6.3 quantifies this difference, showing the change in position error over time. The position error of HINNet shows a general increasing trend over time, while that of the combined method keeps relatively steady as time goes on.

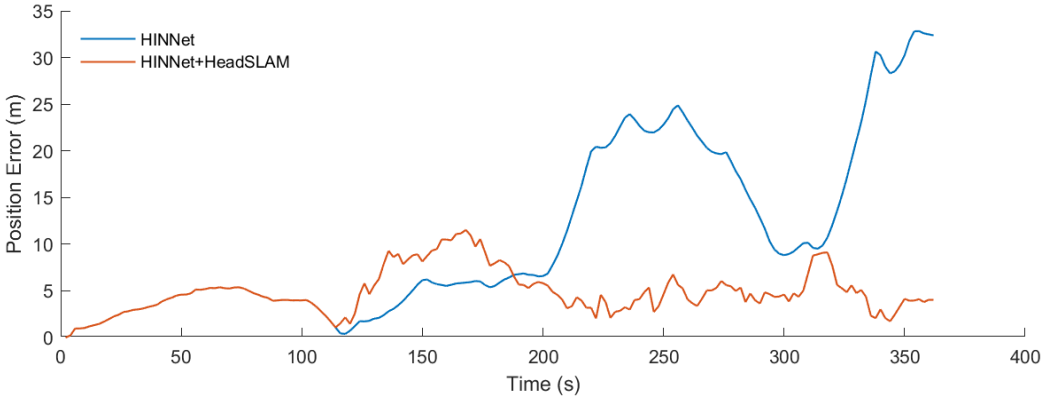
Although the proposed method has the advantage from both HINNet and HeadSLAM, it also obtained part of limitations from them. It could differentiate pure head rotations from walking direction changing as long as they do not happen at



(a)



(b)



(c)

Figure 6.3: Position errors change over time on three sample trajectories.

the same time. Overlaps of head and body rotations lead to confusion in heading estimation and need further information from other sensors or new features to maintain a correct heading direction. It should also be noted that, just like HeadSLAM, the effectiveness of this combined method only exists when walking repeatedly on a predefined restricted path, such as walking in indoor corridors or outdoor tracks for several laps. Because the calibration of HeadSLAM depends on the probability map which is updated in overlaps. Thus, the combined method is applicable for scenarios in which people cover the same path multiple times.

Another problem is that this system is still not a fully automatic system as there is still a human-tuned parameter in the particle filter, which is the deviation of the probability distribution of particles. In theory, this distribution in each step varies, as the inference of neural network contains different accuracy / uncertainty each time. The next chapter will address this problem with uncertainties estimated by the neural network, to establish a fully automatic system.

In addition to the above limitations, there are also other possible future research directions. In real world scenarios, people not just rotate their heads and walk with constant pace, they may run, jump, turn, slide, or stumble in daily activities. Users will have different body data and movement patterns. Sensors may also have different accuracy or other parameters. If a tracking system seeks extensive use in daily real scenarios, datasets with a larger scale and variety should be essential for the generalisation and robustness of the system. The next chapter will also address this problem, starting by proposing a new dataset covering walking and running at different speeds.

Chapter 7

ROCIP: RObust Continues Inertial Position tracking for complex actions emerging from the interaction of human actors and environment

As discussed in previous chapters, inertial navigation is a rapidly developing area and it is becoming more widely used than ever, due to ongoing advances in sensor technologies and tracking algorithms. At the same time, consumer-grade IMUs continue to get smaller, while also becoming cheaper to obtain. Despite developments in PDR, the application for positional tracking using IMUs is still suffering from considerable noise and biases. Various methods have been introduced in the aforementioned chapters to reduce the drift of the signal, ranging from the traditional model-based PDR to more recent machine learning based methods. However, error accumulation needs to be further reduced to allow for long-term operation of these systems. The self-contained property of inertial tracking offers a potential broad adoption across scenarios, but it also is a drawback that limits the system and makes it impossible for it to "connect" to a global reference frame.

To solve this problem, a system that could "introspect its error" and "learn from the past" is proposed. It consists of a neural statistical motion model that regresses both poses and uncertainties with DenseNet, which are then fed into RBPF for cali-

bration with a probabilistic transition map.

An inertial tracking dataset with head-mounted IMUs was collected, including walking and running with different speeds while allowing participants to rotate their heads in a self-selected manner. The dataset consisted of 19 volunteers that generated 151 sequences in 4 scenarios with a total time of 929.8 min.

It was shown that our proposed ROCIP method provided an improved performance (lower error) for localisation compared to previously described methods in the field, such as PDR and HINNET methods. ROCIP had an RTE of 4.94m and an ATE of 4.36m. ROCIP could also solve the problem of error accumulation in dead reckoning and maintain a small and consistent error during long-term tracking.

7.1 Introduction

In previous chapters, solutions for inertial tracking with head-mounted sensors have been proposed for problems with estimating position as head motions occur and the issue of error accumulations during long-term tracking. However, new research questions also emerged after using machine learning techniques to solve these issues. One question that came up is related to the uncertainty of the prediction. The ability to assess the uncertainty would provide us with some insights into how much we should "trust" the obtained outcome. Not knowing this can cause dilemmas for users as they are not able to accurately assess the risk if they decide to rely on the presented estimations. This is particularly relevant for certain use cases in which positional data might affect the safety of the user. People might need to depend on their estimated location to safely navigate their environment. A deterministic estimation result with an unknown error may cause serious problems if the information would be taken at face value. If the uncertainty of the given output could be quantified, then better risk assessments can be performed. The uncertainty can be fed into some kind of decision model that incorporates this information to determine whether to rely on it. Additionally, it can also help with fusing the sensor information with other data

sources, such as GPS, camera, etc. It should be mentioned that quantifying the uncertainty could further support the optimisation of the application of a Kalman or particle filter. Uncertainty estimation in deep learning has been discussed in previous studies [141, 90] and it could be suggested that they can provide additional value to the problem of localisation. The concept of multivariate uncertainty will be discussed further in this chapter. Not only is there a need to understand the (un)certainty of a given estimate, but there is also a need to cover more complex behaviours related to mobility. As discussed in previous chapters, the aim of implementing the developed technology in real-world scenarios requires us to consider experimental environments that are more varied. These scenarios should reflect real-world behaviour as much as possible. Currently, there is hardly any pedestrian inertial navigation method that maintains long-term robust performance when applied in the real world, despite many studies claiming that high accuracies could be obtained. The level of external validity remains limited in most cases. A key consideration that is often ignored is the unpredictability of human behaviour. More importantly, people transition from one mode of locomotion into another (e.g. from walking to jogging to running). Human motions are more complex and unpredictable during daily living [10]. In experiments conducted under laboratory conditions, most studies only require subjects to walk at a constant pace on a predefined trajectory. Yet, in the natural world people may not always walk with a constant pace, they may jog or run with different speeds, walk sideways, stop and start walking unexpectedly, or shown any other complex motion. This is also one of the reasons why the traditional PDR methods fail when efforts are made to apply them in the real world, as they were only modelling the most basic behaviours in human gait. Nonetheless, machine learning has the potential to learn these more realistic complex motions by itself without manual input. A limited number of studies have already seen this problem and started to investigate tracking with different motions. [180] tested its algorithm in walking, fast walking and running. [181] got experimental results for various human motions including straight line walking, circular walking, side stepping, backward walking, running, and

climbing stairs. [108] collected data that show a variety of activities including walking, standing still, organising the kitchen, playing pool, going up and down stairs.

Nonetheless, error accumulation remains an inevitable problem in dead reckoning if there is no opportunity for re-calibration with the true positions in the real world. To overcome this complication, common calibration approaches have been used, which include, for example, the application of a camera or RFID. It is obvious that these solutions introduce external infrastructures and/or extra devices. It provides a solution that is no longer self-contained and reduces the ability to widely utilise this kind of technology. As for using IMUs only, ongoing calibrations for long-term monitoring with the real world is impossible. However, there exists a way to eliminate the error accumulation by calibrating the system with its own prior trajectories. Similar ideas have been used in [4][70]. However, these studies were conducted only in situations where the volunteers were walking at a constant pace. Furthermore, the fixed uncertainty parameters needed to be manually set, which reduce their robustness and generalisability across various users, motions, and environments.

In this chapter, ROCIP, an inertial navigation system for a head-mounted IMU, is proposed. It uses a probabilistic DenseNet model to estimate both pose and uncertainty in a mixed supervised and unsupervised way and is tightly coupled with an RBPF for optimal estimation. The system is tested for "long-term" tracking with various head motions and different speeds of walking and running. The main contributions are as follows.

- We propose a network specially designed for inertial navigation with head-mounted IMUs, based on special input features and outputs that include both displacement and uncertainty.
- We propose a complete state estimation system combining the neural network with the RBPF to control the error accumulations in long-term localisation.
- We introduce an inertial tracking dataset with head-mounted IMUs, with a

total length of 929.8min, 151 sequences, and 19 subjects, which includes human walking and running with different speeds and random head rotations.

ROCIP was compared with two other well-defined methods for head-mounted sensors, and the absolute and relative errors were computed across all three approaches.

7.2 Methods

The system is comprised of two main components: (i) network and (ii) filter. Fig.7.1 shows the main structure of the system. The raw data from the IMU sensor was first transformed into a normalised coordinate system in which the z axis is aligned with the gravity direction, while there is no gravity component on the normalised x axis and the y axis. Then the peak ratio features were calculated, which were specifically designed for tracking with head mounted devices. This ratio helps distinguish between head motion patterns during walking[71]. The peak ratios and normalised IMU data form the input into the subsequent neural network.

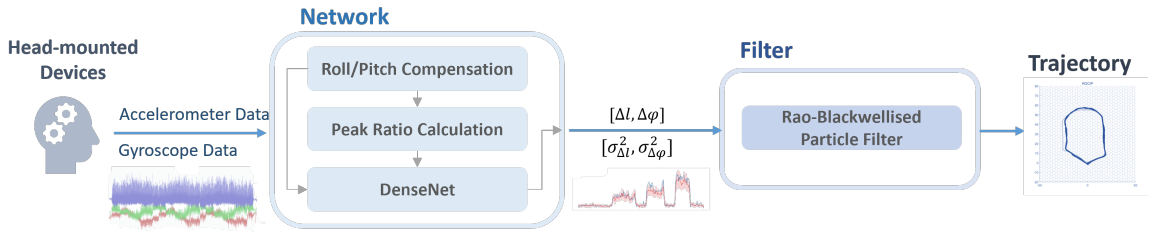


Figure 7.1: Overview of the system. Raw accelerometer and gyroscope data form the input from the head-mounted IMUs. The position and uncertainties are estimated by DenseNet and form the input for the particle filter.

7.2.1 DenseNet

Dense Convolutional Network (DenseNet)[76] is an improved version of CNN. CNN were proposed in the 1980s and have been the dominant machine learning approach in image processing tasks for more than 10 years. More recently, improvements in computer hardware have enabled the training of truly deep CNNs. Yet, the previously

discussed problem of gradient vanishing emerges as the layers are getting deeper. ResNet[61] was proposed to solve this problem by passing signals from one layer to the next via identity connections, and it has been used in RoNIN [62] for inertial navigation. DenseNet connects all layers directly with each other to ensure maximum information flow between layers in the network. It was evaluated on four highly competitive object recognition benchmark tasks and obtained significant improvements over the state-of-the-art with less computational effort. In this research, a DenseNet-BC 100 was adopted as the backbone, which is a DenseNet with bottleneck layers and compression, with 100 convolutional layers. It has three dense blocks. Each block has 16 layers, with each layer having a bottle neck layer and a convolutional layer. DenseNet has been adapted to fit this research, with the layers reshaped to process 1D information. The input and output have also been redesigned for this research. The input could be recognised as a 1D image with a shape of $window\_size \times 1$ and the channels are input features, which could be represented as $(a_{norm}, \omega_{norm}, P_{ratios})_{8*120}$. It outputs the moving distance and orientation variation $(\Delta l, \Delta\varphi)_2$, with the uncertainty of both in a time window. Fig. 7.2 shows the architecture of the adapted neural network.

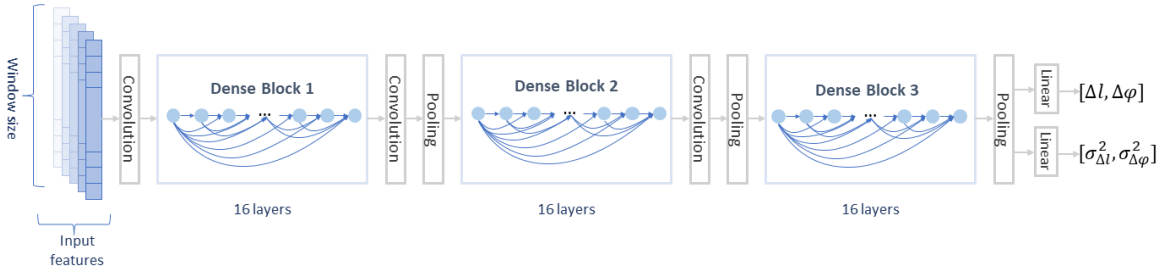


Figure 7.2: Architecture of DenseNet-BC 100 that is used in this chapter.

7.2.2 Uncertainty estimation

In DenseNet, the model outputs are the estimated polar vectors and variances:

$$[\hat{\boldsymbol{y}}, \hat{\boldsymbol{\Sigma}}] = f_\theta(\boldsymbol{x}) \tag{7.1}$$

where $\hat{\boldsymbol{y}} = [\Delta\hat{l}, \Delta\hat{\varphi}]$ and $\hat{\boldsymbol{\Sigma}} = diag(\sigma^2_{\Delta l}, \sigma^2_{\Delta\varphi})$.

Assuming that $\boldsymbol{y}$ follows a multivariate normal distribution:

$$\boldsymbol{y} \sim \mathcal{N}_2(\hat{\boldsymbol{y}}, \hat{\boldsymbol{\Sigma}}) \tag{7.2}$$

with a probability density function:

$$p(\boldsymbol{y}|f_\theta(\boldsymbol{x})) = \frac{1}{\sqrt{(2\pi)^2|\hat{\boldsymbol{\Sigma}}|}} \exp\left(-\frac{1}{2}(\boldsymbol{y}-\hat{\boldsymbol{y}})^T\hat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y}-\hat{\boldsymbol{y}})\right) \tag{7.3}$$

The optimal neural network weights θ^* could be approximated by log-likelihood:

$$\begin{aligned}\theta^* &= \arg\max_\theta p(\boldsymbol{y}|f_\theta(\boldsymbol{x})) \\ &= \arg\min_\theta -2\log p(\boldsymbol{y}|f_\theta(\boldsymbol{x})) \\ &= \arg\min_\theta \log|\hat{\boldsymbol{\Sigma}}| + (\boldsymbol{y}-\hat{\boldsymbol{y}})^T\hat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y}-\hat{\boldsymbol{y}})\end{aligned} \tag{7.4}$$

To make regression more stable, s_l and s_φ are predicted by the neural network instead of $\sigma^2_{\Delta l}, \sigma^2_{\Delta\varphi}$, where $s_l = \log\sigma^2_{\Delta l}$ and $s_\varphi = \log\sigma^2_{\Delta\varphi}$ [90]. Thus, the loss function of neural network could be defined as:

$$\begin{aligned}\mathcal{L} &= \log|\hat{\boldsymbol{\Sigma}}| + (\boldsymbol{y}-\hat{\boldsymbol{y}})^T\hat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{y}-\hat{\boldsymbol{y}}) \\ &= \log\sigma^2_{\Delta l} + \log\sigma^2_{\Delta\varphi} + \frac{(\Delta l - \Delta\hat{l})^2}{\sigma^2_{\Delta l}} + \frac{(\Delta\varphi - \Delta\hat{\varphi})^2}{\sigma^2_{\Delta\varphi}} \\ &= s_l + s_\varphi + \frac{(\Delta l - \Delta\hat{l})^2}{\exp s_l} + \frac{(\Delta\varphi - \Delta\hat{\varphi})^2}{\exp s_\varphi}\end{aligned} \tag{7.5}$$

The loss function shows that the network training contains a mixture of supervised and unsupervised parts. The estimated displacement and rotation $\Delta\hat{l}, \Delta\hat{\varphi}$ have their corresponding ground truth $\Delta l, \Delta\varphi$, so they could be optimised in a supervised way. While there is no ground truth for estimated uncertainty, the estimation of uncertainty is optimised in an unsupervised way.

The loss function can be converged appropriately. Separate the loss function into two parts for displacement Δl and rotation $\Delta\varphi$:

$$\mathcal{L}_{\Delta l} = \log \sigma_{\Delta l}^2 + \frac{(\Delta l - \Delta \hat{l})^2}{\sigma_{\Delta l}^2}$$
$$\mathcal{L}_{\Delta \varphi} = \log \sigma_{\Delta \varphi}^2 + \frac{(\Delta \varphi - \Delta \hat{\varphi})^2}{\sigma_{\Delta \varphi}^2}$$

Use $\mathcal{L}_{\Delta l}$ as the example. If Δl and $\Delta \hat{l}$ are known, as a function of $\sigma_{\Delta l}$, $\mathcal{L}_{\Delta l}$ is convex. The minimum of it will appear where the derivative equals 0:

$$\begin{aligned} \frac{\partial \mathcal{L}_{\Delta l}}{\partial \sigma_{\Delta l}} = \frac{2}{\sigma_{\Delta l}^*} - \frac{2(\Delta l - \Delta \hat{l})^2}{\sigma_{\Delta l}^{*3}} &= 0 \\ 1 - \frac{(\Delta l - \Delta \hat{l})^{*2}}{\sigma_{\Delta l}^2} &= 0 \\ \sigma_{\Delta l}^{*2} &= (\Delta l - \Delta \hat{l})^2 \end{aligned} \tag{7.6}$$

Therefore, during training, $\sigma_{\Delta l}^2$ will try to converge to $(\Delta l - \Delta \hat{l})^2$, which perfectly matches its property as a variance. $\mathcal{L}_{\Delta \varphi}$ converges in the same way.

