A Zero Velocity Detection Method for Soldier Navigation Based on Deep Learning

Tengda Zhu\textsuperscript{1}, Xianfei Pan\textsuperscript{1} and Shufang Zhang\textsuperscript{1}

\textsuperscript{1} College of Intelligence Science and Technology, National University of Defense Technology, Changsha, Hunan, 470073, China

*Corresponding author's e-mail: tendazhu@163.com or afeipan@126.com

Abstract. In this paper, deep learning is introduced into the zero-speed detection of individual navigation. Through the combination of one-dimensional CNN and LSTM, the accuracy of zero-speed detection is further improved, and other cumbersome adjustment procedures based on thresholds and other methods are avoided. At the same time, a method of establishing a zero-speed label using ultrasonic ranging method is also demonstrated. Since the error of running and other sports mainly comes from the device error and not the algorithm itself, in this paper, the 6-axis IMU data is first imported into the one-dimensional CNN and is divided into a part that can perform zero-speed detection and a part that is not suitable for zero-speed detection. It can be used for zero speed detection and reuse LSTM based zero speed detection. Compared with other algorithms, this method uses the deep learning method to train the appropriate model in advance and can be used without adjusting the reference. It is especially suitable for the urgency of the soldier's mission. The experimentally proven SHOE algorithm has a navigation accuracy of less than $3.5m/186.06m$ without adjustment, and the accuracy of the proposed method is stable above $1.0m/186.06m$, which improves the overall navigation accuracy.

1. Introduction

In the current high-tech local war, soldiers are the basic combat unit of the battlefield, and providing high-precision autonomous navigation for soldiers is the basis for digital troops. Single individual navigation system is the focus of current navigation system research. Although GPS and other satellite positioning systems have been widely used in outdoor positioning, they could not solve the jungle, indoor and other satellite signals weak environment accurate positioning. Especially in the real war environment, satellite and other external signals may be destroyed and interfered with by the enemy at any time, so countries have to satellite navigation as an auxiliary means, and other means not dependent on foreign signals as the main method. Navigation positioning by the Inertial Measurement Unit (IMU) is an autonomous navigation method that does not require an external signal. A miniature inertial measurement unit (MIMU) is mounted on the soldier to record high-precision motion trajectories relative to know starting points by recording various data. Unlike other positioning methods such as satellites, the positioning of inertial measurement units (IMU) does not require the assistance of external devices and does not emit signals outwards, especially in the context of covert operations in battlefield environments. Modern MIMUs are known for their small size, light weight, and high accuracy, but all have one common drawback: the drift caused by cumulative errors. There are currently many ways to handle and resolve drift due to error accumulation. One of the reliable and effective methods is to install the IMU on the soldier's foot, and the soldier obtains the pseudo-quantity of the speed during walking. In a walking gait cycle, the soldier's foot is stationary relative to the ground for a short period of time (This short period of time is called zero speed). The amount
of pseudoquantity obtained is incorporated into the navigation altogether by means of Kalman filtering. Therefore, achieving highprecision MIMU positioning relies on accurate zero-speed detection.

The common zero-speed detector is based on a fixed threshold. This type of detector cannot perform zero-speed detection of the asynchronous state efficiently and reliably. It is necessary to set different thresholds for different motion modes, which is cumbersome. Other improved zero-speed detections include adaptive methods to improve zero-speed detection in motion, depending on speed, frequency, etc. However, these adaptive methods still do not solve the correct detection of movements such as up and down stairs, crawling, etc., and at the same time be able to play a role in walking and other movements.

Common sports patterns that soldiers use in investigations and operations include walking, running, and going up and down stairs. During the experiment, we found that when running, the whole sole of the foot is in a process of dynamic balance. It is difficult to judge the true stillness of the sole of the foot, or that the ground is still at a zero speed for a short period of time, which is obviously shorter than that during normal walking. Zero-speed time, especially during the step-by-step phase of running, is higher than 500°/s, which has exceeded the range of most low-cost MIMUs on the market. At this time, the error caused by the device range is significantly higher. The effect of large, zero-speed detection is not significant for accuracy improvement. Therefore, before zero-speed detection, we introduced motion recognition to determine whether to adopt zero-speed correction for different motion patterns. However, the manual method of motion pattern recognition and label is cumbersome and complicated. At the same time, the human body is complex, which means that we need to set up a complex classifier to classify, especially the classifiers with individual differences are not suitable for complex combat environments.

