Posture Recognition and Heading Estimation Based on Machine Learning Using MEMS Sensors

Boyuan Wang[1(✉)], Xuelin Liu[1], Baoguo Yu[2,3], Ruicai Jia[2,3], and Lu Huang[2,3]

[1] College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China
boyuan@hrbeu.edu.cn

[2] The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang 050081, China

[3] State Key Laboratory of Satellite Navigation System and Equipment Technology, Shijiazhuang 050081, China

Abstract. With the popularity of smartphones and the performance improvement of embedded sensor, the smartphone has become the most important terminal device in motion recognition and indoor positioning. In this paper, the methods of the smartphone posture recognition and the pedestrian heading estimation are proposed. We analyze the signal characteristic of the accelerometer and the gyroscope, the representative feature information is extracted and a classifier based on DT model is proposed. Besides, considering the different postures of the smartphone, we propose an improved heading estimation method, which utilizes a weighted-average operation and combines the principal component analysis-based (PCA-based) method and the angle deviation method innovatively. The results of the experiments show that the average accuracy of posture recognition is nearly 97.1%, which can satisfy the pattern recognition in the process of pedestrian navigation. The average error of the proposed heading estimation is 6.2° and the performance is improved than the single PCA-based and angle deviation method.

Keywords: Posture recognition · Machine learning · Heading estimation · MEMS sensors

1 Introduction

Location-based Service (LBS) has become an indispensable service in people's social life, and its key is the acquisition of location information [1]. Traditional LBS relies mainly on global navigation satellite systems (GNSS), such as GPS, Beidou, Glonass and Galileo systems [2]. GNSS can achieve a global wide coverage and the positioning accuracy with civil signals can be better than 10 m. However, due to the influence of signal occlusion, reflection and multipath effect in indoor environment, the effectiveness of satellite navigation signals in indoor environment is difficult to be guaranteed, which greatly restricts the development of LBS in indoor environment [3]. In recent

S. Han et al. (Eds.): AICON 2019, LNICST 287, pp. 496–508, 2019.
https://doi.org/10.1007/978-3-030-22971-9_42

years, more and more attentions have been paid to the indoor positioning, the indoor positioning systems based on different technical systems have also been widely established. Current indoor positioning technologies mainly include WIFI fingerprint [4], Bluetooth [5], pedestrian dead reckoning (PDR) [6], RFID [7], Ultra-wide Bandwidth (UWB) [8], ultrasound [9] and visible light positioning [10].

PDR positioning technology uses the built-in MEMS inertial sensors of the smartphone to calculate the relative position of pedestrian movement, the MEMS inertial sensors include accelerometer, gyroscope and magnetometer. However, in the process of pedestrians using smartphones, the posture of device will constantly change. The traditional PDR method requires pedestrians to keep their device in front of their bodies stably. However, this requirement is unrealistic in the actual navigation process. Therefore, this paper analyses the habits of pedestrians using mobile phones and uses machine learning algorithm to design classifiers for the recognition of different smartphone postures.

Current pattern recognition methods can be divided into two categories [11]: the method based on the environmental sensors and the recognition based on the mobile sensors. The method based on the environmental sensors uses the infrastructure (such as camera networks or wireless access points) to perceive human activities. Mobile sensor-based method usually uses the built-in MEMS inertial sensors in smartphones, such as accelerometers, gyroscopes, magnetometers, etc. to identify pedestrian behavior patterns. At present, mobile sensor-based methods have become very popular. Because they have no coverage restrictions and can work anywhere. At the same time, the smartphone is the devices which is more closely related to people, thus no additional deployment cost is needed [12]. For using the smartphone in indoor positioning, the motions can also be divided into two categories: one is to describe the overall movement of pedestrians, including walking, running, standing, upstairs, downstairs, elevator and so on. The other is different postures that represent people holding the smartphone, such as reading information, watching navigation interface, making calls, swinging with arms, putting it in pockets and so on. This paper mainly aims at the second type of pattern recognition, which can further assist indoor pedestrian navigation by recognizing different phone postures. Some recognition methods of motion mode recognition are also proposed, including artificial neural network (ANN) [13], support vector machine (SVM) [14], decision tree (DT) [15] and other classifiers, they are used to recognize the motion modes, capture the transition between different modes and assist the PDR.