7.2.3 Rao-Blackwellised particle filter

Path inference utilises the Rao-Blackwellised particle filter of the FastSLAM algorithm[118], which decomposes the SLAM problem into a pedestrian localisation problem and a mapping problem conditioned on the estimated poses. Based on the conditional independence property of the SLAM, the posterior pose could be factored as Equation 4.3, where $\mathbf{P}$ and $\mathbf{M}$ represent the pose and the map estimated in the current Bayes inference step. And $\mathbf{Z}_k$, which is intrinsically equal to the $\hat{\boldsymbol{y}}$ estimated from the previous neural network, now recognised as a noisy measurement of the difference between $\mathbf{P}_{k-1}$ and $\mathbf{P}_k$. The pose in time step k could be estimated recursively from the last with the first pose $\mathbf{P}_0$ set arbitrarily, with Equation 4.4. The first factor in it, $p(\mathbf{Z}_k|\mathbf{P}_{k-1:k})$, is the likelihood function of the pose difference estimation in the current part, which is expected to approximate to the real value $\boldsymbol{y}$. Therefore, it is proportional to $p(\boldsymbol{y}|\hat{\boldsymbol{y}})$, which is the posterior and also proportional to Equation

7.3. This likelihood factor is used to sample particles in each step from a Gaussian distribution $\mathcal{N}_2(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, where

$$\boldsymbol{\mu}_k = \hat{\boldsymbol{y}}_k = [\Delta l_k, \Delta\varphi_k]$$
$$\boldsymbol{\Sigma}_k = diag(\sum_{i=0}^{k} \sigma^2_{\Delta l_i}, \sum_{i=0}^{k} \sigma^2_{\Delta\varphi_i})$$

The second factor in Equation 4.4, $p(\mathbf{P}_k|\mathbf{P}_{0:k-1})$, is the pose transition probability, which is used in the particle weight update (Equation 7.7). In this section, the two-dimensional space was divided into a grid of adjacent hexagons of a given radius. Each hexagon $\tilde{h}$ has six edges. Each time a particle makes a transition across an edge $\tilde{e}$ will be counted up and recorded in the number of transition times $N^{\tilde{e}}_{\tilde{h}}$. For each hexagon, the number of transition times of six edges is recorded in a local map $\mathbf{M}_{\tilde{h}}$, which is a vector of length 6. All $\mathbf{M}_{\tilde{h}}$ of every hexagon comprise the probabilistic transition map $\mathbf{M}$. When updating the map after each step, if an edge was crossed, the transition times of this edge will increase in both the incoming and outgoing hexagons. The transition times recorded in the map are then used in the particle weight update:

$$w^i_k \propto p(\mathbf{P}_k|\mathbf{P}_{0:k-1}) \cdot w^i_{k-1} \propto \frac{N^{\tilde{e}}_{\tilde{h}} + \alpha^{\tilde{e}}_{\tilde{h}}}{N_{\tilde{h}} + \alpha_{\tilde{h}}} \cdot w^i_{k-1} \tag{7.7}$$

where w^i_k denotes the weight of the i-th particle at step k, $N_{\tilde{h}}$ is the sum of the crossed times of all 6 edges of the outgoing hexagon $\tilde{h}$ in i-th particle's map counters. $\alpha^{\tilde{e}}_{\tilde{h}}$ and $\alpha_{\tilde{h}} = \sum_{e=0}^{5} \alpha^{\tilde{e}}_{\tilde{h}}$ are the prior counts.

With the above functions, the sequential Bayesian estimation in the particle filter will be conducted in a loop of particle sampling, weight updating, resampling, and probabilistic transition map updating.

7.3 Data collection

Xsens Dot sensors (Xsens Technologies BV, Enschede, Netherlands) were used in the experiments with a recording frequency of 60Hz. An Xsens Dot was firmly attached

to the forehead using a strap that was adjusted for each subject. A phone was also firmly attached on the chest with straps to provide reference data. An application developed in Unity 3D game engine (Unity Technologies, San Francisco, CA, USA) based on Google ARCore was installed in the phone for ground truth collection.

The data collection was conducted in four different environments with different path sizes and trajectories to ensure more varied patterns were captured. This should help to clarify the eternal validity of the results obtained. The trajectories included straight routes, curves, and turnings with different angles, which increased the complexity and diversity of the tests. There were 19 volunteers participating in the experiments with 11 males and 8 females. The participant's ages ranged from 22 to 50 years. Ethical approval was obtained from the University Ethics Committee and this experiment, which was part of a larger study (R70833/RE001). Subjects were asked to walk and run with different speeds on the predetermined trajectories, whilst rotating their head in a random manner.

A total of 151 sequences of data were collected, with a total duration of 929.8 minutes. Each dataset contained both the IMU data from the Xsens Dots, as well as the ground truth trajectory and orientation data from the phone.

7.4 Results

The proposed methods were trained at the University of Oxford Advanced Research Computing (ARC) facility[138], with the dataset split into a training, test, and validation set. Two published methods for tracking with a head-mounted IMU were chosen for comparison: a PDR method [68] that uses peak detection to detect steps, with complementary filter to determine orientations and Weinberg model to detect step lengths; and HINNet [71], which uses Bi-LSTM and peak ratio features.

Figures 7.3 and 7.4 show the output of probabilistic neural network. The motion represented in Figure 7.3 is walking and running in various speed. While Figure 7.4 shows the motion with only walking with constant pace.

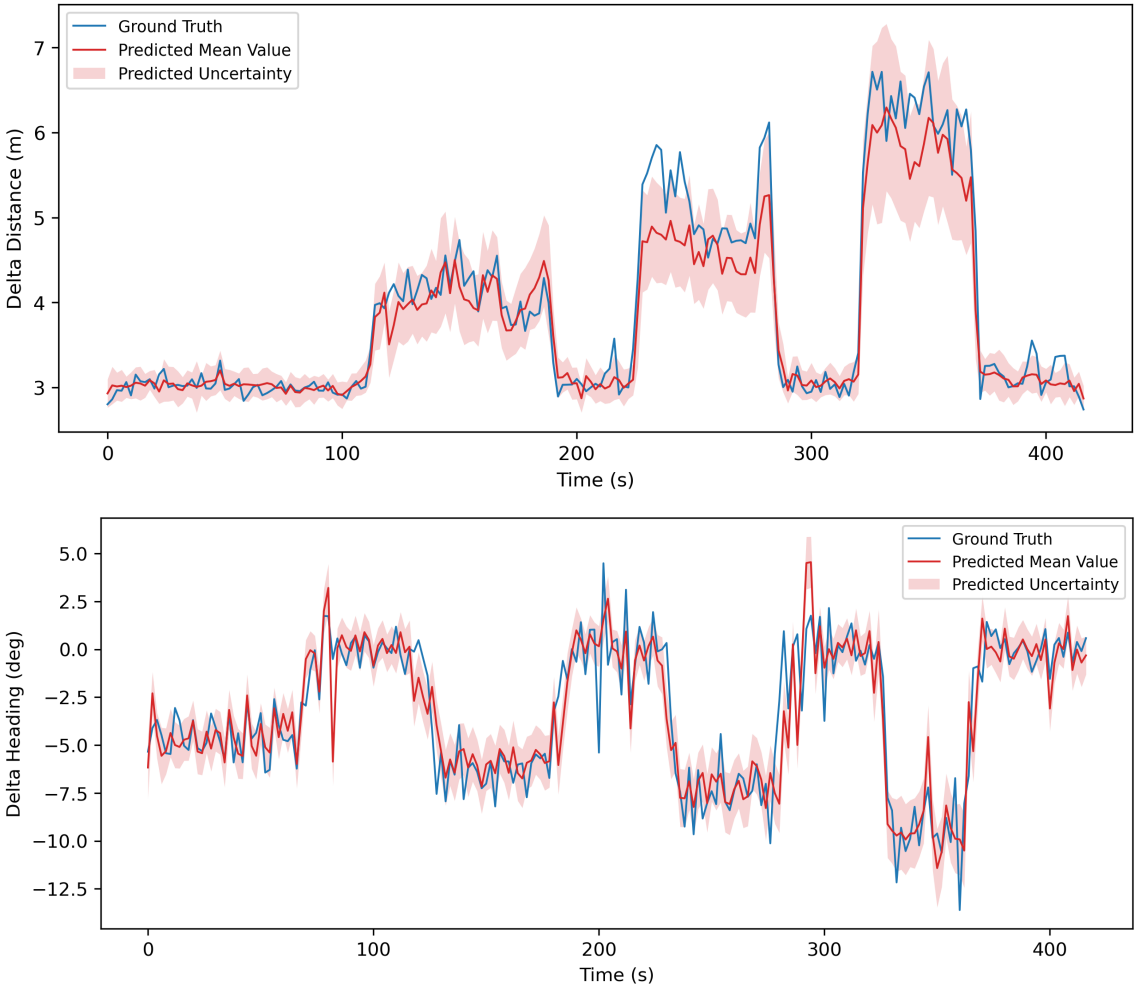


Figure 7.3: Estimated delta distance and heading from probabilistic DenseNet – **walking and running in various speeds**. Blue lines represent the Ground Truth. Red lines are mean values estimated by DenseNet. Light red zones show the estimated uncertainty σ.

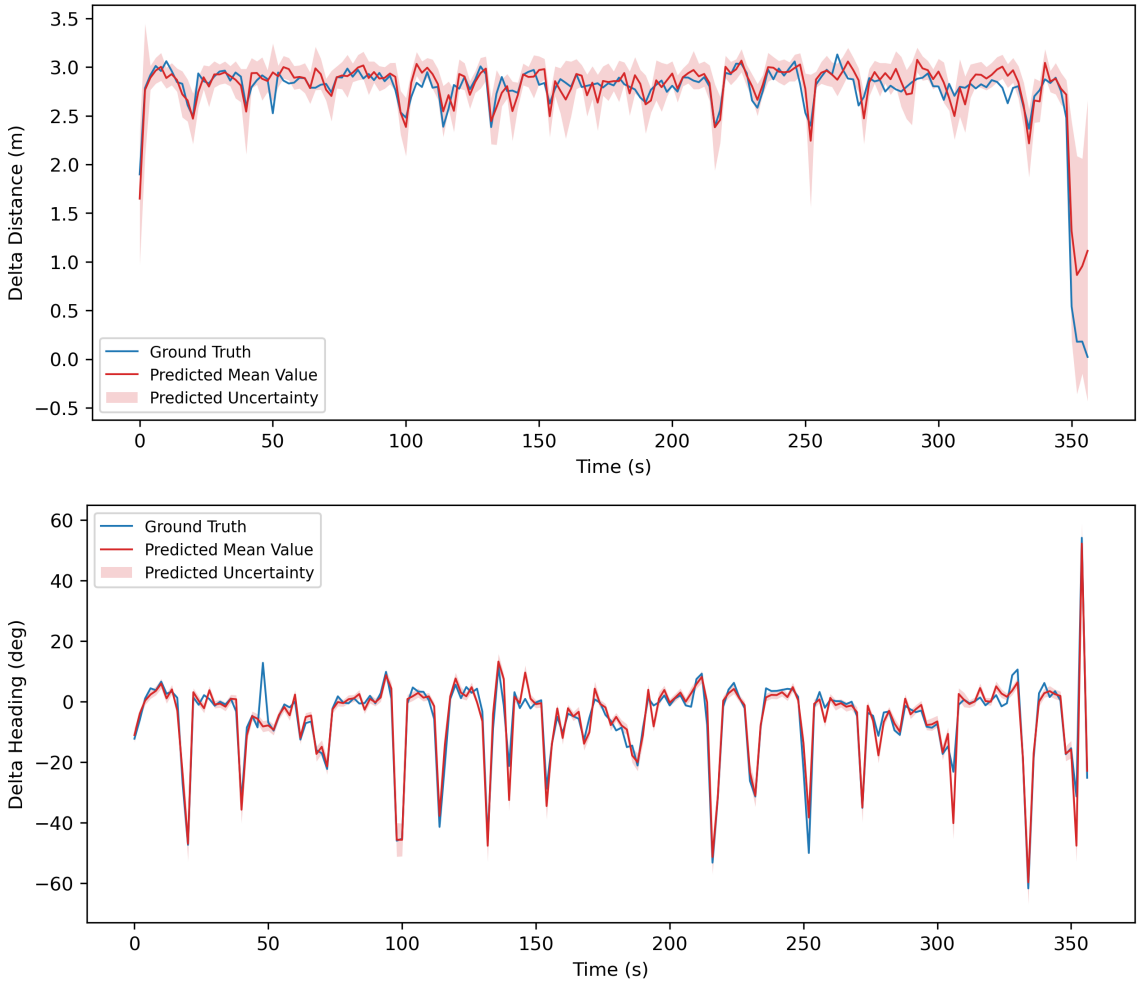


Figure 7.4: Estimated delta distance and heading from probabilistic DenseNet – **walking with constant pace**. Blue lines represent the Ground Truth. Red lines are mean values estimated by DenseNet. Light red zones show the estimated uncertainty σ.

RTE, ATE and percentage error of the total distance covered for each method are summarised in Table 7.1.

Table 7.1: Relative trajectory error (RTE), absolute trajectory error (ATE), and percentage error of total distances of proposed and comparing methods.

Methods	RTE(m)	ATE(m)	Distance Error (%)
PDR	7.99	13.85	7.26
HINNet	5.46	13.41	0.46
ROCIP	4.94	4.36	2.81

The estimated trajectories of different methods in the 4 scenarios are shown in Figure 7.5.

7.5 Discussion and conclusion

This study presented ROCIP, which fuses a DenseNet based neural statistical motion model and a RBPF for long-term inertial tracking with stable errors. The proposed system was trained and evaluated with a newly collected dataset. It outperformed the comparative methods in both ATE and RTE, and generated estimated trajectories that are much closer to the ground truth.

The result also shows the robustness of the proposed system in long-term tracking. ATE is the RMSE of the total trajectory, which is around 7-12 minutes in this study. RTE could be considered as the ATE in one minute. The ATEs of the PDR and HINNet were around a factor of two larger than their RTEs, which suggests that their errors are gradually accumulating as time increases. However, the proposed ROCIP has similar RTE and ATE, indicating that ROCIP error remains relatively stable even during long-term tracking. Fig. 7.5 also provides further evidence of this. The trajectories generated by HINNet and PDR are gradually getting further and further away from the ground truth because of the error accumulations and the lack of calibration. However, the trajectories estimated by ROCIP remain close to the ground truth as time goes on.

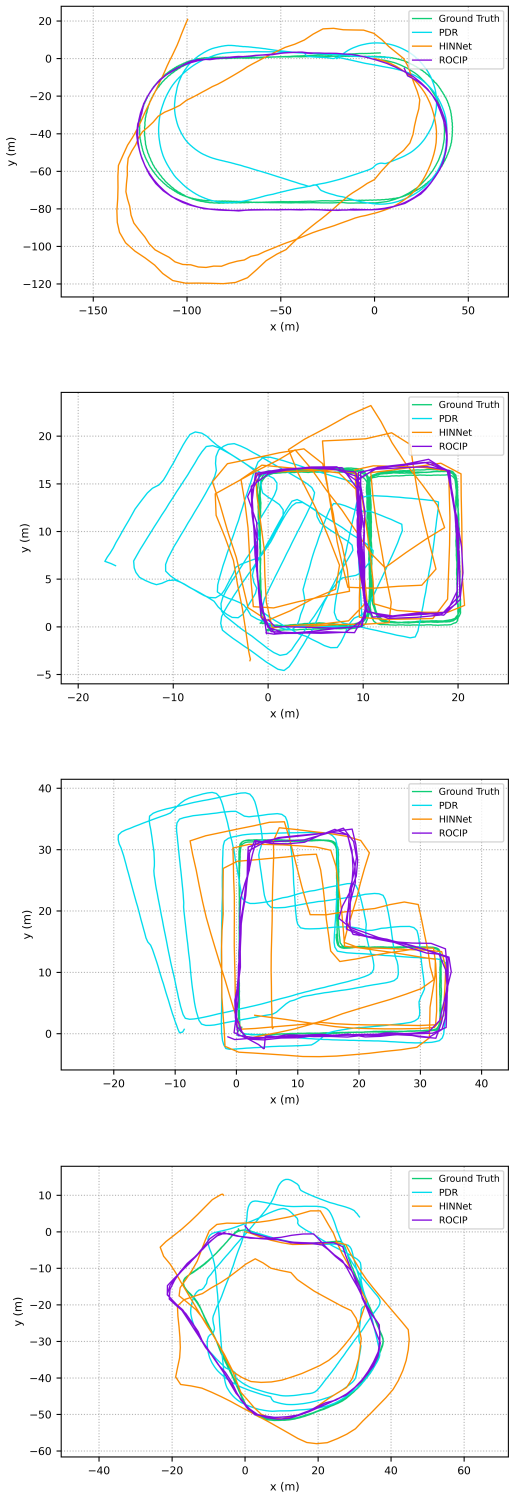


Figure 7.5: Estimated trajectories of different methods on four different paths. The green lines represent the ground truths. Orange lines show trajectories estimated by HINNet. Blue lines are the trajectories generated by the PDR. Purple lines are proposed ROCIP trajectories.

The probabilistic neural network also proves its ability to give reasonable uncertainty estimations. Figure 7.3 shows the estimation result on walking and running at various speeds. 4 walking periods (0-110s, 190-220s, 290-320s, 370-420s) show similar speeds around 3m/s, while 3 running periods (110-190s, 220-290s, 320-370s) show increasing speeds (4m/s, 5m/s, 6m/s). With the speed increasing, the estimated uncertainty (light red region in the figure) shows a trend to increase as well, which infers that the estimation of neural network has larger uncertainty for higher speed. Figure 7.4 shows the estimation result for walking with a constant speed around 2.5-3m/s. The uncertainty also remains in a relatively constant range, and clearly much smaller than that of the running periods in Figure 7.3. It is probably because running introduces more complex motions and higher noise to the network, thus it is reasonable to result in larger uncertainties.

This chapter also proposed a new larger dataset including more subjects, more scenarios, more motions, and speeds. Comparing HINNet results on this new dataset and on the dataset proposed in Chapter 5, in Table 7.1 and Table 5.2, it shows worse performance in a larger dataset, which seems counterintuitive. But it is because the new dataset includes more complex motions such as walking and running at different speeds, and the scenarios are much larger (150*80m) comparing to previous small maps (15*20m). Although accuracy decreased, robustness improved, allowing the system to work with more complex motions and environments.

These kind of approaches can also be generalised to other research areas, such as the tracking of legged robots. The models are also strong candidates for monitoring, e.g., professionals or robots during search and rescue missions, in mining or agriculture [7]. As these environments could require robots that can cope with obstacles, rough outdoor terrain, and steps, which is difficult for wheeled robots. Using IMUs with algorithms similar to human tracking could serve as a component in the fusion of its multi-sensor localisation.

The findings presented here show the ability to use widely available IMU for localisation and monitoring of behaviour with relatively stable errors.

Chapter 8

Conclusions and future work

8.1 Conclusions

This thesis investigated human inertial tracking with wearable devices, especially focusing on being able to apply this to head-mounted sensors.