In order to solve the difficulties encountered in the above individual navigation of, need to develop a kind of can automatically adapt to a variety of human characteristics, various movement modes of the adaptive algorithm to further restrain error accumulation, guarantee the individual navigation system of high precision and reliability, thus put forward in this paper, based on the depth study of the soldier navigation zero-speed detection method, use the deep learning method instead of the traditional sports recognition and zero-speed detection.

Deep learning is one of the most popular fields in machine learning at present. Deep learning refers to the simulation of human learning thinking mode through computer modeling, and the internal rules and representation levels in learning data. Deep learning models are often non-linear and networked, with strong feature extraction ability. Without human intervention, relevant features of the whole data set can be extracted from a few sample data. In this paper, the advantages of deep learning in pattern recognition are utilized. Instead of manual supervision, the original 6-axis IMU data and zero-speed tag are simply imported into the deep learning algorithm for training, so as to obtain the motion mode and zero-speed of binary classification. These two sets of data are used to further improve the accuracy of soldier positioning. It effectively solves the traditional method's dependence on threshold setting. Different individuals need different parameters and different motion modes also need to set different parameters. It has a higher intelligent level and a wider range of application. However, for the same deep learning, a large amount of relevant data needs to be collected in advance for training, and good support of CPU, GPU and other hardware devices is required. However, with the increase of users, the continuous expansion of training data and the continuous development of CPU, GPU and other hardware devices brought by technological progress, all these problems can be well solved.

2. Analysis of the principle of zero-speed detection method
With MIMU (micro inertial measurement unit) as the main inertial measurement unit, its task is to measure the linear angular velocity, linear acceleration and other information in the process of the user's motion, and thus the data can deduce the position and attitude of the soldier relative to the known starting point. The resulting linear angular velocity is integrated to get the direction estimate, while the linear acceleration is integrated once to get the velocity estimate, and again the integration gets the position estimate. Because of the different movement characteristics of different parts of the body, the experimental results show that the dynamic characteristics of the data obtained by placing MIMU in the foot are the clearest, with obvious periodic characteristics. Because the footstep information can directly reflect the speed and direction of walking, at the same time, the two feet in contact with the ground time also has a short zero-speed static state, which facilitates the recognition of the pace and speed attitude correction. But it is also precisely
because the footsteps are dynamic, the movement is complex, but also destined to contain more additional vibration and noise in the data.

As shown in the figure, after the original acceleration \(\tilde{f}\) and angular velocity data output \(\tilde{\omega}\) from MIMU are sent to the computer, the algorithm will first conduct static zero deviation correction and filter feedback compensation, and then send them to the strapdown inertial solution unit for calculation. The zero-speed detection module can be detected at zero speed by adding numerical information of velocity and angular velocity, and then the extended Kalman filter is modified at zero speed based on the test results. The final output navigation information.

The state quantity of the system is taken:

\[
X = \begin{bmatrix} \delta \phi & \delta v & \delta r & \Delta_b & \Delta_f & \eta_b & \eta_f \end{bmatrix}^T
\]

In that, \(\delta \phi\), \(\delta v\), \(\delta r\) are attitude error Angle vector, velocity error vector and position error vector of inertial navigation device respectively. \(\Delta_b\), \(\Delta_f\), \(\eta_b\), \(\eta_f\) are zero deviation error and scale factor error of accelerometer and zero deviation error and scale factor error of gyro respectively. The status models are:

\[
\begin{align*}
\dot{\delta \phi} &= -\omega_m \times \delta \phi + \zeta \\
\dot{\delta v} &= -\delta \phi \times f - (2\omega_e + \omega_m) \times \delta v + \zeta \\
\dot{\delta r} &= -\omega_m \times \delta r + \delta v \\
\dot{\Delta}_b &= 0, \dot{\Delta}_f = 0, \dot{\eta}_b = 0, \dot{\eta}_f = 0
\end{align*}
\]

In the formula, \(\omega_m\) the rotation angle velocity of the navigation system relative to the inertial system, the rotation angle velocity of the navigation system relative to the earth, the error for the gyro, and the output error for the \(\Delta\) accelerometer.