In general, PDR includes three modules: step detection, step length estimation and heading estimation, in which the heading estimation is the most difficult and greatly influenced by the smartphone posture. The heading offset between different phone postures is estimated, and the actual heading is obtained by adding the offset to the orientation, but the offset is not constant in actual environment. In [16], the principal component analysis (PCA) of horizontal acceleration is applied to estimate the heading, but because of the poor precision of gyroscope embedded smartphone, the attitude matrix is not always accurate.

In this paper, a classifier with DT model which can recognize the current posture of the smartphone is designed. A pre-processing for the raw data from the MEMS sensors is carried out to eliminate the influence of noises and 4 phone postures are defined. Simultaneously, the feature data in time and frequency domain are extracted from the filtered signals. Furthermore, a novel heading determination method is proposed. By combining the PCA-based method and the angle deviation method, the estimated heading has higher accuracy.

2 Approach

2.1 MEMS Data Pre-processing

The errors of MEMS IMU can be divided into static error, dynamic error and random error. Static and dynamic errors are generally considered to be deterministic errors related to the velocity and acceleration of the carrier motion, which can be compensated by experimental calibration. Random drift is an important characteristic of the gyroscope, a lot of works have been done in the measurement and modeling of the gyroscope drift. However, due to the low measuring accuracy of the MEMS sensor, the performance of the calibrated IMU cannot satisfy the requirements of the indoor pedestrian positioning system. Therefore, to improve the performance of smartphone posture recognition and pedestrian heading estimation, it is necessary to pre-process the raw data output from the MEMS sensors embedded in the smartphone. This paper uses a 10th order Butterworth filter with a 12 Hz cut-off frequency to effectively eliminate the high-frequency noise interference of the MEMS signal.

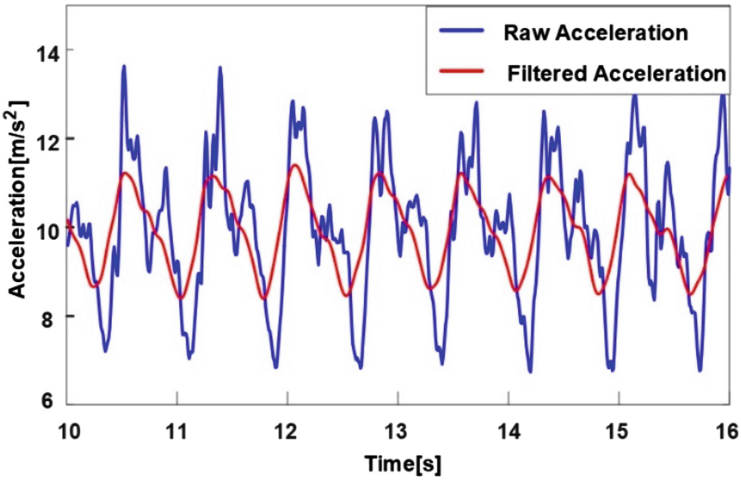

Fig. 1. Raw acceleration signal and filtered acceleration signal.

As shown in Fig. 1, we can see that the raw signals from MEMS sensors contain many clutters and noises, the low-pass filtering removes the high-frequency components of the noise, the filtered signals are smoother and the waveforms can be seen more clearly. This pre-processing operation is necessary for extracting feature information from the MEMS sensor and enables the process of smartphone posture recognition and pedestrian heading estimation.