The systematic review in Chapter 2 provided a comprehensive overview on the current research status of human localisation technology with self-contained wearable sensors, summarised characteristics of all papers in this area, and conducted a quality assessment on them. It showed that only a limited number of papers used head-mounted sensors. This despite that the head could be an appropriate and promising place for wearable sensors to be attached to. It is likely that the additional movements of the head made this location a more challenging sensor placement option. These challenges are further discussed in the subsequent chapters.

Chapter 3 then proposed an inertial navigation solution especially designed for head-mounted IMUs. It was based on the traditional model-based PDR method with a step-and-heading system, which used peak detection to detect steps, a complementary filter to estimate the heading, and Weinberg model to infer the step length. A dataset of 24 sequences was collected with 3 subjects and 2 devices (glasses and instrumented mouthguard with IMUs). Participants were asked to walk in a rectangle with a total length of 68m. Proposed PDR method outperformed the other two comparing methods, with an end-to-end error of 0.88m and total distance error of 2.10%. This indicated that head-mounted systems could perform in a similar fashion as other

sensor placements.

The focus then shifted to the error accumulation problem for long-term tracking. This is a problem of inertial tracking independent of sensor placement, as it is impossible to calibrate the system using a reference position during tracking with only IMUs. However, inertial tracking could leverage information related to its previous trajectory, as it has been shown that the PDR is accurate in the anterior period, which was then further described in Chapter 4 as HeadSLAM. HeadSLAM conducted optimal estimation on the pose estimated by PDR, with a Rao-Blackwellised Particle filter. A dataset of 60 sequences was collected, with 7 subjects and 2 devices (glasses, cap). In this case, participants walked randomly during two scenarios (indoor, outdoor) following a provided template. Each collected sequence was around 10 minutes ($\sim$ 840m). HeadSLAM outperformed original PDR method with a RMSE of 0.34m. It could also generate a probabilistic map for future localisation. This result further confirmed the possibility of using head mounted sensors for localisation.

The difficulty of identifying a head movement from a locomotion-related motion still needed to be further explored, as it forms a fundamental problem for inertial tracking with head-mounted devices. Motions of the head increases the complexity of IMU signals and affect the inertial tracking with traditional model-based PDR methods (e.g. head rotations lead to wrong heading estimations). Chapter 5 proposed HINNet, which disposed of traditional model-based PDR method, instead using a neural network and special "peak ratio" features, to solve the problem of tracking with unpredictable head motions. The dataset in this chapter was collected from 8 subjects over 3 scenarios, accumulating to a total of 79 sequences of around 528 minutes. HINNet was shown to successfully solve the head motion separation problem. Chapter 6 then added a SLAM layer after HINNet and provided some long-term tracking robustness to it. This was now becoming a system that started to become more applicable in the real world.

Nonetheless, the gait patterns explored so far did not really vary much within a subject. The effect of locomotion variation needs to be investigated in order to

truly make these systems robust. Different people walk or run with different speeds and transitions can occur at random times. It is important that inertial tracking is able to cope with that. Additionally, uncertainty estimation in tracking with machine learning is also worth researching, as it would feed into further sensor fusion or optimal positional estimation with SLAM. ROCIP was then proposed in Chapter 8 using a DenseNet to estimate both pose and uncertainty, which were then inputted into the SLAM layer for re-calibration during long-term tracking. It provided a robust method even with volunteers walking or running in different speeds. The dataset used in this chapter was collected from 19 subjects walking and running at different speeds in 4 scenarios, with a total of 151 sequences lasting 929.8 min. ROCIP has proved to be able to handle random head movements as well as changes in walking patterns, and could maintain a stable small error during the long-term tracking assessment.

8.2 Discussion

This thesis fills the gap in the research of low-cost head-mounted human tracking, with papers focused on different aspects, together solving the problem of long-term, robust pedestrian localisation with a consumer-grade head-mounted IMU. Before, wearable tracking relied largely on sensors on the foot, smartphone, and waist, with additional requirements for attaching specific sensors to body parts which rarely wear any device or are not convenient to carry all the time like the phone. Now people can track themselves with the device they already wear everyday without extra effort, such as the earphone, glasses, smart mouthguard, etc. Possible applications include: tracking older adults living alone with the sensors in removable dentures, hearing aids, or other daily worn devices, to help them live a healthier lifestyle and in case of emergency; tracking atheletes with smart mouthguards to help increase performance and estimate fatigue level; providing a dead reckoning tracking solution for hikers to know their relative position during hiking in wilderness.

In this thesis, different state estimation methods have been covered, which could

be roughly divided as model-based and data-driven machine learning methods. Due to the simplicity and ease of implementation, Kalman filter has been wildly used. EKF has been implemented in this thesis for its extended ability to deal with nonlinear systems. But the linearisation process will lead to errors in the nonlinear system due to the calculation of Jacobian matrix, resulting in the accuracy decreasing in the final state estimation. Sigma point Kalman filter (SPKF) avoids the linearisation errors by directly propagating the probability distribution through the non-linear functions. This makes it more accurate for highly non-linear systems and can provide better estimates of the state and its uncertainty. It includes UKF, central difference Kalman filter (CDKF), square-root unscented Kalman filter(SRUKF), etc. The Cubature Kalman Filter (CKF) is another variant of the Kalman Filter for nonlinear systems, which avoids linearisation errors by directly integrating nonlinear functions into the filter framework. So, besides the EKF used in the thesis, there are more kinds of Kalman filters that may achieve better nonlinear estimations. But when the system becomes more complex and challenging to obtain system models, another kind of state estimation method shows its advantage by random sampling, calculating the conditional probability of the system state and realizes the conditional probability transfer through Bayes' theorem, which is sequential Monte Carlo method or particle filter method. It has been used in HeadSLAM and ROCIP for trajectory calibration. Particle filter has the advantage of its wide applicability for any complex system and overcoming the problem in EKF or other Gaussian filters being sensitive to initial values. However, its larger calculation amount brings challenges to real-time applications like real-time tracking on edge devices.

Data-driven state estimation methods are based on learning of neural networks, which is also adopted in this thesis, LSTM and CNN (DenseNet). It does not require other prior physical information or model assumptions, learning the models from the hidden relations in the data capable of fitting almost any nonlinear relationship. Chapters 5 and 6 have proven their excellence compared to model-based methods. But one of the biggest problems faced by data-driven learning based state estimation

methods is the complex noise in the system. One of the solutions is considering the noise characteristics in the input, the Bayesian training process. It learns the distribution of weights and bias rather than the fixed values, to model the uncertainty of the system, and thus improves the reliability of the results. Although data-driven methods have the generalisability to fit any system, they ignored the established physical knowledge of the system but only relies on data, the performance of which will be easily affected if the data are insufficient or too noisy.

State estimation methods based on hybrid models and data-driven algorithms could benefit from both. Chapter 7 tries this hybrid way by coupling neural network with particle filter. There are other research studies focus on Kalman filter and neural network hybrid. For example, research [145] used augmentation of the radial basis function (RBF) neural network with error-state Kalman filter (ESKF) for underwater vehicle multi-sensor localisation. Research [188] used the learning Kalman network (LKN), which achieves the non-linearity of the observation model and the transition model through deep neural networks, for monocular visual odometry estimation. Similar methods also have potential applications in human tracking tasks.

The ability to use head-mounted sensors that are integrated into everyday objects to track position of people and groups can greatly improve a range of different operations. It allows for accurate behavioural tracking in anything from sports, industry, and robotics. This work opens up the ability for smart objects (such as clothing and assistive technologies) to be used in new ways to increase safety, security, and health. It can be envisioned that these techniques will become as ubiquitous as mobile phones are at the moment.

8.3 Future work

The current research of robust localisation with wearable devices could be extended by exploring the following aspects.

Large comprehensive motion model.

Recently, ChatGPT was introduced and quickly gained recognition for its comprehensive replies and eloquent solutions in various fields of expertise. It is mostly attributed to GPT-3, which was the largest language model with 175 billion parameters and a training dataset with 499 billion tokens[17]. Prior to GPT-3, the first GPT launched in 2018 used 117 million parameters, and the second version GPT-2 released in 2019 took a huge jump to 1.5 billion parameters, then the number was increased to 175 billion in GPT-3. The recently introduced GPT-4[121] shows better performance and, of course, with a larger model. Multimodal tasks also benefit from larger models. Text-to-image generation models are also developing rapidly, with models and training data becoming extremely larger. OpenAI image generation model "DALL-E 2" a 3.5 billion-parameter model trained on approximately 650 million images. DALL-E 3 [12] is the newest release, which could understand significantly more nuance and detail than previous systems. The text-to-video model is another kind of multimodal task that is greatly improved by large models. The recent release of OpenAI, Sora [16], is capable of generating a minute of high-fidelity video, which is their largest model with large-scale training. Its results suggest that, to build general purpose simulators of the physical world, scaling models is a promising path. One point worth noting is that, during the Sora training, AI is learning to understand and simulate the motions in the real physical world, which contains great potential in AI understanding human motions, activities, and interactions with environments, for various applications.

It all proved that an immense deep learning model has the potential to represent and capture extremely complicated information and patterns, and could generate more diverse outputs for a variety of applications. Reflecting again on deep learning in human motion tasks, it is promising and reasonable that a motion model with a similar size could fulfil almost any tasks related to human motions, such as tracking, motion recognition, behaviour analysis, intention prediction, etc. In practical operation, training such a huge model is not a problem because of the development of computing

hardware and the previous experience in training largest language models. But the problem also arises that when model size increases, the running time will increase. Human motion analysis in real time may be challenged if the calculation time is longer than the sampling time of each sensor data group. The current cutting-edge models for text or images are training and running on GPUs with high speed. However, human motion tracking with wearable sensors is usually calculated directly on edge devices like the phone or wearable devices. The calculation speed is impossible to compete with GPUs. Meanwhile, it also suffers from the problem of limited battery capacity. More complex models will lead to more energy consumption, thus decreasing the battery life of the device, which could be very inconvenient for users. Besides the problem of calculation time and energy consumption, there is also a question to think about. Do we really need it to be that accurate? If a similar size of the model is used in inertial motion trackings, it is possible that the error could be reduced to extremely small, like only several centimetres. But in what kinds of scenario do we need such high accuracy? One of the possible applications is the human motion analysis during operating instruments like people operating precise machines in factories. In some other scenarios, such as tracking human position at home or in even larger spaces like airports, errors around or less than 1 m would be enough.

Data are another important aspect that needs to be focused on. Large models require large dataset for training. Huge human labelling work is essential for the datasets. Companies that focus on large models for texts or images usually crowd-source data labelling or even hire people for labeling. As for the motion data, the activities can be labelled by human, but the trajectory groud truth can not but need more accurate tools to collect. Montion capture systems like Vicon (Vicon Motion Systems Ltd., Oxford, UK) are usually used for indoor groudtruth generation. However, the activity range is limited in the Vicon room, and it can not be used for outdoor activities. So, the way to get trajectory ground truth is also need to be considered. The pretraining could use a vast amount of unlabelled IMU data collected from various wearable devices. The fine-tuning for each specific task could then use

the limited labelled data. To obtain more labelled data, besides crowd sourcing, there are other technique ways like data augmentation, transfer learning, active learning, intelligent labelling, and generative learning.

Intelligent sensor fusion.

Inertial navigation is low-cost, unobtrusive, and self-contained, which allows it to be used in almost any scenario, with any wearable device or smart daily worn objects. However, it suffers from lack of calibration way with real world locations. It is impossible for inertial navigation systems to know their initial absolute position in the real world, not to mention calibrating the estimated position with real position during tracking. Therefore, it is hard to start inertial navigation from a random unknown position, and even hard to keep localisation accuracy in long-term tracking.

One of the possible solutions is the fusion of more IMU sensors attached to different body parts. Research [27] used 12 IMUs separately mounted on the centre of the pelvis, thighs, shanks, feet, center of shoulder, upper arms, and lower arms, calibrated sensors based on the human skeleton model. Research [178] adopted 6 IMUs for human pose estimation based on deep learning. Research [185] used two IMUs mounted on the foot and chest to estimate the accurate position based on the different motion patterns of the two IMUs. The fusion of more IMUs is a promising way to reduce or even eliminate drift. The convenience of using devices may need to be considered if a lot of devices are required to be worn.

Fusing other sensors for tracking is another practical way to calibrate its accumulated errors, such as using Wi-Fi, Bluetooth beacons, RFID tags, GPS, etc. Research [60] used passive RFID tags placed around the building to confirm the reference location to correct the drift in inertial tracking. Research [170] used BLE beacons placed in specific locations indoors to correct the accumulated inertial tracking errors. Research [159] proposed a WiFi-assisted inertial odometry technique that uses WiFi signals as an auxiliary source of information to correct drift errors. Research [111] adopted the BLE signal system on lampposts, together with GNSS and iner-

tial sensors, for real-time positioning in urban environments. These researches all required sensors pre-installed at the known locations in the specific environments to provide reference positions for tracking systems. The reference positions help calibrate the estimated positions with real locations and avoid drift and accumulated errors in inertial tracking systems. But on the other hand, the application scenarios of these systems are limited to only the environments with the required pre-installed infrastructures. It is impossible to use them in an unknown environment or when the infrastructure is faulty or powered off.

If a tracking system with more generalisable application scenarios is needed, other infrastructureless sensors could be adopted as well, such as cameras, Lidar, radar, etc. RGB cameras are basically the essential sensor for tracking in autonomous driving and for other autonomous vehicles or robots. Combining with inertial sensors, visual-inertial localisation is an important task in the computer vision area, reaching high accuracies with the rapid development of deep learning in recent years. Research [190] is one of the examples that uses semantic segmentation technique in computer vision for visual-inertial odometry estimation. Other kinds of cameras also help in different situations, such as thermal cameras, event cameras, night vision cameras, etc, with similar computer vision techniques. Other range sensors like Lidar and radar could generate point clouds to sense 3D information of the environment relative to the current position. Research [37] estimates radar inertial odometry for localisation in visually degraded and GNSS denied environments. Research [177] tightly coupled lidar and inertial sensor data for odometry and mapping. These sensors do not require infrastructures, and thus can be used in a variety of scenarios. However, the aspects that limit their application include running time and energy consumptions. 2D image processing with deep learning would require more calculation resources compared to 1D inertial sensor data, leading to longer calculation time and energy consumption, which need careful consideration for real-time or long-term trackings.

How to fuse them intelligently is the main question in this research. As these infrastructure-based methods do not always work: sometimes they are unavailable;

sometimes their signals are severely affected and not reliable. In addition, for long-term tracking, battery life is valued. It is not practical to always turn on all the energy-consuming sensors. A decision making solution is essential for the balance of accuracy and energy with the questions like: when to turn on each sensor, how long each time lasts, and to what extent should the result be trusted from some sensor at some time. Probabilistic models are the first steps to generate uncertainties of results. The decision optimisation could then be carried out by deep learning.

Human motion analysis for more applications

In addition to position tracking, inertial sensors could be used for human pose estimation, human activity recognition, and other applications benefited from motion analysis. Human motion analysis encompasses a wide range of techniques aimed at understanding, quantifying, and interpreting human movement patterns. With the integration of the knowledge from other areas like biomechanics, neuroscience, rehabilitation, etc., human motion analysis has various applications including the following:

- Understanding complex human movements.
- Predicting the risk of injury.
- Optimising rehabilitation protocols.
- Design of assistive devices.
- Mental health monitoring.
- Emotion recognition.
- Human-computer interaction.

Traditional methods such as optical motion capture systems are often considered to be the gold standard method for motion capture and have been widely used but

are often limited by factors such as cost, complexity, and invasiveness. For example, three-dimensional motion capture systems like Vicon (Vicon Motion Systems Ltd., Oxford, UK) or Optotrak (Northern Digital Inc, Ontario, Canada) have a high accuracy, but require conduction in controlled environments with several fixed cameras pre-installed, calibrated, and correlated. Meanwhile, human subjects are required to place up to 50 markers at anatomically specific locations, which makes it even harder to conduct motion analysis in daily life environment for daily activities. Two-dimensional motion capture system, like Kinect (Microsoft Corporation, WA, USA), is a more affordable alternative, and can be used in almost any relatively uncontrolled environments. However, multiple cameras may still be required for a full motion analysis because of the problem of a restricted field of view of the camera. Parallax error and perspective error will also lead to wrong perception of spatial relationships. The privacy issue of video recording makes it hardly acceptable to conduct motion capture in private places. Besides, with the limited maneuverability margins, these camera-based systems are designed primarily for indoor use. On the contrary, wearable inertial sensors benefited from their portability and can be used in any indoor and outdoor scenarios, even in a private space. Many research investigations have shown strong alignment between IMU and the gold standard 3D optical motion capture systems on the spatiotemporal and kinematic measure, across various domains such as clinical, ergonomics, and sports analysis [13, 126, 184, 65, 153]. To extract the motion information from the raw noisy 1D IMU data, complex sensor fusion and pose estimation methodologies are necessary. Combining with machine learning algorithms, wearable inertial sensors could offer the potential for accurate real-time motion tracking in diverse environments without the need for specialised equipment.

Beyond pose and activity measurements, if IMUs are combined with other wearable monitoring sensors, motion analysis can be used in a variety of applications. Integrating IMU data with other modalities such as EMG, electroencephalography (EEG), and heart rate monitoring can provide deeper insights into human movement patterns and physiological responses. Research [169] developed a wearable system

for American sign language recognition with 4 channel surface EMG on the right forearm and an IMU on the right wrist. EMG and IMU combination has also been used to develop an upper limb exoskeleton for rehabilitation [48], multi-grip prosthesis control [93], omnidirectional wheelchair control [95], etc. For the combination of IMU and EEG, research [104] has used it to create a smart safety helmet to detect the fatigue level of the workers. Research [88] adopted the sensor fusion of the IMU and EEG for drowsiness detection. Research [89] estimated respiration rate from imu with Photoplethysmogram (PPG). With its characteristic of recording any subtle or obvious motions directly, IMU has already shown great potential in various applications of different areas. In the future, IMU motion capture and analysis could have more application areas, such as healthcare in low- and middle-income countries for its availability and popularity, or personal care for older adults living alone.