Select the North East (N, E, D) coordinate system as the navigation system, and the system equation is:

\[
\dot{X} = FX + W
\]

In the formula, \(F\) is transfer matrix, \(W\) is the system noise, \(F\) can be obtained by linearizing the above equation. The observational equations are:

\[
Z = HX + V
\]

\[
H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}
\]

\(V\) to measure noise. The Kalman filtering is carried out by the established model system, the amount of error state is estimated and zero-speed correction is realized. It can be seen that zero-speed detection and correction is the basis for accurate pedestrian navigation.
The commonly used zero-speed detection algorithm is to take an appropriate sliding window, traverse the entire acceleration or angular velocity data over time, and seek variance seeking the acceleration or angular velocity sequence within each window, using the established threshold to filter the results. If it is greater than the threshold, the footsteps are considered to be swinging at the moment, otherwise they are considered to be at a stationary zero. At present, zero-speed detection algorithms can be broadly divided into three categories: zero-speed detection based on acceleration information, zero-speed detection based on angular velocity information, and gait detection algorithms that combine acceleration and angular velocity information. The method is based on the model variance and periodic variation characteristics of the data by setting a threshold to determine the current moment at the zero-speed moment of landing stationary or lift-step moment, and the threshold setting is generally based on the user's physical characteristics and walking habits in advance. Most scholars set multiple thresholds to detect the zero-speed state at different paces according to the gait characteristics of human walking, and constructed a zero-speed detection algorithm adapted to multigait. However, the commonly used zero-speed detection algorithm has some limitations. No matter what method there is a single threshold, the degree of adaptability to different walking states is not high.

In the zero-speed detection based on deep learning, an important part is the acquisition of training data and the establishment of corresponding labels, in order to obtain more accurate zero-speed labels, including manual label, shoe pressure sensor induction and other methods, but these methods are cumbersome and require a lot of manpower and material resources, And the accuracy to be verified. Through a large number of experiments, the author has simplified the method of using ultrasonic directly to use ultrasonic waves by placing sensors on the soles of shoes, detecting the phase displacement variance of the ground reflected wave and comparing the threshold. Compared with other zero-speed label establishment methods, the improved ultrasonic ranging method is easy to use, good synchronization, can be used in real-time third-party zero-speed detection and navigation solutions. The figure a result of IMU and ultrasonic zero-speed testing. From the figure we can see that compared to the traditional algorithm, the label obtained by ultrasonic ranging is closer to the true zero-speed stationary.

3. Application of 1D CNN in motion mode classification
In general, most technicians focus on the application of two-dimensional CNN networks, which are particularly suitable for image recognition applications. Articles including Brandon W and others have forced the conversion of 6-axis IMU data into an "IMU image" to deal with, removing the feature of time series issues that are closely related to chronological order, and disrupting the data to find the features in it. After such a cast, the lack of chronological characteristics, whether CNN is applied to zero-speed detection or motion. The recognition accuracy rate is certainly not high. Therefore, this paper adopts the retention time sequence one-dimensional CNN, the IMU data collected in the experiment with time order related to one-dimensional CNN processing, does not disturb its original
sequence. One-dimensional CNN and 2D CNN are mainly different in the input data dimension and feature detector sliding. As shown in the figure:

In one-dimensional CNN, this paper arranges 6-axis IMU in sequence as shown in the figure above, and because of the different motion modes, the time it takes to complete a repeat motion is different, and we do not use the commonly used by Jiang W et al. proposed to arrange the collected data into fixed length and width "active picture", and the length is also considered as a variable, updating the different lengths by the final recognition accuracy rate. Since the width is a 6-axis IMU sequence, the width remains constant. This gives the Hx6 matrix, where most previous experiments and articles have been conducted using fixed windows, while Oresti B et al. point out that there is no so-called optimal window, which depends on the particular model used, the relevant nature of the sensor, and the type of motion. In general, smaller windows can complete detection tasks more quickly, but because there is less data in each window and human activity is complex, it is difficult to identify complex movements. Javier O et al. point out that incorrect lengths can truncate a cycle of continuous motion, and errors usually occur at the beginning or end of the activity, so this article points out that with an overlap of 50%. But too much overlap can also make the model over-fit. Therefore, this paper adopts variable window and variable overlap rate, in order to find and determine the most suitable window size and overlap rate for each motion mode through motion pattern recognition. Therefore, the importance of classifying exercise patterns before the implementation of zero-speed detection is obvious.

To prevent the effects of rotation, scaling, and jitter on model accuracy due to the position, orientation, different types of shoes, and differences in each person's movement, Terry T. introduced here. The wearable data enhancement method proposed by others pre-processes the data. At the same time, the data of the gyroscope is taken as the initial error before each actual measurement for 1 minute to obtain the temperature drift of the gyroscope; considering that the data collected by the acceleration sensor only has low frequency components, pass the Butterworth low-pass filter. Filter it out:

\[
|H(\omega)|^2 = \frac{1}{1 + \left(\frac{\omega}{\omega_c}\right)^{2n}} = \frac{1}{1 + e^{2\left(\frac{\omega}{\omega_p}\right)^{2n}}}, \quad (6)
\]

and use the formula \(y_i = \frac{x_i}{\sum_{i=1}^{n} x_i}\) to normalize the number, so that the data is unified to 0 and 1.