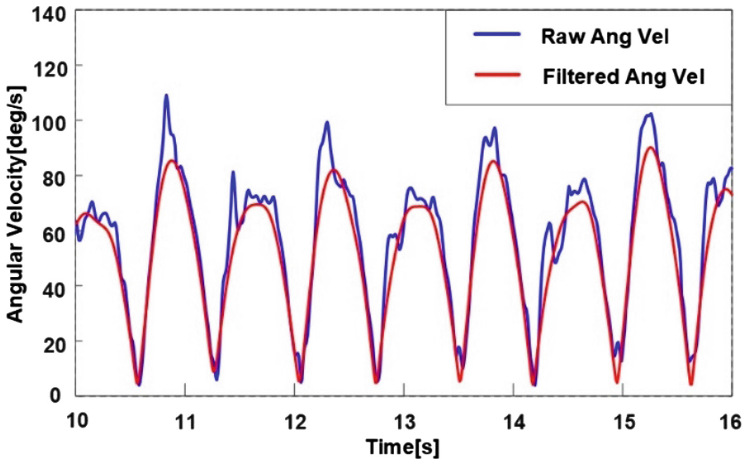

Fig. 2. Raw angel velocity signal and filtered angel velocity signal.

In general, all the coordinate systems of the accelerometer, gyroscope and magnetometer embedded in the smartphone are consistent, which are also the same as the coordinate system of the smartphone. Thus, we introduce the coordinate system of the smartphone, as shown in Fig. 2. We place the smartphone on the horizontal plane, the screen center is the origin of the coordinate system; the X axis is parallel to the short side of smartphone and the direction is horizontal to the right; the Y axis is parallel to the long side of smartphone and the direction is horizontal to the forward; the Z axis is upward and perpendicular to the screen (Fig. 3).

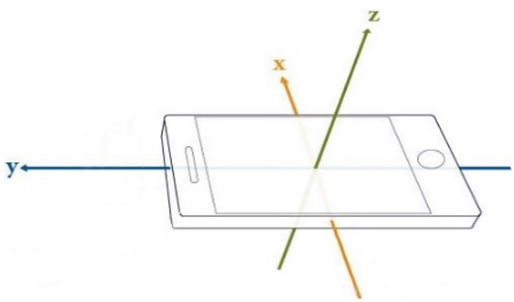

Fig. 3. The coordinate system of the embedded sensors.

2.2 Posture Recognition

The posture recognition is that we collect the current sensor signal, based on our predefined pattern categories and the data characteristics of the sensors, we can determine the current smartphone postures. The posture recognition in this paper includes three modules: postures definition, feature information extraction and class determination.

Posture Definition. According to the daily custom of using smartphones for people, we consider 4 common postures in this paper, which represent the posture for holding or placing the smartphone, including Holding, Calling, Swinging and Pocket. The indoor positioning system is related to different postures. The detailed description is listed as follow:

Holding: the case that the phone is held steadily in front of the body. In this case, the phone is stable relative to the body, and the direction of the phone represents the direction of pedestrian motion.

Calling: the case that the pedestrian makes a call and the phone screen points to the side of the body. In this case, the posture can be further divided into calling with left-hand and right-hand, that is, left-calling and right-calling.

Swinging: the case that pedestrian swings the phone with the hand. In this case, the phone approximately points to the direction of pedestrian motion.

Pocket: the case that the phone is carried in the front pocket of the trousers. In this case, we define the phone plane is approximately perpendicular to the ground when pedestrian is in static state.

Feature Information Extraction. The filtered IMU signal cannot completely satisfy the requirement of the posture recognition. We still need to extract feature information from the filtered accelerometer and gyroscope signals in a sliding window. The size of the sliding window is set to 256 samples with 50% overlap.

The module values of acceleration and angular velocity are denoted by

$$a_{mv} = \sqrt{(a_x^2 + a_y^2 + a_z^2)} \tag{1}$$

$$\omega_{mv} = \sqrt{(\omega_x^2 + \omega_y^2 + \omega_z^2)} \tag{2}$$

where a_x, a_y and a_z are the measurements from 3-axis accelerometer, ω_x, ω_y and ω_z are the measurements from 3-axis gyroscope.