Appendix A

Characteristic summary and quality assessment of the systematic review

Title	Date	Author	Method	Algorithm	Mean error (m)	Number of subjects	Error (m)	Error	Indoor(1) /Outdoor(0) 2d / 3d	1. Does the study address a clearly focused question?	2. Do the authors discuss how they decided which method to use?	3. Is there sufficient detail of the methods used?	4. Is the method of data collection well described?	5. Are the explanations for the results plausible and coherent?	6. Are the results of the study compared with those from other studies?	7. Did the authors identify any limitations?	8.Was ethical approval sought?
3D ActionSLAM: wearable person tracking in multi-floor environments	2015	Hardegger, Michael; Roggen, Daniel; Troester, Gerhard	a single foot- mounted IMU and a hip-worn smartphone that acts as both, sensor and computing platform.	ZUPT-PDR, action recognition algorithm, Particle filter	-	23	282	7.00%	1	3	Y	Y	Y	Y	N	N	Y	N
3D pedestrian dead reckoning and activity classification using waist-mounted inertial measurement unit	2015	F. Inderst; F. Pascucci; M. Santoni	Waist-mounted IMU	Cascade Pedestrian Dead Reckoning (C-PDR), activity classification, Extended Kalman Filter,	26	1	280	10.00%	1	3	Y	Y	Y	Y	Y	N	Y	N
A comparison of Pedestrian Dead-Reckoning algorithms using a low-cost MEMS IMU	2009	A. R. Jimenez; F. Seco; C. Prieto; J. Guevara	a single foot- mounted IMU	The Weiberg SL Algorithm, ZUPT	5-15	1	100-320	5%	10	2	Y	Y	Y	N	N	N	Y	N
A foot-mounted PDR system based on IMU/EKF+HMM+ZUPT+ZARU+HDR+compass algorithm	2017	W. Zhang; X. Li; D. Wei; X. Ji; H. Yuan	a single foot- mounted IMU	IMU/EKF+HMM+ZUPT+ZARU+HDR+Compass algorithm	0.643	1	310	0.21%	1	3	N	N	Y	Y	N	N	N	N
A Handheld Inertial Pedestrian Navigation System With Accurate Step Modes and Device Poses Recognition	2015	H. Zhang; W. Yuan; Q. Shen; T. Li; H. Chang	A handheld inertial pedestrian navigation system (IPNS). The system consists of a sensor module and a processor module. The sensor module includes a tri-axial accelerometer, a tri-axial gyroscope, a tri-axial magnetometer and a pressure sensor.	Step Modes and Device Poses Recognition, Step Detector With Band-Pass Filter	4.5indoor, 40outdoor	1	2100outdoor	2.00%	10	32	Y	Y	Y	Y	N	Y	Y	N
A Hybrid Heading Estimation Scheme Exploiting Smart-phone Inertial Sensors for PDR based Indoor Navigation	2015	Kim, Nammoon; Zeng, Qingchi; Kim, Youngok	A smartphone with inertial sensors	Hybrid heading estimation	-	1	50	error decrease 17%	1	2	N	N	N	N	N	Y	N	N
A hybrid step model and new azimuth estimation method for pedestrian dead reckoning	2014	Y. Zhu; R. Zhang; W. Xia; Z. Jia; L. Shen	a head-mounted IMU	Hybrid Step Length Model, New Azimuth Estimation Method	0.6178	1	25	-	1	2	Y	Y	Y	N	N	Y	N	N
A localization system using inertial measurement units from wireless commercial hand-held devices	2013	A. Mikov; A. Moschevikin; A. Fedorov; A. Sikora	A hand-held device equipped with a gyroscope and an accelerometer,which can be attached on hand, front trousers pocket, front shirt/jacket pocket or back trousers pocket.	ZUPT	-	10	270	5%	1	2	Y	Y	Y	N	N	Y	N	N
A low-cost shoe-mounted Inertial Navigation System with magnetic disturbance compensation	2013	R. Ashkar; M. Romanovas; V. Goridko; M. Schwaab; M. Traechtler; Y. Manoli	An IMU placed on the shoe, magnetic sensor	UKF, EEKF, EKF, ZUPT, Zero Angular Rate Update (ZARU), magnetic disturbance compensation	0.0194indoor, 3.0882outdoor	1	-	-	10	3	Y	Y	Y	N	N	N	N	N
A method of pedestrian dead reckoning for smartphones using frequency domain analysis on patterns of acceleration and angular velocity	2014	M. Kourogi; T. Kurata	A smartphone	-	4.5	10	270	1.70%	1	2	Y	Y	N	Y	Y	N	Y	N
A Multi-Mode Dead Reckoning System for Pedestrian Tracking Using Smartphones	2016	Q. Tian; Z. Salcic; K. I. Wang; Y. Pan	A smartphone	mode classification,	0.6indoor, 13outdoor	5	96.33indoor, 348outdoor	1.57%indoor, 3.7%outdoor	1	2	Y	Y	Y	Y	Y	Y	N	N
A Novel Body Motion Model based Personal Dead-reckoning System	2011	Xiang, Zhiyu; Qi, Baozhen; Wang, Jiafeng	A compass, a cheap MEMS IMU worn at the waist	Kalman filter	1.11	1	41.96	2.58%	1	2	Y	Y	Y	N	N	N	N	N
A novel heading estimation algorithm for pedestrian using a smartphone without attitude constraints	2016	Donghui Liu; L. Pei; J. Qian; L. Wang; P. Liu; Zhenjiang Dong; Siyuan Xie; Wei Wei	A smartphone	Attitude and Heading Reference System (AHRS)	0.43	1	-	-	1	2	Y	Y	N	N	N	N	N	N
A Novel Pedestrian Dead Reckoning Algorithm for Multi-Mode Recognition Based on Smartphones	2019	Xu, Limin; Xiong, Zhi; Liu, Jianye; Wang, Zhengchun; Ding, Yiming	A smartphone	zero angular velocity, lateral velocity limitation (LV)	1.53	1	198.4	0.90%	1	2	Y	Y	Y	N	Y	Y	N	N
A Novel Step Detection and Step Length Estimation Algorithm for Hand-held Smartphones	2018	N. Strozzi; F. Parisi; G. Ferrari	A smartphone	Weinberg step length estimation formula	-	1	188	3.00%	1	2	Y	Y	Y	N	N	N	N	N
A Novel Technique for Gait Analysis Using Two Waist Mounted Gyroscopes	2019	A. Nasr; T. Nadeem	2 IMUs mounted on the waist	peak detection, Weinberg method,	-	3	32	5.15%	1	3	Y	Y	Y	N	Y	Y	N	N
A pedestrian dead reckoning system using a foot kinematic constraint and shoe modeling for various motions	2018	Ju, Hojin; Park, Chan Gook	a foot-Mounted IMU and a pressure sensor	-	0.11	1	60	0.18%	1	2	Y	Y	Y	Y	Y	Y	N	N
A Pedestrian Dead-Reckoning System for Walking and Marking Time Mixed Movement Using an SHSs Scheme and a Foot-Mounted IMU	2019	Y. Wu; H. Zhu; Q. Du; S. Tang	a Foot-Mounted IMU	extended Kalman filter (EKF), zero angular rate measurement (ZARM), Heuristic Heading Reduction (HDR), Flat-ground Hypothesis (FGH) and Cardinal Heading Aided Inertial Navigation (CHAIN), back propagation (BP) algorithm	4.38	7	320.8	0.50%	1	2	Y	Y	Y	Y	Y	N	Y	N
A pedestrian dead-reckoning system that considers the heel-strike and toe-off phases when using a foot-mounted IMU	2016	Ju, Hojin; Lee, Min Su; Park, So Young; Song, Jin Woo; Park, Chan Gook	a foot-Mounted IMU	EKF, ZUPT	indoor1.24, outdoor7.63	1	indoor340, outdoor700	indoor0.36%, outdoor1.09%	10	3	Y	N	Y	Y	Y	Y	N	N
A robust dead-reckoning pedestrian tracking system with low cost sensors	2011	Y. Jin; Hong-Song Toh; W. Soh; Wai-Choong Wong	Three smartphones, two are placed in the pedestrian's two trousers pockets, one is held in hand.	low-pass filters (LPF)	-	1	-	reduction in average tracking error up to 73.7%	1	2	Y	Y	Y	N	N	Y	N	N
A robust humanoid robot navigation algorithm with ZUPT	2012	Y. Li; X. Luo; X. T. Ren; J. J. Wang	A shoe-mounted IMU	ZUPT, EKF	-	1	240	0.33%	1	3	Y	Y	Y	N	N	N	Y	N
A Robust Indoor/Outdoor Navigation Filter Fusing Data from Vision and Magneto-Inertial Measurement Unit	2017	Caruso, David; Eudes, Alexandre; Sanfourche, Martial; Vissiere, David; Le Besnerais, Guy	monocular camera, MIMU	MSCKF algorithm, KF; error-state kalman	1.64	1	530	0.31%	1	3	Y	Y	Y	N	Y	Y	N	N
A Robust Pedestrian Dead Reckoning System Using Low-Cost Magnetic and Inertial Sensors	2019	L. Shi; Y. Zhao; G. Liu; S. Chen; Y. Wang; Y. Shi	An inertial and magnetic measurement unit (IMMU) mounted on foot	zero velocity update (ZVU), Gait Phase Detection Algorithm, KF	0.74indoor, 2.33outdoor	1	103.86indoor, 402.7outdoor	0.73%indoor, 0.58%outdoor	10	2	Y	Y	Y	Y	Y	Y	Y	N
A robust pedestrian dead-reckoning positioning based on pedestrian behavior and sensor validity	2012	S. Asano; Y. Wakuda; N. Koshizuka; K. Sakamura	A three-axis accelerometer, a three-axis gyroscope, a three-axis magnetometer and a barometer, mounted on waist	particle filter,	0.36	1	600	-	10	2	Y	Y	N	N	N	N	Y	N
A sensor-to-segment calibration method for motion capture system based on low cost MIMU	2018	Choe, Namchol; Zhao, Hongyu; Qiu, Sen; So, Yongguk	twelve MIMUs separately mounted on the center of pelvis, thighs, shanks, feet, center of shoulder, upper arms, and lower arms	a novel calibration method	-	1	100	1.00%	1	3	Y	Y	N	Y	Y	N	Y	N

A Smartphone-Based Pedestrian Dead Reckoning System With Multiple Virtual Tracking for Indoor Navigation	2018	H. Ju; S. Y. Park; C. G. Park	A smartphone	low-pass filters (LPF), ARS, EKF,	3.85	3	710	0.77%	1	2	Y	Y	Y	Y	Y	Y	N	N
A Wearable Inertial Pedestrian Navigation System With Quaternion-Based Extended Kalman Filter for Pedestrian Localization	2017	Y. Hsu; J. Wang; C. Chang	A shoe-mounted IMU	zero-velocity compensation (ZVC), Lowpass Filter, EKF, probabilistic neural network (PNN),	8.33indoor, 5.28outdoor	1	239.9indoor, 1020outdoor	3.47%indoor, 0.52%outdoor	10	3	Y	Y	Y	Y	Y	Y	N	N
Accurate 3D-Tracking System for Wireless Indoor Personal Positioning	2018	M. Asadi; S. Sadeghi; A. Kargar; A. Mahani	A shoe-mounted IMU	zero velocity update (ZVU), Enhanced Heuristic Drift Elimination (EHDE)	0.22	10	40	6.00%	1	3	Y	Y	N	Y	Y	Y	N	N
ActionSLAM on a smartphone: At-home tracking with a fully wearable system	2013	M. Hardegger; S. Mazilu; D. Caraci; F. Hess; D. Roggen; G. Tr??ster	a smartphone placed in the hip pocket	Action Recognition, partical filter	0.39	1	130	0.30%	1	2	Y	Y	N	N	N	Y	Y	N
ActionSLAM: Using location-related actions as landmarks in pedestrian SLAM	2012	M. Hardegger; D. Roggen; S. Mazilu; G. Tr??ster	A foot-mounted IMU	a specific instantiation of the FastSLAM framework optimized to operate with action landmarks, ZUPT, EKF,	1.2	1	960	0.13%	1	2	Y	N	Y	N	N	Y	N	N
Activity and location recognition using wearable sensors	2002	Seon-Woo Lee; K. Mase	The leg module, contains the biaxial accelerometer and the gyroscope located in the user's right or left trouser pocket. the waist module, contains a digital compass and a microcontroller attached to the middle of the user's waist.	Unit motion recognition	-	8	90	8.20%	1	3	Y	N	Y	N	N	N	Y	N
Activity classification and dead reckoning for pedestrian navigation with wearable sensors	2009	Sun, Zuolei; Mao, Xuchu; Tian, Weifeng; Zhang, Xiangfen	A MEMS, a tri-axial accelerometer and an electronic magnetic compass, mounted on the belt.	fast Fourier transforms (FFTs), Probabilistic neural network (PNN), wavelet transform (WT)	14.7	1	434.21	3.39%	0	2	Y	Y	Y	Y	Y	Y	Y	N
Adaptive Zero Velocity Update Based on Velocity Classification for Pedestrian Tracking	2017	R. Zhang; H. Yang; F. H??flinger; L. M. Reindl	Two IMUs mounted on foot and chest	ZUPT, Adaptive Neuro-Fuzzy Inference System (ANFIS)	1	10	40	2.50%	1	3	Y	Y	N	Y	N	N	Y	N
An advanced method for pedestrian dead reckoning using BLSTM-RNNs	2015	M. Edel; E. K??ppe	a Foot-Mounted IMU	BLSTM-RNNs, linear regression	-	10	-	0.80%	10	3	N	N	N	Y	Y	Y	N	N
An alternative approach to vision techniques: Pedestrian navigation system based on digital magnetic compass and gyroscope integration	2002	Ladetto, Q; Merminod, B	A compass and a gyroscope worn vertical at the belt level	-	10	1	1905	0.53%	0	3	N	N	N	N	N	N	N	N
An autonomic indoor positioning application based on smartphone	2014	Y. Sun; Y. Zhao; J. Schiller	A smartphone	Zero Velocity Compensation (ZVC), KF,	0.298	10	40	0.75%	1	2	Y	Y	N	Y	N	N	Y	N
An efficient method for evaluating the performance of integrated multiple pedestrian navigation systems	2015	H. Lan; C. Yu; Y. Li; Y. Zhuang; N. El-Sheimy	Two smartphones mounted on the user's upper body	KF	1.289	1	193	0.67%	1	2	N	Y	Y	Y	Y	N	N	N
An enhanced pedestrian dead reckoning approach for pedestrian tracking using smartphones	2015	Qinglin Tian; Z. Salcic; K. I. Wang; Yun Pan	A smartphone	classification algorithm, False Peak Rejection (FPR)	0.3	10	96.33	0.30%	1	2	Y	Y	N	N	N	N	N	N
An Experimental Heuristic Approach to Multi-Pose Pedestrian Dead Reckoning Without Using Magnetometers for Indoor Localization	2019	J. Lee; S. Huang	A smartphone	classification algorithm,	3.51	4	89	2.01%	1	2	Y	Y	Y	Y	Y	Y	N	N
An Improved Pedestrian Navigation System Using IMU and Magnetometer	2017	Li, Zhenwei; Song, Chunlei; Cai, Jingyi; Hua, Rui; Yu, Pei	A foot-mounted IMU, magnetometer	KF, Gait detection algorithm,	4.27	1	400	1.00%	0	2	Y	Y	Y	N	N	N	N	N
An indoor self-localization algorithm using the calibration of the online magnetic fingerprints and indoor landmarks	2016	Q. Wang; H. Luo; F. Zhao; W. Shao	A smartphone, a foot-mounted IMU	zero velocity update, Indoor landmark detection, magnetic trajectory matching	2.55	1	-	80 percentile localization accuracy of 1.4m	1	2	Y	Y	N	N	N	Y	N	N
An inverse square root filter for robust indoor/outdoor magneto-visual-inertial odometry	2017	D. Caruso; A. Eudes; M. Sanfourche; D. Vissiere; G. le Besnerais	camera and MIMU	MSCKF VINS algorithm, Inverse Square root Filter	1.89	1	-	0.70%	1	3	Y	N	Y	N	N	Y	N	N
Application of multisensor fusion to develop a personal location and 3D mapping system	2018	Hsu, Ya-Wen; Huang, Shiang-Shuang; Perng, Jau-Woei	an IMU, a laser range finder, a Microsoft Kinect RGB-D sensor	RGB-D SLAM, IMU/laser SLAM	1.15	1	30	3.83%	1	3	Y	Y	Y	Y	N	Y	Y	N
Comparison of EMG-based and Accelerometer-based Speed Estimation Methods in Pedestrian Dead Reckoning	2011	Chen, Wei; Chen, Ruizhi; Chen, Xiang; Zhang, Xu; Chen, Yuwei; Wang, Jianyu; Fu, Zhongqian	EMG sensors attached to both legs, an accelerometer and a compass attached on the abdominal area	-	-	1	1592.8	0.69%	0	2	Y	Y	Y	Y	Y	Y	Y	N
Comparison of Pedestrian Tracking Methods Based on Foot- and Waist-Mounted Inertial Sensors and Handheld Smartphones	2019	Yu, Ning; Li, Yunfei; Ma, Xiaofeng; Wu, Yinfeng; Feng, Renjian	IMU mounted on foot, waist, hand respectively	Madgwick, Mahony, SVO-LKF, FCF, EKF, EHDE, VLKF	0.8	1	340	0.30%	1	2	Y	Y	N	Y	Y	Y	Y	N
Context-Aided Inertial Navigation via Belief Condensation	2016	J. Prieto; S. Mazuelas; M. Z. Win	a Foot-Mounted IMU	belief condensation (BC), Hidden Markov Model	0.55	1	435	0.12%	1	2	Y	Y	Y	Y	Y	Y	N	N
Continuous Motion Recognition for Natural Pedestrian Dead Reckoning Using Smartphone Sensors	2014	Qian, Jiuchao; Pei, Ling; Ying, Rendong; Chen, Xin; Zou, Danping; Liu, Peilin; Yu, Wenxian	A smartphone	continuous motion recognition algorithm, Conditional Random Fields (CRFs) algorithm	0.27	1	-	-	1	2	Y	Y	N	N	N	N	N	N
Dead Reckoning in Structured Environments for Human Indoor Navigation	2017	G. Giorgi; G. Frigo; C. Narduzzi	A smartphone	SD, HDE, KF	3.8	2	68.23	4.00%	1	2	Y	Y	N	Y	N	N	Y	N
Dead reckoning navigation with Constant Velocity Update (CUPT)	2012	Y. Li; J. J. Wang; S. Xiao; X. Luo	a Foot-Mounted IMU	Constant Velocity Update (CUPT)	-	1	315.3	0.37%	1	3	Y	Y	Y	N	N	N	N	N
Design and implementation of an inertial navigation system for pedestrians based on a low-cost MEMS IMU	2013	F. Montorsi; F. Pancaldi; G. M. Vitetta	a Foot-Mounted IMU	EKF, ZUPT	10	10	300	3.33%	1	2	Y	N	Y	N	N	N	Y	N
Design and Implementation of MARG Sensors Based Positioning Method Using a Mobile Phone	2015	Tian, Zengshan; Qian, Guang; Zhou, Mu	A smartphone	EKF	-	1	-	0.20%	1	2	Y	Y	N	N	N	N	N	N
Design of an infrastructureless in-door localization device using an IMU sensor	2015	T. Do; R. Liu; C. Yuen; U. Tan	a Foot-Mounted IMU	quaternion-based indirect Kalman filter (iKF)	0.84	4	40	0.44%	1	2	Y	Y	Y	Y	Y	N	N	N
DEVELOPMENT OF A PEDESTRIAN INDOOR NAVIGATION SYSTEM BASED ON MULTI-SENSOR FUSION AND FUZZY LOGIC ESTIMATION ALGORITHMS	2015	Lai, Y. C.; Chang, C. C.; Tsai, C. M.; Lin, S. Y.; Huang, S. C.	handheld/waist mounted IMU	multi-sensor fusion and fuzzy logic estimation algorithms	0.33	1	45	0.73%	1	2	Y	Y	N	N	N	N	N	N
Displacement estimation in micro-sensor motion capture	2010	Xiaoli Meng; Shuyan Sun; Lianying Ji; Jiankang Wu; Wai-Choong Wong	seven SMUs, which are mounted to the human pelvis, thighs, shanks and feet respectively	complementary Kalman filter (CKF)	0.72	1	15	4.80%	1	2	N	Y	Y	Y	N	N	N	N
Enhanced Heuristic Drift Elimination with Adaptive Zero-Velocity Detection and Heading Correction Algorithms for Pedestrian Navigation.	2020	Zhu R, Wang Y, Yu B, Gan X, Jia H, Wang B.	a foot mounted IMU	HDE, ZUPT	-	2	-	1.06%	0	2	Y	Y	Y	N	Y	Y	N	N
Enhanced Pedestrian Navigation Based on Course Angle Error Estimation Using Cascaded Kalman Filters.	2018	Song JW, Park CG.	a Foot-Mounted IMU	two cascaded Kalman filters (TCKF), inertial navigation system-extended Kalman filter-ZUPT (INS-EKF-ZUPT)	0.794outdoor, 0.978indoor	1	400outdoor	0.198%outdoor	10	2	Y	Y	Y	Y	Y	Y	N	N