In this paper, the IMU data flow first passes through two one-dimensional CNN filters, each column of the output matrix represents the weight of the filter, in order to prevent overfitting to reduce the complexity of the output data, the data input maximum pooling layer after two one-dimensional CNN filters. Because the human body movement is relatively complex, want to achieve the judgment of the movement attitude through the 6-axis IMU data must learn higher-level characteristics, so the introduction of the third and fourth one-dimensional CNN layer, further after the one-dimensional CNN layer usually have to go into the pool layer, here choose the average pool layer. Further according to the experimental results, we will determine whether more layers are needed to improve the accuracy, after these steps, each feature detector has only one weight left. Finally, through the Dropout layer, randomly assign the weight of neurons to zero, that is, randomly discard the data, which can be very good to remove the impact of small changes on the entire network, improve the robustness of the entire network structure. Finally, the Soft max activation function is fully connected so that the output value is the probability that the different gait corresponds to a different category, each probability and 1, and the motion pattern is judged empirically The confidence level of a category must exceed 0.85 when it belongs to a category, otherwise it is an invalid judgment that further reduces the error caused by invalid judgment.

Due to the complexity of the human movement itself, it is difficult to find a large and uniform way to navigate all the different modes of motion. This paper has been discussed earlier, in order to
improve the overall navigation accuracy, but not because of running and other sports way its accuracy is limited by the device range, resulting in the overall accuracy of the decline. When using zero-speed detection for correction processing, we do not use the zero-speed correction method for running and other feasible navigation methods.

4. Application of LSTM in zero-speed detection

Understanding LSTM must first say recursive neural network, recursive neural network internal contains a circular structure, each neuron is connected to this layer and upper and lower neurons, and the calculation results are related to the current input and historical input, the network can effectively utilize the time characteristics and thus particularly suitable for the time series of data. Because recursive neural networks are related to historical input, the problem of long-term neural networks will have serious gradient disappearance. In order to solve this problem, proposed LSTM (long-term memory network), adding a door structure to manage the input and output of data. Multiple Cell units form the entire LSTM learning network.

LSTM with its unique door structure, can be a good solution to the traditional RNN and other network gradients disappearing, each step of the time has long and short characteristics. And the time-and-step relationship is \( C_t \) connected through each cell final output. As described earlier, there is no so-called best window, depending on the specific model, the relevant nature of the sensor, and the type of motion. Therefore, before the implementation of LSTM zero-speed detection, the classification of motion mode is particularly important, using the same data pre-processing as the previous one-dimensional CNN for motion mode classification including rotational invariance processing and normalized processing, here is not repeated.

This experiment builds an experimental model through the LSTM network, and the structure of the model is shown as the figure.

![Figure 6. Flaw of LSTM.](image)

The experimentally designed network consists of 1 hidden layer, two fully connected layers, the F1 layer consists of 80 nodes, and the hidden layer has a total of 120 nodes, F3 for 2 states, zero-speed stationary and non-zero speed. Also to prevent the network from overfitting, a Dropout method is applied to randomly discard some of the data. The common Adam optimizer is also used in order to find the optimal solution more quickly. The data obtained from the previous state of motion can be used to learn using zero-speed detection, in the hope that two deep learning frameworks will be combined, improve the accuracy of zero-speed judgment and further improve the accuracy of individual navigation.

All in all, this paper introduces two deep learning models into the zero-speed detection of single-man navigation. After the data obtained by the 6-axis IMU is properly normalized and rotated invariance, the classification of motion mode recognition is imported into the 1D CNN network, if the class is suitable for zero-speed detection, the data will be imported into the LSTM network for zero-speed detection, and if the motion mode is not suitable for zero-speed detection, another method is used. Including up and down stairs. For the motion mode that can use zero speed detection, the zero-speed label and raw data are finally imported into the extended Kalman filter for navigation almost navigation almost, so as to obtain the precise position of the soldier and further improve the overall navigation accuracy.

5. Data acquisition, calculation testing and experimental verification

In order to obtain correct and reliable data, verify the correctness and feasibility of the algorithm, and compare the advantages and disadvantages of this method and the previous method, we have
developed a set of data acquisition equipment. Then the algorithm test is carried out by the method mentioned in the paper, and the feasibility of the new method proposed in this paper is finally verified.

Data acquisition equipment includes a pair of PLA 07 combat boots (to secure data acquisition equipment) and anMTi-G-700 The habitual lysage equipment and a self-developed ultrasonic ranging device (used to make standard zero-speed labels).