We choose variance and energy spectral density as the feature information for posture recognition. The variances of accelerations and angular velocities describe the amplitude change of the pedestrian in a motion period, which are calculated by

$$\sigma_a^2 = \frac{\sum (a - \overline{a})}{N} \tag{3}$$

$$\sigma_\omega^2 = \frac{\sum (\omega - \overline{\omega})}{N} \tag{4}$$

In this paper, for the accelerometer and gyroscope signals, the variances of the module value and each axis value are extracted, which are denoted by

$$\sigma_a^2 = [\sigma_{ax}^2, \sigma_{ay}^2, \sigma_{az}^2, \sigma_{amv}^2] \tag{5}$$

$$\sigma_\omega^2 = [\sigma_{\omega x}^2, \sigma_{\omega y}^2, \sigma_{\omega z}^2, \sigma_{\omega mv}^2] \tag{6}$$

where σ_a^2 and σ_ω^2 are the acceleration variance vector and the angular velocity variance vector, respectively. They form the time-domain feature. The characteristics of acceleration variances with four postures during walking are shown in Fig. 4.

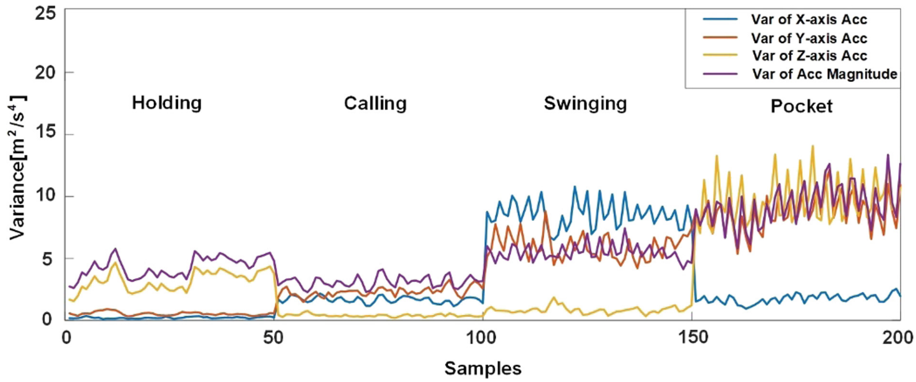

Fig. 4. The variances of accelerations with four phone poses while pedestrian is walking.

The time-frequency analysis of accelerations and angular velocities is performed by the Short Time Fourier Transform (STFT). As shown in Figs. 5 and 6, the subject walks with four phone postures, in the order of Holding, Calling, Swinging and Pocket. The period of each posture is 100 s. As shown in Fig. 6, it can be found that the energy spectral density of Z-axis angular velocities $e_{\omega z}$ shows significant peaks from 200 s to 300 s, this is due to the phone swings around Z-axis during the swinging of the hand. When the phone is carried in pocket, the phone is fixed relative to the thigh and rotates with the thigh. The energy spectral density of X-axis angular velocities $e_{\omega x}$ shows significant peaks from 300 s to 400 s in Fig. 5. The frequency-domain feature vector is expressed as $e = [e_a, e_{\omega x}, e_{\omega z}]$.

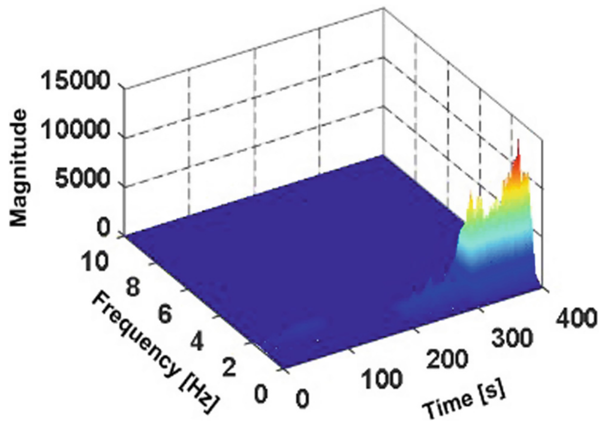

Fig. 5. The energy spectral density of X-axis angular velocities.