Enhancing Improved Heuristic Drift Elimination for Wrist-Worn PDR Systems in Buildings	2016	L. E. Diez; A. Bahillo; S. Bataineh; A. D. Masegosa; A. Perallos	a wrist-worn IMU	modified improved heuristic drift elimination (iHDE)	0.57	1	256	0.22%	1	2	Y	Y	N	N	N	Y	N	N
Error Modelling for Multi-Sensor Measurements in Infrastructure-Free Indoor Navigation	2018	Ruotsalainen, Laura; Kirkko-Jaakkola, Martti; Rantanen, Jesperi; Makela, Maija	a monocular camera, a foot-mounted Inertial Measurement Unit (IMU), sonar, and a bar	particle filter	2.54	1	170	1.49%	1	2	Y	Y	N	Y	Y	N	N	N
Evaluation of a new method of heading estimation for pedestrian dead reckoning using shoe mounted sensors	2005	Stirling, R; Fyfe, K; Lachapelle, G	a Foot-Mounted IMU	-	25	1	900	2.77%	0	2	Y	N	N	N	N	Y	Y	N
Foot-mounted INS and selected geomagnetic information constraint for individual localization	2017	W. Duo; P. Xianfei; H. Xiaoping; H. Xiaofeng	foot-mounted MEMS inertial/ magnetic measurement unit	ZUPT, recursive least square (RLS) with forgetting factor algorithm	3	1	185.8	1.60%	0	2	Y	Y	N	Y	Y	Y	N	N
Foot-Mounted Pedestrian Navigation Algorithm Based on BOR/MINS Integrated Framework	2019	Z. Den; P. Wang; T. Liu; Y. Cao; B. Wang	a Foot-Mounted MIMU	heuristic drift reduction (HDR), complementary filter	2.01	5	408	0.49%	0	2	Y	Y	Y	Y	Y	Y	N	N
Heading Drift Reduction for Foot-Mounted Inertial Navigation System via Multi-Sensor Fusion and Dual-Gait Analysis	2019	H. Zhao; Z. Wang; S. Qiu; Y. Shen; L. Zhang; K. Tang; G. Fortino	two IMUs mounted on each foot	zero velocity updates (ZUPT)	3	4	346	0.86%	0	2	Y	Y	Y	Y	Y	N	N	N
Heading Estimation for Pedestrian Dead Reckoning Based on Robust Adaptive Kalman Filtering.	2018	Wu D, Xia L, Geng J.	A smartphone	robust adaptive Kalman filtering (RAKF)	1.35	5	150.4	0.89%	1	2	Y	Y	Y	Y	Y	Y	N	N
Heading Estimation with Real-time Compensation Based on Kalman Filter Algorithm for an Indoor Positioning System	2016	Li, Xin; Wang, Jian; Liu, Chunyan	MEMS, IMU of smartphone	KF	2.31	1	-	-	1	2	Y	Y	N	N	N	Y	N	N
Hierarchical calibration architecture based on Inertial/magnetic sensors for indoor positioning	2018	G. Liu; L. Shi; J. Xun; S. Chen; H. Liu; Y. Shi	IMU mounted on the ankle	EKF, machine learning	0.5	1	220	0.23%	1	3	Y	Y	N	N	N	Y	N	N
Human velocity tracking and localization using 3 IMU sensors	2013	Q. Yuan; I. Chen; A. Caus	Three IMU mounted on the right shank, right thigh and the pelvis, Four force sensing resistors (FSR) in the insole shoe pad	EKF	0.1	1	15	0.67%	1	3	Y	Y	Y	N	N	N	Y	N
Hybrid Indoor Localization Using IMU Sensors and Smartphone Camera	2019	Poulose, Alwin; Han, Dong Seog	IMU and camera on a smartphone	UcoSLAM, KF	0.069	1	18.2	0.38%	1	2	Y	N	Y	Y	Y	Y	N	N
Hybrid Navigation Method of INS/PDR Based on Action Recognition	2018	J. Lu; K. Chen; B. Li; M. Dai	a Foot-Mounted MIMU	Motion Classification, ZUPT	-	4	52.2	1.13%	1	2	Y	Y	Y	Y	Y	Y	Y	N
Hybrid orientation filter aided indoor tracking for pedestrians using a smartphone	2017	Z. Yang; Y. Pan; L. Zhang	A smartphone (camera)	Hybrid Orientation Filter	0.83	5	207	0.41%	1	2	Y	Y	N	Y	Y	Y	N	N
iDFusion: Globally Consistent Dense 3D Reconstruction from RGB-D and Inertial Measurements	2019	Zhong, Dawei; Han, Lei; Fang, Lu	IMU and camera	TSDF fusion	0.01	1	-	-	1	3	Y	N	Y	N	N	Y	Y	N
Implementation and performance analysis of smartphone-based 3D PDR system with hybrid motion and heading classifier	2014	B. Shin; S. Lee; C. Kim; J. Kim; T. Lee; C. Kee; S. Heo; H. Rhee	A smartphone	motion recognition, ZUPT, KF, SVM	2.25	3	133.2	5.00%	1	3	Y	Y	N	Y	Y	N	N	N
Improved Pedestrian Dead Reckoning Based on a Robust Adaptive Kalman Filter for Indoor Inertial Location System	2019	Fan, Qigao; Zhang, Hai; Pan, Peng; Zhuang, Xiangpeng; Jia, Jie; Zhang, Pengsong; Zhao, Zhengqing; Zhu, Gaowen; Tang, Yuanyuan	a Foot-Mounted MIMU	robust adaptive Kalman filter (RAKF) algorithm, Complementary Filter	0.27	1	231.8	2.50%	1	2	Y	Y	Y	N	Y	Y	N	N
Improving pedestrian navigation system performance through the use of non-orthogonal redundant inertial measurement units	2017	P. Marinushkin; A. Levitskiy; F. Zograf	a Foot-Mounted IMU	extended Kalman filter	-	1	170	2.50%	1	2	Y	Y	N	N	N	Y	N	N
In situ heading drift correction for human position tracking using foot-mounted inertial/magnetic sensors	2012	E. Bachmann; J. Calusdian; E. Hodgson; X. Yun	2 Foot-Mounted Inertial/Magnetic Sensors	zero position update	1.07	4	400	0.20%	0	2	Y	Y	N	Y	Y	N	Y	N
Indoor infrastructure-less solution based on sensor augmented smartphone for pedestrian localisation	2012	G. Trehard; S. Lamy-Perbal; M. Boukallel	a smartphone with anemometer gyroscopic sensors	Pitot tube theory	3	1	50	6.00%	1	2	Y	Y	Y	N	N	N	Y	N
Indoor Multi-Floor 3D Target Tracking Based on the Multi-Sensor Fusion	2020	J. Luo; C. Zhang; C. Wang	smartphone	KF, FCD	0.499	1	-	-	1	3	Y	Y	Y	N	Y	Y	N	N
Indoor pedestrian localisation solution based on anemometry sensor integration with a smartphone	2012	G. Trehard; M. Boukallel; S. Lamy-Perbal	a smartphone with anemometer and gyroscopic sensors	Pitot tube theory	3	1	50	6.00%	1	2	Y	Y	Y	N	N	N	Y	N
Indoor Pedestrian Localization With a Smartphone: A Comparison of Inertial and Vision-Based Methods	2016	W. Elloumi; A. Latoui; R. Canals; A. Chetouani; S. Treuillet	A smartphone	Extended Kalman Filter (EKF)	0.519	1	59	0.88%	1	2	Y	Y	N	Y	Y	Y	Y	N
Indoor pedestrian navigation using an INS/EKF framework for yaw drift reduction and a foot-mounted IMU	2010	A. R. Jim??nez; F. Seco; J. C. Prieto; J. Guevara	a Foot-Mounted IMU	INS- EKF-ZUPT (IEZ), ZARU, HDR	0.3 - 1.5	1	125	1.00%	1	2	Y	Y	Y	N	N	Y	N	N
Indoor positioning based on foot-mounted IMU	2015	Guo, H.; Uradzinski, M.; Yin, H.; Yu, M.	a Foot-Mounted IMU	ZARU, ZUPT	0.88	1	90.48	1.00%	1	2	Y	Y	Y	Y	Y	Y	Y	N
Indoor positioning system using walking pattern classification	2014	F. De Cillis; F. De Simio; L. Faramondi; F. Inderst; F. Pascucci; R. Setola	Waist-mounted IMU	classification, Extended Kalman Filter	2	1	200	1.00%	1	3	Y	Y	N	N	N	N	Y	N
Indoor Positioning System: A New Approach Based on LSTM and Two Stage Activity Classification	2019	Hussain, Ghulam; Jabbar, Muhammad Shahid; Cho, Jun-Dong; Bae, Sangmin	A smartphone	LSTM and Two Stage Activity Classification	0.782	19	-	-	1	2	Y	Y	Y	N	Y	Y	N	N
IndoorGuide ?€? A multi sensor pedestrian navigation system for precise and robust indoor localization	2016	J. Ruppelt; P. Merz; G. F. Trommer	A IMU and a camera or a rotating laser scanner mounted on the upper body, a IMU mounted on the foot	stochastic cloning error state space Kalman filter (SCKF), FSM-based ZUPT, Multi-frame Feature Integration (MfFI), RANdom SAmple Consensus (RANSAC) strategy	0.1	1	-	-	1	2	Y	Y	N	N	N	Y	N	N
Inertial Sensor Based Indoor Localization and Monitoring System for Emergency Responders	2013	R. Zhang; F. Hoflinger; L. Reindl	six IMUs mounted on waist, thighs, shanks, and one foot	modified Kalman filter, zero velocity update (ZUPT)	2.35	1	-	-	1	3	Y	Y	Y	Y	Y	N	Y	N
Inertial/magnetic sensors based pedestrian dead reckoning by means of multi-sensor fusion	2018	Qiu, Sen; Wang, Zhelong; Zhao, Hongyu; Qin, Kairong; Li, Zhenglin; Hu, Huosheng	a Foot-Mounted IMU	EKF, zero velocity update(ZVU)	0.49indoor, 2.59outdoor	4	50indoor, 423outdoor	0.41%indoor, 0.65%outdoor	10	23	Y	Y	Y	Y	Y	Y	Y	N
Inertial-vision sensor fusion for pedestrian localization	2011	D. Chdid; R. Oueis; H. Khoury; D. Asmar; I. Elhajj	An IMU mounted on the foot and camera rigs attached to the waist	EKF incorporates three correction techniques: ZUPT, ZARU and HDR	0.18645	1	40	0.47%	1	2	Y	Y	N	Y	Y	Y	N	N
In-plane dead reckoning with knee and waist attached gyroscopes	2011	Liu, Zexi; Aduba, Chukwuemeka; Won, Chang-Hee	two gyroscopes attached to the waist and knee	Zero Angular Displacement Up- date (ZADU) algorithm	0.1935	1	62.32	0.31%	1	2	Y	Y	N	N	Y	Y	N	N