The mounting position of the 07 combat boots and the ultrasonic ranging equipment is fixed. When zero speed is stationary, that is, the ultrasonic distance equipment and the ground in a certain range can be determined that the moment is zero speed stationary. Ultrasonic module distance information triggered by MTi-G-700 module, This ensures that the inertial data and ultrasonic time are synchronized. The image is an example of the original output data of the ultrasonic ranging module.

Experimental comparisons show that the sampling frequency of 100hz is relatively stable, with relatively small jitters and relatively little interference. As the sampling rate increases, the small jitter and interference become smaller, and as the amount of data increases, the computing power required by the equipment will also increase, and the power consumption will increase, which is not conducive to the application of the real single-man combat environment.

Because human posture, walk length, motion attitude, etc. are different, these differences will affect the classification of motion patterns and zero speed detection, the main method sly euphemisms to solve these differences in the past are different people use different parameters, thresholds, need to different people different movement categories of parameter adjustment and threshold modification. And the use of deep learning methods of sports mode classification and zero-speed detection can save these cumbersome steps, only need to train the corresponding deep learning model before the equipment factory, can be applied to different people, without the need for the factory user to set up additional, which has to be said to be a clever method, is an example of artificial intelligence application. In order to expand the training set, a number of different height and weight of personnel were selected for data collection. Let it complete a variety of movements including walking, up and down stairs, running, moving forward and other different sports patterns, as shown in the figure.

The experiment uses Python as a programming language and Pytorch as the basic environment for deep learning, in which a one-dimensional CNN network and LSTM network are configured. To speed up the training process, two NVIDIA GeForce GTX 1080Ti graphics cards were selected for acceleration. As shown, this is a two-dimensional map of walking along the first floor of a university and returning to the starting point. In this experiment, a person who has never participated in the training set data
collection is selected as the object to be tested. It can be clearly found that the red dotted line is a planar two-dimensional map obtained by using the traditional SHOE zero-speed detection method in the case of unadjusted reference, and the black solid line is subjected to zero-speed detection using deep learning. A planar two-dimensional map obtained by navigation. In addition to the methods used for zero-speed detection, the two methods are different and the others remain unchanged. It can be clearly seen that the black solid line finally returns to the starting point, while the red dotted line still has a slight drift. This is mainly reflected in the accuracy of the two-speed judgment. Deep learning is based on a large amount of data learning, and generates a certain model, which can be applied to most situations. Traditionally, the threshold method is used. Before each experiment, a complicated parameter adjustment process is carried out. In this experiment, the parameters of SHOE are not adjusted, and the final navigation result shows a large deviation. As a soldier's particularity, emergency missions may come at any time. It is not possible to delay the smooth execution of tasks because of various parameters. The accuracy of the traditional SHOE zero-speed detection method is 3.5m/186.06m, and the depth of the navigation method based on the depth learning method proposed in this paper is less than 1m/186.06m. It can be seen that the method of using the depth learning proposed in this paper for zero-speed correction significantly improves the accuracy of navigation positioning without adjusting the parameters.

6. Summary and outlook

This article is a small attempt to introduce deep learning into individual navigation. It can be seen that deep learning is not only a proper noun for making images and other directions, but also a good application in other fields. Using the method proposed in this paper, the accuracy of individual navigation and positioning is improved to a certain extent, and various complicated parameters and threshold adjustments are eliminated. It is expected that the application of a large number of samples plus the application of deep learning algorithms can solve the cumbersome problem of zero-speed detection that has plagued the academic community for many years. However, the biggest problem for individual navigation is the accuracy of MIMU devices. With the advancement of technology, micro-inertial devices will certainly have better development.

References

[1] Wagstaff, B., & Kelly, J. (2018). Lstm-based zero-velocity detection for robust inertial navigation.
[2] Cortés, Santiago, Solin, A., & Kannala, J. (2018). Deep learning based speed estimation for constraining strapdown inertial navigation on smartphones.
[3] Skog, I., Handel, P., Nilsson, J. O., & Rantakokko, J. (2010). Zero-velocity detection—an algorithm evaluation. IEEE Transactions on Biomedical Engineering, 57(11), 2657-2666.
[4] Ribeiro, D., Mateus, A., Nascimento, J. C., & Miraldo, P. (2016). A real-time pedestrian detector using deep learning for human-aware navigation.
[5] Chen, C., Lu, X., Markham, A., & Trigoni, N. (2018). Ionet: learning to cure the curse of drift in inertial odometry.