Kinematic Model-Based Pedestrian Dead Reckoning for Heading Correction and Lower Body Motion Tracking	2015	Lee, Min Su; Ju, Hojin; Song, Jin Woo; Park, Chan Gook	7 IMUs mounted on waist, each upper leg, each lower leg and each foot	Extended Kalman Filter (EKF), ZUPT	0.2085indoor, 5.35outdoor	5	280indoor, 400outdoor	0.07%indoor, 1.33%outdoor	10	2	Y	Y	Y	Y	Y	Y	Y	N
Localization of Walking or Running User with Wearable 3D Position Sensor	2007	Kazuki Yamanaka, Masayuki Kanbara and Naokazu Yokoya	an electromagnetic sensor , an inertial sensor at user's hip, push button switches attached to user's both heels and toes, electromagnetic sensor are fixed on cnemis of both users' legs	-	1.054	1	30	1.40%	1	3	Y	Y	N	Y	N	Y	N	N
Loose Coupling of Wearable-Based INSs with Automatic Heading Evaluation	2017	Ahmed, Dina Bousdar; Diaz, Estefania Munoz	2 IMUs located on the upper-front part of the foot and on the upper thigh of the same leg	ZUPTs, exteded Kalman filter (EKF) is	-	30	-	10cm/s	0	2	Y	Y	N	Y	N	N	N	N
Low cost infrastructure free form of indoor positioning	2014	S. K. Gupta; S. Box; R. E. Wilson	A smartphone	Fast Fourier Transform (FFT)	1.7	7	32.7	2.70%	1	3	Y	Y	N	Y	Y	Y	Y	N
Low-cost Bluetooth foot-mounted IMU for pedestrian tracking in industrial environments	2015	A. Bahillo; I. Angulo; E. Onieva; A. Perallos; P. Fern??ndez	a Foot-Mounted IMU	ZUPT, KF, heuristic drif elimination (HDE)	2	1	800	0.25%	1	3	Y	Y	N	N	Y	N	N	N
LSTM-Based Zero-Velocity Detection for Robust Inertial Navigation	2018	B. Wagstaff; J. Kelly	a Foot-Mounted IMU	EKF, LSTM	1.083	5	220	0.49%	1	3	Y	Y	N	Y	Y	Y	N	N
Magnetometer-Free Sensor Fusion Applied to Pedestrian Tracking: A Feasibility Study	2019	S. Cardarelli; P. di Florio; A. Mengarelli; A. Tigrini; S. Fioretti; F. Verdini	a foot mounted IMU	KF, ZUPT	0.96	1	43.16	2.22%	1	2	Y	N	Y	Y	Y	N	N	Y
Motion Recognition-Based 3D Pedestrian Navigation System Using Smartphone	2016	Shin, Beomju; Kim, Chulki; Kim, Jaehun; Lee, Seok; Kee, Changdon; Kim, Hyoung Seok; Lee, Taikjin	A smartphone	artificial neural network, KF	-	3	144	6.00%	1	3	Y	Y	Y	N	N	N	N	N
Multimode Pedestrian Dead Reckoning Gait Detection Algorithm Based on Identification of Pedestrian Phone Carrying Position	2019	Guo, Ying; Liu, Qinghua; Ji, Xianlei; Wang, Shengli; Feng, Mingyang; Sun, Yuxi	smartphone	PCA, random forest algorithm	0.261	1	45	0.58%	1	2	Y	Y	Y	Y	Y	Y	N	N
Novel Velocity Update Applied for IMU-based Wearable Device to Estimate the Vertical Distance	2019	T. N. Do; U. Tan	an IMU mounted on waist	-	-	4	-	0.77%	1	3	Y	Y	N	Y	N	N	N	N
On single sensor-based inertial navigation	2016	N. Strozzi; F. Parisi; G. Ferrari	a single wearable Magnetic, Angular Rate, and Gravity (MARG) sensor on foot/chest	Enhanced Pedestrian Dead Reckoning (EPDR) and De-Drifted Propagation (DDP)	1.73indoor, 6.14outdoor	16	90indoor, 400outdoor	2.01%indoor, 1.54%outdoor	10	2	Y	Y	N	N	N	N	N	N
Optimal Heading Estimation Based Multidimensional Particle Filter for Pedestrian Indoor Positioning	2018	L. Pei; D. Liu; D. Zou; R. Lee Fook Choy; Y. Chen; Z. He	A smartphone	multidimensional particle filter (MPF) algorithm	2.5	5	300	0.83%	1	2	Y	Y	Y	N	Y	Y	Y	N
Pedestrian Dead Reckoning Based on Frequency Self-Synchronization and Body Kinematics	2017	M. Basso; M. Galanti; G. Innocenti; D. Miceli	IMU embedded in a smartphone and a more expensive stand-alone IMU freely positioned not too far from the waist	Phase Locked Loop (PLL), PCA	8	2	3262	0.30%	0	2	Y	Y	N	Y	N	N	N	N
Pedestrian Dead Reckoning in Handheld Terminal with Inertial Measurement Unit	2014	Wang, Keji; Deng, Zhongliang; Luo, Shengmei; Yu, Yanpei; Ruan, Fengli	A smartphone	-	3	1	-	-	1	2	Y	Y	N	N	N	N	N	N
Pedestrian dead reckoning on smartphones with varying walking speed	2016	R. Zhou	A smartphone	Simple Moving Average (SMA) algorithm, KF	2.7	1	-	-	1	2	Y	Y	N	N	N	N	Y	N
Pedestrian Dead Reckoning System Using Dual IMU to Consider Heel Strike Impact	2018	H. Ju; J. H. Lee; C. G. Park	shoe and calf mounted IMUs	ZUPT	0.3	3	138	0.22%	1	2	Y	Y	N	Y	Y	Y	N	N
Pedestrian dead reckoning with attitude estimation using a fuzzy logic tuned adaptive kalman filter	2013	M. N. Ibarra-Bonilla; P. Jorge Escamilla-Ambrosio; J. Manuel Ramirez-Cortes; C. Vianchada	a laptop	fuzzy logic tuned adaptive Kalman filter	-	1	-	6.40%	10	2	Y	N	N	N	N	N	N	N
Pedestrian dead reckoning with waist-worn inertial sensors	2012	J. C. Alvarez; A. M. L??pez; R. C. Gonz??lez; D. ??lvarez	Waist-mounted IMU	-	-	1	178.58	7.00%	1	2	Y	Y	N	N	N	N	Y	N
Pedestrian Dead Reckoning-Assisted Visual Inertial Odometry Integrity Monitoring.	2019	Wang Y, Peng A, Lin Z, Zheng L, Zheng H.	a camera and an IMU	Multi-State constrained Kalman Filter (MSCKF)	2.5838	1	1400	0.18%	1	3	Y	Y	Y	Y	Y	Y	N	N
Pedestrian dead-reckoning algorithms for dual foot-mounted inertial sensors	2019	I. A. Chistiakov; A. A. Nikulin; I. B. Gartseev	two foot-mounted inertial measurement units (IMU)	ZUPT	0.241	6	70	0.34%	1	2	Y	N	Y	N	N	Y	N	N
Pedestrian inertial navigation with gait phase detection assisted zero velocity updating	2009	Young Soo Suh; S. Park	An IMU and 4 force sensor mounted on foot	hidden Markov model (HMM) filter, Complementary Kalman filter, Zero velocity updating	0.2422	1	6	4.03%	1	3	Y	N	Y	N	N	N	N	N
Pedestrian Navigation with Foot-Mounted Inertial Sensors in Wearable Body Area Networks	2014	Zhou, Xuan-cheng; Chen, Jian-xin; Dong, Yi; Lu, Xi-ruo; Cui, Jing-wu; Zheng, Bao-yu	a Foot-Mounted IMU	EKF, ZUPT	1.5	1	206	0.73%	0	2	Y	Y	Y	Y	Y	Y	N	N
Pedestrian Tracking by Acoustic Doppler Effects	2019	T. Chiang; K. Ou; J. Qiu; Y. Tseng	A microphone mounted on the left foot and two buzzers mounted on the right foot	short-time Fourier transform (STFT)	0.3646	1	24.4	1.49%	1	2	Y	Y	N	N	Y	Y	Y	N
Performance Enhancement of Pedestrian Navigation Systems Based on Low-Cost Foot-Mounted MEMS-IMU/Ultrasonic Sensor	2019	Xia, Ming; Xiu, Chundi; Yang, Dongkai; Wang, Li	An IMU and an ultrasonic sensor	a straight motion heading update (SMHU), fuzzy adaptive EKF(FAEKF)	12.2	1	800	1.54%	0	2	Y	Y	Y	Y	Y	Y	Y	N
Performance testing of PDR using common sensors on a smartphone	2013	A. R. Pratama; Widyawan; R. Hidayat	A smartphone	Scarlet experimental method	3.58	30	82.67	4.40%	1	2	Y	Y	N	Y	Y	N	N	N
Personal dead reckoning using IMU device at upper torso for walking and running	2016	T. N. Do; R. Liu; C. Yuen; U. Tan	An IMU attached to pedestrian's lower back	quaternion-based Kalman Filter	3.22	1	68.28	4.72%	1	2	Y	N	N	Y	Y	N	N	N
Personal Dead Reckoning Using IMU Mounted on Upper Torso and Inverted Pendulum Model	2016	T. Do; R. Liu; C. Yuen; M. Zhang; U. Tan	IMU mounted on the waist/back/chest	Pythagorean theorem	5.7	4	400	1.42%	0	2	Y	Y	Y	Y	Y	N	Y	N
Personal position measurement using dead reckoning	2003	Randell, C; Djallis, C; Muller, H	compass sensor taped to shoulder; accelerometer sensors attached to pedestrian's shoes; Bitsy motherboard carried in pocket; and accelerometer sensor, extracted from sole holster to show size	-	52	1	4029	1.29%	0	2	Y	Y	N	Y	Y	N	N	N
Position Tracking During Human Walking Using an Integrated Wearable Sensing System	2017	Zizzo, Giulio; Ren, Lei	five ultrasound receivers mounted on the left foot along with the IMUs, two ultrasound transmitters mounted on the right foot	EKF, ZUPT, HDR	0.423	1	55	0.77%	1	2	Y	Y	N	Y	Y	Y	N	N
Real-Time Human Foot Motion Localization Algorithm With Dynamic Speed	2016	L. Van Nguyen; H. M. La	a Foot-Mounted IMU	Extended Kalman filter (EKF), zero velocity update (ZVU)	outdoor3.59, indoor0.45	1	outdoor645, indoor120	outdoor0.55%, indoor0.375%	10	2	Y	Y	Y	Y	Y	N	N	N
Real-Time Infrastructureless Indoor Tracking for Pedestrian Using a Smartphone	2019	Z. Yang; Y. Pan; Q. Tian; R. Huan	A smartphone	Complementary Filter	0.83	6	207	0.40%	1	2	Y	Y	Y	Y	Y	Y	N	N
Real-time Pedestrian Tracking in Indoor Environments	2014	A. Vidal; J. J. Marron; M. A. Labrador	A smartphone	-	5	1	111	10.07%	1	2	Y	Y	N	N	N	N	N	N
Research for pedestrian navigation positioning method based on MEMS sensors	2015	X. Yunqiang; Z. Yanshun; W. Z. Qing; L. Ming	an IMU worn in the back of waist	-	8.7outdoor	1	1560.9outdoor	0.56%outdoor	10	2	Y	Y	Y	N	Y	N	N	N

Research on PF-SLAM Indoor Pedestrian Localization Algorithm Based on Feature Point Map	2018	Shi, Jingjing; Ren, Mingrong; Wang, Pu; Meng, Juan	a foot-Mounted IMU	SLAM, KF, ZUPT, particle filter	-	1	184.6	0.26%	1	2	Y	Y	Y	Y	Y	Y	N	N
Research on Robust PDR Algorithm Based on Smart Phone	2018	Kuan, Jian; Chen, Xingeng; Niu, XiaoJi	A smartphone	EKF	1.9indoor, 4.3outdoor	1	205indoor, 390outdoor	0.9%indoor, 1.1%outdoor	10	2	Y	Y	Y	Y	Y	Y	N	N
Robust pedestrian dead reckoning for indoor positioning using smartphone	2015	Myung Chul Park; V. V. Chirakkal; Dong Seog Han	A smartphone	-	2	1	47	4.25%	1	2	Y	N	N	N	N	Y	N	N
Self-contained Position Tracking of Human Movement Using Small Inertial/Magnetic Sensor Modules	2007	X. Yun; E. R. Bachmann; H. Moore; J. Calusdian	a Foot-Mounted IMU	Zero velocity update	-	2	24indoor, 120outdoor	5.5%indoor, 1.3%outdoor	10	3	Y	Y	N	N	N	N	N	N
Sensing strides using EMG signal for pedestrian navigation	2010	Chen, Ruizhi; Chen, Wei; Chen, Xiang; Zhang, Xu; Chen, Yuwei	EMG sensors attached to the left and right leg	artificial neural network (ANN)	11	3	1581.8	0.69%	0	2	Y	Y	N	Y	Y	Y	N	N
Sensor-based dead-reckoning for indoor positioning	2014	Sharp, Ian; Yu, Kegen	a smartphone mounted on the hip	ZUPT	0.85	12	122.6	0.69%	1	2	Y	Y	N	N	N	N	N	N
Simulated-Zero Velocity Update Method for smartphone Navigation	2018	S. Zeng; Q. Zeng; R. Chen; Q. Meng; H. Huang; J. Liu	A smartphone	Zero Velocity Update (ZUPT), Kalman Filter	3.27	1	200	1.63%	1	2	Y	Y	Y	N	Y	Y	N	N
Small and low-cost navigation device for pedestrian	2018	X. Xing; Z. Yating	MEMS accelerometer, gyroscope	-	6.28	1	400	1.57%	1	2	N	N	N	N	N	N	N	N
Smart Insole-Based Indoor Localization System for Internet of Things Applications	2019	D. Chen; H. Cao; H. Chen; Z. Zhu; X. Qian; W. Xu; M. Huang	pressure sensors and IMUs on shoes	known veloc- ity update (KUPT) and double-foot position calibration (DFPC)	0.78	1	91	0.86%	1	2	Y	Y	Y	Y	Y	Y	N	N
SmartMTra: Put the Limit of the Inertial Sensor Based Indoor Motion Trajectory Tracing Using Smartphone	2016	Zhang, Pengyan; Chen, Xiaojiang; Fang, Dingyi; Wang, Wei; Tang, Zhanyong; Ma, Yang	A smartphone	domain-divided statistics model	5.2	5	150	3.46%	1	3	Y	Y	N	Y	Y	Y	N	N
Smartphone Indoor Localization with Accelerometer and Gyroscope	2014	H. Hsu; W. Peng; T. K. Shih; T. Pai; K. L. Man	A smartphone	-	0.43	1	10	6.90%	1	2	Y	Y	N	Y	Y	N	N	N
Smartphone sensor based algorithms for Dead Reckoning using magnetic field sensor and accelerometer for localization purposes	2014	D. Caspari; L. Riedhammer; M. Strutu; U. Grossmann	A smartphone	-	2.1	1	34.2	6.14%	1	2	Y	N	N	Y	N	N	Y	N
Smartphone-based Pedestrian Dead Reckoning as an indoor positioning system	2012	A. R. Pratama; Widyawan; R. Hidayat	A smartphone	peak detection method, Weinberg approach, Scarlet approach, Kim approach	1.39	15	20	6.95%	1	2	Y	Y	N	Y	Y	Y	N	N
Spatial direction corrections to improve indoor localization using inertial navigation with sensors on a smartphone	2016	H. D. R. Lakmal; J. Samarabandu	A smartphone	Scarlet method, Gaussian model	3.3	1	124.36	2.65%	1	2	Y	Y	N	N	Y	Y	N	N
Step Detection for ZUPT-Aided Inertial Pedestrian Navigation System Using Foot-Mounted Permanent Magnet	2016	A. Norrdine; Z. Kasmi; J. Blankenbach	IMU and permanent magnet attached respectively on the right and left foot	ZUPT, Kalman filter	0.11	1	30.5	0.36%	1	2	Y	Y	Y	Y	Y	N	Y	N
Strap-down Pedestrian Dead-Reckoning system	2011	P. Goyal; V. J. Ribeiro; H. Saran; A. Kumar	a waist-worn IMU	quaternion-based extended Kalman filter (EKF), Wienberg method	3.61	9	94.37	4.00%	1	3	Y	N	Y	Y	Y	Y	N	N
Synergism of INS and PDR in Self-Contained Pedestrian Tracking With a Miniature Sensor Module	2010	C. Huang; Z. Liao; L. Zhao	a Foot-Mounted IMU	line-phase finite impulse response (FIR) low-pass filter (LPF), stance phase updates (SPU)	1.2	1	60	2.00%	0	2	Y	Y	Y	Y	N	N	N	N
The multi-mode inertial tracking system for unconstrained indoor positioning	2016	A. Mikov	A smartphone	Stochastic Gradient Descent (SGD), Markov matrix,	-	5	-	7.00%	1	2	Y	Y	N	N	N	N	N	N
The Standing Calibration Method of MEMS Gyro Bias for Autonomous Pedestrian Navigation System	2017	Zhang, Yanshun; Yang, Xu; Xing, Xiangming; Wang, Zhanqing; Xiong, Yunqiang	Waist-mounted IMU	-	32.72	1	890	3.68%	1	2	Y	Y	Y	N	Y	Y	N	N
Two-mode navigation method for low-cost inertial measurement unit-based indoor pedestrian navigation	2016	Xu, Yuan; Chen, Xiyuan; Wang, Yimin	foot-mounted IMU and the shoulder- mounted IMU	Kalman filter (KF)	1.9265	1	176	1.09%	1	2	Y	Y	N	N	N	Y	N	N
User Localization Using Wearable Electromagnetic Tracker and Orientation Sensor	2006	A. Hamaguchi; M. Kanbara; N. Yokoya	an electromagnetic sensor, an inertial sensor at hip, push button switches on both heels	-	-	1	200	5.20%	1	3	Y	N	N	N	N	N	N	N
Using Step Size and Lower Limb Segment Orientation from Multiple Low-Cost Wearable Inertial/Magnetic Sensors for Pedestrian Navigation	2019	Tjhai, Chandra; O'Keefe, Kyle	7 inertial sensors mounted on the back of pelvis and on the front of each thigh, shank, and foot	Weinberg's model, Skog's generalized likelihood ratio test (GLRT) detector	-	1	94.761	4.90%	1	3	Y	Y	Y	N	N	N	Y	N
Waist mounted Pedestrian Dead-Reckoning system	2012	Jaehyun Park; Yunki Kim; Jangmyung Lee	a waist-worn IMU	HDR(Heuristic Drift Reduction) algorithm	-	1	64.48	3.00%	1	2	Y	Y	N	N	N	N	N	N
Walking compass with head-mounted IMU sensor	2016	J. Windau; L. Itti	a head-mounted IMU (Google Glass)	walking pattern analysis, EKF fusion algorithm	-	1	35	2.50%	0	2	Y	Y	Y	N	Y	Y	N	N
Walking Direction Estimation Based on Statistical Modeling of Human Gait Features With Handheld MIMU	2017	C. Combettes; V. Renaudin	Handheld MIMU	Gaussian MixtureModel (GMM), WAlking direction estimation based on Inertial Signal Statistics (WAISS) algorithm	-	4	1000	5.60%	0	2	Y	Y	Y	Y	N	Y	Y	N

Bibliography

[1] Ossama Abdel-Hamid, Abdel-rahman Mohamed, Hui Jiang, Li Deng, Gerald Penn, and Dong Yu. Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on audio, speech, and language processing*, 22(10):1533–1545, 2014.

[2] Muhammad Haris Afzal, Valérie Renaudin, and Gérard Lachapelle. Use of earth's magnetic field for mitigating gyroscope errors regardless of magnetic perturbation. *Sensors*, 11(12):11390–11414, 2011.

[3] Abdulrahman Alarifi, AbdulMalik Al-Salman, Mansour Alsaleh, Ahmad Alnafessah, Suheer Al-Hadhrami, Mai A Al-Ammar, and Hend S Al-Khalifa. Ultra wideband indoor positioning technologies: Analysis and recent advances. *Sensors*, 16(5):707, 2016.

[4] Michael Angermann and Patrick Robertson. Footslam: Pedestrian simultaneous localization and mapping without exteroceptive sensors—hitchhiking on human perception and cognition. *Proceedings of the IEEE*, 100(Special Centennial Issue):1840–1848, 2012.

[5] Satoshi Asano, Yuki Wakuda, Noboru Koshizuka, and Ken Sakamura. A robust pedestrian dead-reckoning positioning based on pedestrian behavior and sensor validity. In *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium*, pages 328–333. IEEE, 2012.

[6] Rania Ashkar, Michailas Romanovas, Vadim Goridko, Manuel Schwaab, Martin Traechtler, and Yiannos Manoli. A low-cost shoe-mounted inertial navigation

system with magnetic disturbance compensation. In *International Conference on Indoor Positioning and Indoor Navigation*, pages 1–10. IEEE, 2013.

[7] C Dario Bellicoso, Marko Bjelonic, Lorenz Wellhausen, Kai Holtmann, Fabian Günther, Marco Tranzatto, Peter Fankhauser, and Marco Hutter. Advances in real-world applications for legged robots. *Journal of Field Robotics*, 35(8):1311–1326, 2018.

[8] Jeroen HM Bergmann, Caroline Alexiou, and Ian CH Smith. Procedural differences directly affect timed up and go times. *Journal of the American Geriatrics Society*, 57(11):2168–2169, 2009.

[9] Jeroen HM Bergmann, Vikesh Chandaria, and Alison McGregor. Wearable and implantable sensors: the patient's perspective. *Sensors*, 12(12):16695–16709, 2012.

[10] Jeroen HM Bergmann, Patrick M Langdon, Ruth E Mayagoitia, and Newton Howard. Exploring the use of sensors to measure behavioral interactions: an experimental evaluation of using hand trajectories. *PLoS One*, 9(2):e88080, 2014.

[11] JHM Bergmann and AH McGregor. Body-worn sensor design: what do patients and clinicians want? *Annals of biomedical engineering*, 39:2299–2312, 2011.

[12] James Betker, Gabriel Goh, Li Jing, Tim Brooks, Jianfeng Wang, Linjie Li, Long Ouyang, Juntang Zhuang, Joyce Lee, Yufei Guo, Wesam Manassra, Prafulla Dhariwal, Casey Chu, Yunxin Jiao, and Aditya Ramesh.

[13] Stephanie Blair, Grant Duthie, Sam Robertson, William Hopkins, and Kevin Ball. Concurrent validation of an inertial measurement system to quantify kicking biomechanics in four football codes. *Journal of biomechanics*, 73:24–32, 2018.

[14] Johann Borenstein, Lauro Ojeda, and Surat Kwanmuang. Heuristic reduction of gyro drift for personnel tracking systems. *The Journal of navigation*, 62(1):41–58, 2009.

[15] Dan Brett. Mouthguard evolution. `https://www.dentistryiq.com/dentistry/pediatric-dentistry/article/16348536/mouthguard-evolution`. Accessed January 1st, 2003.

[16] Tim Brooks, Bill Peebles, Connor Homes, Will DePue, Yufei Guo, Li Jing, David Schnurr, Joe Taylor, Troy Luhman, Eric Luhman, Clarence Yin Ng Wing, Ricky Wang, and Aditya Ramesh. Video generation models as world simulators. *OpenAI*, 2024.

[17] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[18] Raymond C Browning, Emily A Baker, Jessica A Herron, and Rodger Kram. Effects of obesity and sex on the energetic cost and preferred speed of walking. *Journal of applied physiology*, 100(2):390–398, 2006.

[19] Wallace H Campbell. *Introduction to geomagnetic fields*. Cambridge University Press, 2003.

[20] Stefano Cardarelli, Paola di Florio, Alessandro Mengarelli, Andrea Tigrini, Sandro Fioretti, and Federica Verdini. Magnetometer-free sensor fusion applied to pedestrian tracking: A feasibility study. In *2019 IEEE 23rd International Symposium on Consumer Technologies (ISCT)*, pages 238–242. IEEE, 2019.

[21] MA Chattha and Ijaz Haider Naqvi. Pilot: A precise imu based localization technique for smart phone users. In *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, pages 1–5. IEEE, 2016.

[22] Changhao Chen, Xiaoxuan Lu, Andrew Markham, and Niki Trigoni. Ionet: Learning to cure the curse of drift in inertial odometry. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[23] Diliang Chen, Huiyi Cao, Huan Chen, Zetao Zhu, Xiaoye Qian, Wenyao Xu, and Ming-Chun Huang. Smart insole-based indoor localization system for internet of things applications. *IEEE Internet of Things Journal*, 6(4):7253–7265, 2019.

[24] Wei Chen, Ruizhi Chen, Xiang Chen, Xu Zhang, Yuwei Chen, Jianyu Wang, and Zhongqian Fu. Comparison of emg-based and accelerometer-based speed estimation methods in pedestrian dead reckoning. *The Journal of Navigation*, 64(2):265–280, 2011.

[25] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

[26] Namchol Choe, Hongyu Zhao, Sen Qiu, and Yongguk So. A sensor-to-segment calibration method for motion capture system based on low cost mimu. *Measurement*, 131:490–500, 2019.

[27] Namchol Choe, Hongyu Zhao, Sen Qiu, and Yongguk So. A sensor-to-segment calibration method for motion capture system based on low cost mimu. *Measurement*, 131:490–500, 2019.

[28] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.

[29] Alejandro Correa, Marc Barcelo, Antoni Morell, and Jose Lopez Vicario. A review of pedestrian indoor positioning systems for mass market applications. *Sensors*, 17(8):1927, 2017.

[30] Petterson CR, van Wichen AE, Antoun JS, et al. Sports mouthguards: a review. *New Zealand Dental Journal*, 116(1), 2020.

[31] Diep Dao, Chris Rizos, and Jinling Wang. Location-based services: technical and business issues. *Gps Solutions*, 6(3):169–178, 2002.

[32] William Davies, Hua Ye, and Jeroen Bergmann. Unobtrusive bioanalytics for impact–related sport activities. In *International Conference on Applied Human Factors and Ergonomics*, pages 285–293. Springer, 2018.

[33] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[34] James Diebel. Representing attitude: Euler angles, unit quaternions, and rotation vectors. *Matrix*, 58(15-16):1–35, 2006.

[35] Luis Enrique Díez, Alfonso Bahillo, Jon Otegui, and Timothy Otim. Step length estimation methods based on inertial sensors: A review. *IEEE Sensors Journal*, 18(17):6908–6926, 2018.

[36] Tri-Nhut Do, Ran Liu, Chau Yuen, Meng Zhang, and U-Xuan Tan. Personal dead reckoning using imu mounted on upper torso and inverted pendulum model. *IEEE Sensors Journal*, 16(21):7600–7608, 2016.

[37] Christopher Doer and Gert F Trommer. Radar inertial odometry with online calibration. In *2020 European Navigation Conference (ENC)*, pages 1–10. IEEE, 2020.

[38] Patrick Doetsch, Michal Kozielski, and Hermann Ney. Fast and robust training of recurrent neural networks for offline handwriting recognition. In *2014 14th international conference on frontiers in handwriting recognition*, pages 279–284. IEEE, 2014.

[39] Jeffrey Donahue, Lisa Anne Hendricks, Sergio Guadarrama, Marcus Rohrbach, Subhashini Venugopalan, Kate Saenko, and Trevor Darrell. Long-term recurrent convolutional networks for visual recognition and description. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2625–2634, 2015.

[40] Hugh Durrant-Whyte and Tim Bailey. Simultaneous localization and mapping: part i. *IEEE robotics & automation magazine*, 13(2):99–110, 2006.

[41] Leonardo de Almeida e Bueno, Man Ting Kwong, William RF Milnthorpe, Runbei Cheng, and Jeroen HM Bergmann. Applying ubiquitous sensing to estimate perceived exertion based on cardiorespiratory features. *Sports Engineering*, 24(1):1–9, 2021.

[42] Marcus Edel and Enrico Köppe. An advanced method for pedestrian dead reckoning using blstm-rnns. In *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–6. IEEE, 2015.

[43] Wael Elloumi, Abdelhakim Latoui, Raphaël Canals, Aladine Chetouani, and Sylvie Treuillet. Indoor pedestrian localization with a smartphone: A comparison of inertial and vision-based methods. *IEEE Sensors Journal*, 16(13):5376–5388, 2016.

[44] Qigao Fan, Hai Zhang, Peng Pan, Xiangpeng Zhuang, Jie Jia, Pengsong Zhang, Zhengqing Zhao, Gaowen Zhu, and Yuanyuan Tang. Improved pedestrian dead reckoning based on a robust adaptive kalman filter for indoor inertial location system. *Sensors*, 19(2):294, 2019.

[45] Yuchen Fan, Yao Qian, Feng-Long Xie, and Frank K Soong. Tts synthesis with bidirectional lstm based recurrent neural networks. In *Fifteenth annual conference of the international speech communication association*, 2014.

[46] Brian Ferris, Dieter Fox, and Neil D Lawrence. Wifi-slam using gaussian process latent variable models. In *IJCAI*, volume 7, pages 2480–2485, 2007.

[47] Eric Foxlin. Pedestrian tracking with shoe-mounted inertial sensors. *IEEE Computer graphics and applications*, 25(6):38–46, 2005.

[48] Yogeswaran Ganesan, Suresh Gobee, and Vickneswary Durairajah. Development of an upper limb exoskeleton for rehabilitation with feedback from emg and imu sensor. *Procedia Computer Science*, 76:53–59, 2015.

[49] Feyissa Woyano Gobana. Survey of inertial/magnetic sensors based pedestrian dead reckoning by multi-sensor fusion method. In *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 1327–1334. IEEE, 2018.

[50] André Filipe Gonçalves Gonçalves Ferreira, Duarte Manuel Azevedo Fernandes, André Paulo Catarino, and João L. Monteiro. Localization and positioning systems for emergency responders: A survey. *IEEE Communications Surveys & Tutorials*, 19(4):2836–2870, 2017.

[51] Pragun Goyal, Vinay J Ribeiro, Huzur Saran, and Anshul Kumar. Strap-down pedestrian dead-reckoning system. In *2011 international conference on indoor positioning and indoor navigation*, pages 1–7. IEEE, 2011.

[52] Alex Graves. Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*, 2013.

[53] Fuqiang Gu, Kourosh Khoshelham, Chunyang Yu, and Jianga Shang. Accurate step length estimation for pedestrian dead reckoning localization using stacked autoencoders. *IEEE Transactions on Instrumentation and Measurement*, 68(8):2705–2713, 2019.

[54] H Guo, M Uradziński, H Yin, and M Yu. Indoor positioning based on foot-mounted imu. *Bulletin of the Polish Academy of Sciences. Technical Sciences*, 63(3):629–634, 2015.

[55] Ying Guo, Qinghua Liu, Xianlei Ji, Shengli Wang, Mingyang Feng, and Yuxi Sun. Multimode pedestrian dead reckoning gait detection algorithm based on identification of pedestrian phone carrying position. *Mobile Information Systems*, 2019:1–14, 2019.

[56] Michael Hardegger, Daniel Roggen, Sinziana Mazilu, and Gerhard Tröster. Actionslam: Using location-related actions as landmarks in pedestrian slam. In *2012 international conference on indoor positioning and indoor navigation (IPIN)*, pages 1–10. IEEE, 2012.

[57] Michael Hardegger, Daniel Roggen, and Gerhard Tröster. 3D ActionSLAM: wearable person tracking in multi-floor environments. *Personal and ubiquitous computing*, 19(1):123–141, 2015.

[58] Robert Harle. A survey of indoor inertial positioning systems for pedestrians. *IEEE Communications Surveys & Tutorials*, 15(3):1281–1293, 2013.

[59] Md Abid Hasan and Mohammad Nasimuzzaman Mishuk. Mems imu based pedestrian indoor navigation for smart glass. *Wireless Personal Communications*, 101(1):287–303, 2018.

[60] SJ Hayward, Joel Earps, R Sharpe, Kate van Lopik, J Tribe, and AA West. A novel inertial positioning update method, using passive rfid tags, for indoor asset localisation. *CIRP Journal of Manufacturing Science and Technology*, 35:968–982, 2021.

[61] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[62] Sachini Herath, Hang Yan, and Yasutaka Furukawa. Ronin: Robust neural inertial navigation in the wild: Benchmark, evaluations, & new methods. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3146–3152. IEEE, 2020.

[63] Wolfgang Hess, Damon Kohler, Holger Rapp, and Daniel Andor. Real-time loop closure in 2d lidar slam. In *2016 IEEE international conference on robotics and automation (ICRA)*, pages 1271–1278. IEEE, 2016.

[64] Jeffrey Hightower and Gaetano Borriello. Particle filters for location estimation in ubiquitous computing: A case study. In *UbiComp 2004: Ubiquitous Computing: 6th International Conference, Nottingham, UK, September 7-10, 2004. Proceedings 6*, pages 88–106. Springer, 2004.

[65] Benjamin R Hindle, Justin WL Keogh, and Anna V Lorimer. Validation of spatiotemporal and kinematic measures in functional exercises using a minimal modeling inertial sensor methodology. *Sensors*, 20(16):4586, 2020.

[66] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[67] Brien A Holden, Timothy R Fricke, David A Wilson, Monica Jong, Kovin S Naidoo, Padmaja Sankaridurg, Tien Y Wong, Thomas J Naduvilath, and Serge Resnikoff. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology*, 123(5):1036–1042, 2016.

[68] Xinyu Hou and Jeroen Bergmann. A pedestrian dead reckoning method for head-mounted sensors. *Sensors*, 20(21):6349, 2020.

[69] Xinyu Hou and Jeroen Bergmann. Pedestrian dead reckoning with wearable sensors: A systematic review. *IEEE Sensors Journal*, 21(1):143–152, 2020.

[70] Xinyu Hou and Jeroen Bergmann. Headslam: pedestrian slam with head-mounted sensors. *Sensors*, 22(4):1593, 2022.

[71] Xinyu Hou and Jeroen HM Bergmann. Hinnet: Inertial navigation with head-mounted sensors using a neural network. *Engineering Applications of Artificial Intelligence*, 123:106066, 2023.

[72] Xinyu Hou, Jiaying You, and Pingzhao Hu. Predicting Drug-Drug Interactions Using Deep Neural Network. In *Proceedings of the 2019 11th International Conference on Machine Learning and Computing*, pages 168–172, 2019.

[73] Ya-Wen Hsu, Shiang-Shuang Huang, and Jau-Woei Perng. Application of multisensor fusion to develop a personal location and 3d mapping system. *Optik*, 172:328–339, 2018.

[74] Yu-Liang Hsu, Jeen-Shing Wang, and Che-Wei Chang. A wearable inertial pedestrian navigation system with quaternion-based extended kalman filter for pedestrian localization. *IEEE Sensors Journal*, 17(10):3193–3206, 2017.

[75] Chengkai Huang, Shanbao He, Zhuqing Jiang, Chao Li, Yupeng Wang, and Xueyang Wang. Indoor positioning system based on improved pdr and magnetic calibration using smartphone. In *2014 IEEE 25th annual international symposium on personal, indoor, and mobile radio communication (PIMRC)*, pages 2099–2103. IEEE, 2014.

[76] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[77] Jian Huang, Xiaoqiang Yu, Yuan Wang, and Xiling Xiao. An integrated wireless wearable sensor system for posture recognition and indoor localization. *Sensors*, 16(11):1825, 2016.

[78] Jingxuan Huang, Zesong Fei, Tianxiong Wang, Xinyi Wang, Fan Liu, Haijun Zhou, J Andrew Zhang, and Guohua Wei. V2x-communication assisted interfer-

ence minimization for automotive radars. *China Communications*, 16(10):100–111, 2019.

[79] David H Hubel and Torsten N Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, 148(3):574, 1959.

[80] Ghulam Hussain, Muhammad Shahid Jabbar, Jun-Dong Cho, and Sangmin Bae. Indoor positioning system: A new approach based on lstm and two stage activity classification. *Electronics*, 8(4):375, 2019.

[81] Mariana N Ibarra-Bonilla, P Jorge Escamilla-Ambrosio, J Manuel Ramirez-Cortes, and Carlos Vianchada. Pedestrian dead reckoning with attitude estimation using a fuzzy logic tuned adaptive kalman filter. In *2013 IEEE 4th Latin American Symposium on Circuits and Systems (LASCAS)*, pages 1–4. IEEE, 2013.

[82] Federica Inderst, Federica Pascucci, and Marco Santoni. 3d pedestrian dead reckoning and activity classification using waist-mounted inertial measurement unit. In *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–9. IEEE, 2015.

[83] Ho Jun Jang, Jae Min Shin, and Lynn Choi. Geomagnetic field based indoor localization using recurrent neural networks. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pages 1–6. IEEE, 2017.

[84] Antonio Ramón Jiménez and Fernando Seco. Finding objects using uwb or ble localization technology: A museum-like use case. In *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8. IEEE, 2017.

[85] Susanna Kaiser and Estefania Munoz Diaz. Pocketslam based on the principle of the footslam algorithm. In *2015 International Conference on Localization and GNSS (ICL-GNSS)*, pages 1–5. IEEE, 2015.

[86] Nal Kalchbrenner, Edward Grefenstette, and Phil Blunsom. A convolutional neural network for modelling sentences. *arXiv preprint arXiv:1404.2188*, 2014.

[87] Wonho Kang and Youngnam Han. Smartpdr: Smartphone-based pedestrian dead reckoning for indoor localization. *IEEE Sensors Journal*, 15(5):2906–2916, 2015.

[88] Victor Javier Kartsch, Simone Benatti, Pasquale Davide Schiavone, Davide Rossi, and Luca Benini. A sensor fusion approach for drowsiness detection in wearable ultra-low-power systems. *Information Fusion*, 43:66–76, 2018.

[89] Kianoosh Kazemi, Iman Azimi, Pasi Liljeberg, and Amir M Rahmani. Robust cnn-based respiration rate estimation for smartwatch ppg and imu. *arXiv preprint arXiv:2401.05469*, 2024.

[90] Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in neural information processing systems*, 30, 2017.

[91] Jeong Won Kim, Han Jin Jang, Dong-Hwan Hwang, and Chansik Park. A step, stride and heading determination for the pedestrian navigation system. *Journal of Global Positioning Systems*, 3(1-2):273–279, 2004.

[92] Itzik Klein and Omri Asraf. Stepnet—deep learning approaches for step length estimation. *IEEE Access*, 8:85706–85713, 2020.

[93] Agamemnon Krasoulis, Sethu Vijayakumar, and Kianoush Nazarpour. Multi-grip classification-based prosthesis control with two emg-imu sensors. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(2):508–518, 2020.

[94] Jian Kuan, Xingeng Chen, and XiaoJi Niu. Research on robust pdr algorithm based on smart phone. In *China Satellite Navigation Conference (CSNC) 2018 Proceedings: Volume III*, pages 673–684. Springer, 2018.

[95] Ananda Sankar Kundu, Oishee Mazumder, Prasanna Kumar Lenka, and Subhasis Bhaumik. Hand gesture recognition based omnidirectional wheelchair control using imu and emg sensors. *Journal of Intelligent & Robotic Systems*, 91:529–541, 2018.

[96] Quentin Ladetto and Bertrand Merminod. An alternative approach to vision techniques-pedestrian navigation system based on digital magnetic compass and gyroscope integration. In *6th World Multiconference on Systemics, Cybernetics and Information, Orlando, USA*, 2002.

[97] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995, 1995.

[98] Yann LeCun, Larry Jackel, Leon Bottou, A Brunot, Corinna Cortes, John Denker, Harris Drucker, Isabelle Guyon, UA Muller, Eduard Sackinger, et al. Comparison of learning algorithms for handwritten digit recognition. In *International conference on artificial neural networks*, volume 60, pages 53–60. Perth, Australia, 1995.

[99] Jin-Shyan Lee and Shih-Min Huang. An experimental heuristic approach to multi-pose pedestrian dead reckoning without using magnetometers for indoor localization. *IEEE Sensors Journal*, 19(20):9532–9542, 2019.

[100] Min Su Lee, Hojin Ju, Jin Woo Song, and Chan Gook Park. Kinematic model-based pedestrian dead reckoning for heading correction and lower body motion tracking. *Sensors*, 15(11):28129–28153, 2015.

[101] Seon-Woo Lee and Kenji Mase. Activity and location recognition using wearable sensors. *IEEE pervasive computing*, 1(3):24–32, 2002.

[102] James Lenz and S Edelstein. Magnetic sensors and their applications. *IEEE Sensors journal*, 6(3):631–649, 2006.

[103] Fan Li, Chunshui Zhao, Guanzhong Ding, Jian Gong, Chenxing Liu, and Feng Zhao. A reliable and accurate indoor localization method using phone inertial sensors. In *Proceedings of the 2012 ACM conference on ubiquitous computing*, pages 421–430, 2012.

[104] Ping Li, Ramy Meziane, Martin J-D Otis, Hassan Ezzaidi, and Philippe Cardou. A smart safety helmet using imu and eeg sensors for worker fatigue detection. In *2014 IEEE International Symposium on Robotic and Sensors Environments (ROSE) Proceedings*, pages 55–60. IEEE, 2014.

[105] Zhenwei Li, Chunlei Song, Jingyi Cai, Rui Hua, and Pei Yu. An improved pedestrian navigation system using imu and magnetometer. In *2017 International Conference on Computer Systems, Electronics and Control (ICCSEC)*, pages 1639–1642. IEEE, 2017.

[106] Dario G Liebermann, Larry Katz, Mike D Hughes, Roger M Bartlett, Jim McClements, and Ian M Franks. Advances in the application of information technology to sport performance. *Journal of sports sciences*, 20(10):755–769, 2002.

[107] Guangjun Liu, Feng Yang, Xiaofan Bao, and Tao Jiang. Robust optimization of a mems accelerometer considering temperature variations. *Sensors*, 15(3):6342–6359, 2015.

[108] Wenxin Liu, David Caruso, Eddy Ilg, Jing Dong, Anastasios I Mourikis, Kostas Daniilidis, Vijay Kumar, and Jakob Engel. Tlio: Tight learned inertial odometry. *IEEE Robotics and Automation Letters*, 5(4):5653–5660, 2020.

[109] Yuzhou Liu and DeLiang Wang. Time and frequency domain long short-term memory for noise robust pitch tracking. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5600–5604. IEEE, 2017.

[110] Lennart Ljung. Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(1):36–50, 1979.

[111] Huan Luo, Yaxin Li, Jingxian Wang, Duojie Weng, Junhua Ye, Li-Ta Hsu, and Wu Chen. Integration of gnss and ble technology with inertial sensors for real-time positioning in urban environments. *IEEE Access*, 9:15744–15763, 2021.

[112] Minh-Thang Luong, Ilya Sutskever, Quoc V Le, Oriol Vinyals, and Wojciech Zaremba. Addressing the rare word problem in neural machine translation. *arXiv preprint arXiv:1410.8206*, 2014.

[113] Robert Mahony, Tarek Hamel, and Jean-Michel Pflimlin. Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on Automatic Control*, 53(5):1203–1218, 2008.

[114] Robert Mahony, Tarek Hamel, and Jean-Michel Pflimlin. Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on automatic control*, 53(5):1203–1218, 2008.

[115] Xiaoli Meng, Shuyan Sun, Lianying Ji, Jiankang Wu, and Wai-Choong Wong. Displacement estimation in micro-sensor motion capture. In *2010 IEEE International Conference on Systems, Man and Cybernetics*, pages 2611–2618. IEEE, 2010.

[116] Aleksandr Mikov, Alex Moschevikin, Alexander Fedorov, and Axel Sikora. A localization system using inertial measurement units from wireless commercial hand-held devices. In *International Conference on Indoor Positioning and Indoor Navigation*, pages 1–7. IEEE, 2013.

[117] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, Prisma Group, et al. Preferred reporting items for systematic reviews and meta-

analyses: the prisma statement. *International journal of surgery*, 8(5):336–341, 2010.

[118] Michael Montemerlo, Sebastian Thrun, Daphne Koller, Ben Wegbreit, et al. Fastslam: A factored solution to the simultaneous localization and mapping problem. *Aaai/iaai*, 593598, 2002.

[119] Raul Mur-Artal, Jose Maria Martinez Montiel, and Juan D Tardos. Orb-slam: a versatile and accurate monocular slam system. *IEEE transactions on robotics*, 31(5):1147–1163, 2015.

[120] Paul Newman, David Cole, and Kin Ho. Outdoor slam using visual appearance and laser ranging. In *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, pages 1180–1187. IEEE, 2006.

[121] OpenAI. Gpt-4 technical report, 2023.

[122] Jaehyun Park, Yunki Kim, and Jangmyung Lee. Waist mounted pedestrian dead-reckoning system. In *2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pages 335–336. IEEE, 2012.

[123] Soyoung Park, Jae Hong Lee, and Chan Gook Park. Robust pedestrian dead reckoning for multiple poses in smartphones. *IEEE Access*, 9:54498–54508, 2021.

[124] Harri Pensas, Miika Valtonen, and Jukka Vanhala. Wireless sensor networks energy optimization using user location information in smart homes. In *2011 International Conference on Broadband and Wireless Computing, Communication and Applications*, pages 351–356, 2011.

[125] Johan Perul and Valérie Renaudin. Learning individual models to estimate the walking direction of mobile phone users. *IEEE Sensors Journal*, 19(24):12306–12315, 2019.

[126] Pietro Picerno, Andrea Cereatti, and Aurelio Cappozzo. Joint kinematics estimate using wearable inertial and magnetic sensing modules. *Gait & posture*, 28(4):588–595, 2008.

[127] Sebastian Porebski, Piotr Porwik, Ewa Straszecka, and Tomasz Orczyk. Liver fibrosis diagnosis support using the Dempster–Shafer theory extended for fuzzy focal elements. *Engineering Applications of Artificial Intelligence*, 76:67–79, 2018.

[128] Alwin Poulose, Odongo Steven Eyobu, and Dong Seog Han. An indoor position-estimation algorithm using smartphone imu sensor data. *Ieee Access*, 7:11165–11177, 2019.

[129] Thierry Pozzo, Alain Berthoz, and Loic Lefort. Head stabilization during various locomotor tasks in humans. *Experimental brain research*, 82(1):97–106, 1990.

[130] Javier Prieto, Santiago Mazuelas, and Moe Z Win. Context-aided inertial navigation via belief condensation. *IEEE Transactions on Signal Processing*, 64(12):3250–3261, 2016.

[131] Maria Garcia Puyol, Dmytro Bobkov, Patrick Robertson, and Thomas Jost. Pedestrian simultaneous localization and mapping in multistory buildings using inertial sensors. *IEEE Transactions on Intelligent Transportation Systems*, 15(4):1714–1727, 2014.

[132] Maria Garcia Puyol, Dmytro Bobkov, Patrick Robertson, and Thomas Jost. Pedestrian simultaneous localization and mapping in multistory buildings using inertial sensors. *IEEE Transactions on Intelligent Transportation Systems*, 15(4):1714–1727, 2014.

[133] Wen Qi and Hang Su. A Cybertwin Based Multimodal Network for ECG Patterns Monitoring Using Deep Learning. *IEEE Transactions on Industrial Informatics*, 18(10):6663–6670, 2022.

[134] Sen Qiu, Zhelong Wang, Hongyu Zhao, Kairong Qin, Zhenglin Li, and Huosheng Hu. Inertial/magnetic sensors based pedestrian dead reckoning by means of multi-sensor fusion. *Information Fusion*, 39:108–119, 2018.

[135] Cliff Randell, Chris Djiallis, and Henk Muller. Personal position measurement using dead reckoning. In *Seventh IEEE International Symposium on Wearable Computers, 2003. Proceedings.*, pages 166–166. IEEE Computer Society, 2003.

[136] Valérie Renaudin, Melania Susi, and Gérard Lachapelle. Step length estimation using handheld inertial sensors. *Sensors*, 12(7):8507–8525, 2012.

[137] David Ribas, Pere Ridao, Jose Neira, and Juan D Tardos. Slam using an imaging sonar for partially structured underwater environments. In *2006 IEEE/RSJ international conference on intelligent robots and systems*, pages 5040–5045. IEEE, 2006.

[138] Andrew Richards. *University of Oxford Advanced Research Computing*, August 2015.

[139] J Romme, JHC van den Heuvel, G Dolmans, G Selimis, K Philips, and H De Groot. Measurement and analysis of UWB radio channel for indoor localization in a hospital environment. In *2014 IEEE International Conference on Ultra-WideBand (ICUWB)*, pages 274–279. IEEE, 2014.

[140] Laura Ruotsalainen, Martti Kirkko-Jaakkola, Jesperi Rantanen, and Maija Mäkelä. Error modelling for multi-sensor measurements in infrastructure-free indoor navigation. *Sensors*, 18(2):590, 2018.

[141] Rebecca L Russell and Christopher Reale. Multivariate uncertainty in deep learning. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12):7937–7943, 2021.

[142] Haşim Sak, Andrew Senior, and Françoise Beaufays. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. *Interspeech 2014*, 2014.

[143] Jim Scarlett. Enhancing the performance of pedometers using a single accelerometer. *Application Note, Analog Devices*, 41, 2007.

[144] Mike Schuster and Kuldip K Paliwal. Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11):2673–2681, 1997.

[145] Nabil Shaukat, Ahmed Ali, Muhammad Javed Iqbal, Muhammad Moinuddin, and Pablo Otero. Multi-sensor fusion for underwater vehicle localization by augmentation of rbf neural network and error-state kalman filter. *Sensors*, 21(4):1149, 2021.

[146] Ling-Feng Shi, Yu-Le Zhao, Gong-Xu Liu, Sen Chen, Yue Wang, and Yi-Fan Shi. A robust pedestrian dead reckoning system using low-cost magnetic and inertial sensors. *IEEE Transactions on Instrumentation and Measurement*, 68(8):2996–3003, 2018.

[147] Beomju Shin, Chulki Kim, Jaehun Kim, Seok Lee, Changdon Kee, Hyoung Seok Kim, and Taikjin Lee. Motion recognition-based 3d pedestrian navigation system using smartphone. *IEEE Sensors Journal*, 16(18):6977–6989, 2016.

[148] Beomju Shin, Seok Lee, Chulki Kim, Jaehun Kim, Taikjin Lee, Changdon Kee, Sujeong Heo, and Heonsoo Rhee. Implementation and performance analysis of smartphone-based 3d pdr system with hybrid motion and heading classifier. In *2014 IEEE/ION Position, Location and Navigation Symposium-PLANS 2014*, pages 201–204. IEEE, 2014.

[149] SH Shin, CG Park, JW Kim, HS Hong, and JM Lee. Adaptive step length estimation algorithm using low-cost mems inertial sensors. In *2007 ieee sensors applications symposium*, pages 1–5. IEEE, 2007.

[150] Søren Kaae Sønderby and Ole Winther. Protein secondary structure prediction with long short term memory networks. *arXiv preprint arXiv:1412.7828*, 2014.

[151] Young Soo Suh and Sangkyung Park. Pedestrian inertial navigation with gait phase detection assisted zero velocity updating. In *2009 4th International Conference on Autonomous Robots and Agents*, pages 336–341. IEEE, 2009.

[152] Zuolei Sun, Xuchu Mao, Weifeng Tian, and Xiangfen Zhang. Activity classification and dead reckoning for pedestrian navigation with wearable sensors. *Measurement science and technology*, 20(1):015203, 2008.

[153] Wolfgang Teufl, Markus Miezal, Bertram Taetz, Michael Fröhlich, and Gabriele Bleser. Validity of inertial sensor based 3d joint kinematics of static and dynamic sport and physiotherapy specific movements. *PloS one*, 14(2):e0213064, 2019.

[154] Ye Tian, Ao Peng, Xueting Xu, and Weicheng Zhang. A heading estimation algorithm for wrist device assisted by sequential geomagnetic observations. *IEEE Sensors Journal*, 22(6):5309–5317, 2022.

[155] Guillaume Trehard, Sylvie Lamy-Perbal, and Mehdi Boukallel. Indoor infrastructure-less solution based on sensor augmented smartphone for pedestrian localisation. In *2012 Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS)*, pages 1–7. IEEE, 2012.

[156] Xinling Tuo and Hezhi Xie. Effectiveness of acute: Chronic workload ratio and oslo sports trauma research center questionnaire on health problems in monitoring sports load and injury of track and field athletes. In *2021 International Conference on Information Technology and Contemporary Sports (TCS)*, pages 520–523, 2021.

[157] Cardiff University. Specialist unit for review evidence (sure) 2018. questions to assist with the critical appraisal of systematic reviews.

[158] Luan Van Nguyen and Hung Manh La. Real-time human foot motion localization algorithm with dynamic speed. *IEEE Transactions on Human-Machine Systems*, 46(6):822–833, 2016.

[159] Raghav H Venkatnarayan and Muhammad Shahzad. Enhancing indoor inertial odometry with wifi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(2):1–27, 2019.

[160] Melanija Vezočnik and Matjaz B. Juric. Average step length estimation models' evaluation using inertial sensors: A review. *IEEE Sensors Journal*, 19(2):396–403, 2019.

[161] Johan Wahlström, Andrew Markham, and Niki Trigoni. Footslam meets adaptive thresholding. *IEEE Sensors Journal*, 20(16):9351–9358, 2020.

[162] Guohua Wang, Xinyu Wang, Jing Nie, and Liwei Lin. Magnetic-based indoor localization using smartphone via a fusion algorithm. *IEEE Sensors Journal*, 19(15):6477–6485, 2019.

[163] He Wang, Souvik Sen, Ahmed Elgohary, Moustafa Farid, Moustafa Youssef, and Romit Roy Choudhury. No need to war-drive: Unsupervised indoor localization. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, pages 197–210, 2012.

[164] Qu Wang, Haiyong Luo, Fang Zhao, and Wenhua Shao. An indoor self-localization algorithm using the calibration of the online magnetic fingerprints and indoor landmarks. In *2016 international conference on indoor positioning and indoor navigation (IPIN)*, pages 1–8. IEEE, 2016.

[165] Qu Wang, Langlang Ye, Haiyong Luo, Aidong Men, Fang Zhao, and Yan Huang. Pedestrian stride-length estimation based on lstm and denoising autoencoders. *Sensors*, 19(4):840, 2019.

[166] Harvey Weinberg. Using the adxl202 in pedometer and personal navigation applications. *Analog Devices AN-602 application note*, 2(2):1–6, 2002.

[167] Jens Windau and Laurent Itti. Walking compass with head-mounted imu sensor. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5542–5547. IEEE, 2016.

[168] Dongjin Wu, Linyuan Xia, and Jijun Geng. Heading estimation for pedestrian dead reckoning based on robust adaptive kalman filtering. *Sensors*, 18(6):1970, 2018.

[169] Jian Wu, Lu Sun, and Roozbeh Jafari. A wearable system for recognizing american sign language in real-time using imu and surface emg sensors. *IEEE Journal of Biomedical and Health Informatics*, 20(5):1281–1290, 2016.

[170] Xudong Wu, Ruofei Shen, Luoyi Fu, Xiaohua Tian, Peng Liu, and Xinbing Wang. ibill: Using ibeacon and inertial sensors for accurate indoor localization in large open areas. *IEEE Access*, 5:14589–14599, 2017.

[171] Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, et al. Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*, 2016.

[172] Yuan Wu, Hai-Bing Zhu, Qing-Xiu Du, and Shu-Ming Tang. A survey of the research status of pedestrian dead reckoning systems based on inertial sensors. *International Journal of Automation and Computing*, 16:65–83, 2019.

[173] Yuan Wu, Haibing Zhu, Qingxiu Du, and Shuming Tang. A pedestrian dead-reckoning system for walking and marking time mixed movement using an shss scheme and a foot-mounted imu. *IEEE Sensors Journal*, 19(5):1661–1671, 2018.

[174] Rohan Kumar Yadav, Bimal Bhattarai, Hui-Seon Gang, and Jae-Young Pyun. Trusted k nearest bayesian estimation for indoor positioning system. *IEEE Access*, 7:51484–51498, 2019.

[175] Hang Yan, Qi Shan, and Yasutaka Furukawa. Ridi: Robust imu double integration. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 621–636, 2018.

[176] Zhe Yang, Yun Pan, and Ling Zhang. Hybrid orientation filter aided indoor tracking for pedestrians using a smartphone. In *2017 13th IEEE International Conference on Control & Automation (ICCA)*, pages 107–112. IEEE, 2017.

[177] Haoyang Ye, Yuying Chen, and Ming Liu. Tightly coupled 3d lidar inertial odometry and mapping. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 3144–3150. IEEE, 2019.

[178] Xinyu Yi, Yuxiao Zhou, Marc Habermann, Soshi Shimada, Vladislav Golyanik, Christian Theobalt, and Feng Xu. Physical inertial poser (pip): Physics-aware real-time human motion tracking from sparse inertial sensors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 13167–13178, June 2022.

[179] Ning Yu, Yunfei Li, Xiaofeng Ma, Yinfeng Wu, and Renjian Feng. Comparison of pedestrian tracking methods based on foot-and waist-mounted inertial sensors and handheld smartphones. *IEEE Sensors Journal*, 19(18):8160–8173, 2019.

[180] Xinguo Yu, Ben Liu, Xinyue Lan, Zhuoling Xiao, Shuisheng Lin, Bo Yan, and Liang Zhou. Azupt: Adaptive zero velocity update based on neural networks

for pedestrian tracking. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, 2019.

[181] Xiaoping Yun, Eric R Bachmann, Hyatt Moore, and James Calusdian. Self-contained position tracking of human movement using small inertial/magnetic sensor modules. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 2526–2533. IEEE, 2007.

[182] Francisco Zampella, Fernando Seco, et al. Robust indoor positioning fusing pdr and rf technologies: The rfid and uwb case. In *International conference on indoor positioning and indoor navigation*, pages 1–10. IEEE, 2013.

[183] Hemin Zhang, Weizheng Yuan, Qiang Shen, Tai Li, and Honglong Chang. A handheld inertial pedestrian navigation system with accurate step modes and device poses recognition. *IEEE Sensors Journal*, 15(3):1421–1429, 2014.

[184] Jun-Tian Zhang, Alison C Novak, Brenda Brouwer, and Qingguo Li. Concurrent validation of xsens mvn measurement of lower limb joint angular kinematics. *Physiological measurement*, 34(8):N63, 2013.

[185] Rui Zhang, Hai Yang, Fabian Höflinger, and Leonhard M Reindl. Adaptive zero velocity update based on velocity classification for pedestrian tracking. *IEEE Sensors journal*, 17(7):2137–2145, 2017.

[186] Wenchao Zhang, Xianghong Li, Dongyan Wei, Xinchun Ji, and Hong Yuan. A foot-mounted pdr system based on imu/ekf+ hmm+ zupt+ zaru+ hdr+ compass algorithm. In *2017 International conference on indoor positioning and indoor navigation (IPIN)*, pages 1–5. IEEE, 2017.

[187] Zichao Zhang and Davide Scaramuzza. A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7244–7251. IEEE, 2018.

[188] Cheng Zhao, Li Sun, Zhi Yan, Gerhard Neumann, Tom Duckett, and Rustam Stolkin. Learning kalman network: A deep monocular visual odometry for on-road driving. *Robotics and Autonomous Systems*, 121:103234, 2019.

[189] Hongyu Zhao, Zhelong Wang, Sen Qiu, Yanming Shen, Luyao Zhang, Kai Tang, and Giancarlo Fortino. Heading drift reduction for foot-mounted inertial navigation system via multi-sensor fusion and dual-gait analysis. *IEEE Sensors Journal*, 19(19):8514–8521, 2018.

[190] Xinyang Zhao, Changhong Wang, and Marcelo H Ang. Real-time visual-inertial localization using semantic segmentation towards dynamic environments. *IEEE Access*, 8:155047–155059, 2020.

[191] Jingchun Zhou, Tongyu Yang, Weishen Chu, and Weishi Zhang. Underwater image restoration via backscatter pixel prior and color compensation. *Engineering Applications of Artificial Intelligence*, 111:104785, 2022.

[192] Yaping Zhu, Rui Zhang, Weiwei Xia, Ziyan Jia, and Lianfeng Shen. A hybrid step model and new azimuth estimation method for pedestrian dead reckoning. In *2014 Sixth International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–5. IEEE, 2014.

[193] Giulio Zizzo and Lei Ren. Position tracking during human walking using an integrated wearable sensing system. *Sensors*, 17(12):2866, 2017.