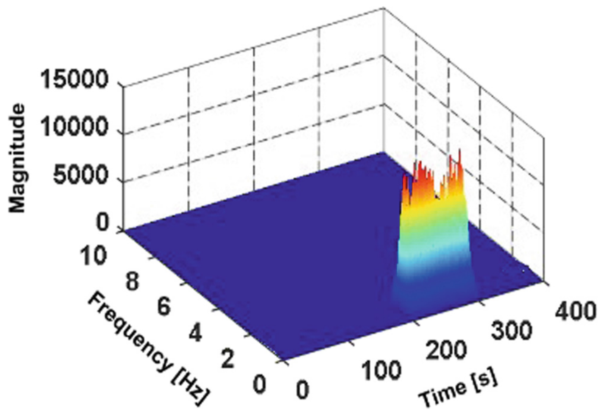

Fig. 6. The energy spectral density of Z-axis angular velocities.

Posture Determination. After the device posture definition and the feature extraction, we begin to identify the postures according to the defined categories. DT is a non-parametric classifier that has tree structure, each node represents a classification, the advantage of DT is it can directly reflect the classification process and the characteristics of the data, is easy to understand and implement. In this paper, the DT model is utilized as the classifier to classify four device postures.

The feature vector $f = [\sigma_a^2, \sigma_\omega^2, e]$ consists of the variances and the energy spectral density of accelerometer and gyroscope signals, which is chosen as the input vector of classifier and the output is the current posture. All feature data are divided into two groups, one group is training dataset for training the classifier parameters and the other group is testing dataset for verifying the recognition accuracy of the trained classifier. ε_a, $\varepsilon_{\omega z}$, λ_a, λ_ω, η_a and η_ω are the parameters of classifier. The flowchart of the classifier based on DT model is shown in Fig. 7.

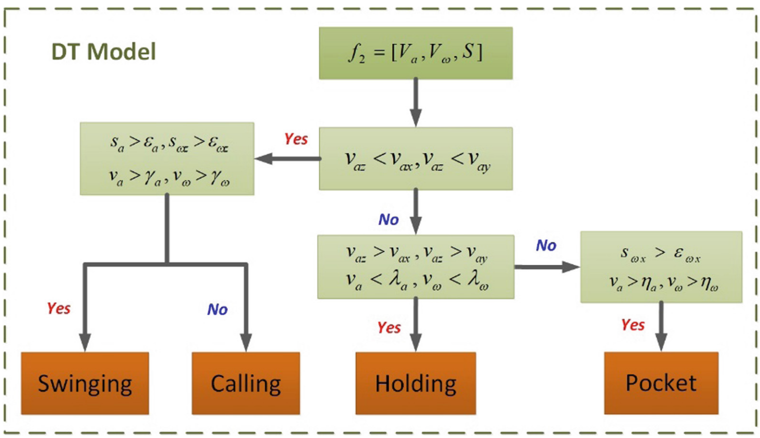

Fig. 7. The DT classifier for smartphone posture classification.

2.3 Heading Estimation Based on Smartphone Posture

To simplify the computational complexity and avoid the cumulative error caused by the integral operation of gyroscope data. The pedestrian's heading is estimated only by the accelerometer and magnetometer in this paper. Due to the magnetometer is easily disturbed by the environment, thus, our positioning work is carried out in an approximate laboratory environment to guarantee the availability of magnetometer.

We propose a novel method for estimating the heading, which combines the heading deviation determination and the PCA-based heading estimation. The Degrees of Freedom (DOF) of the smartphone is very high during the navigation process, there is a deviation between the direction of smartphone and the direction of pedestrian motion. The orientation readings from the smartphone cannot always represent the actual headings.

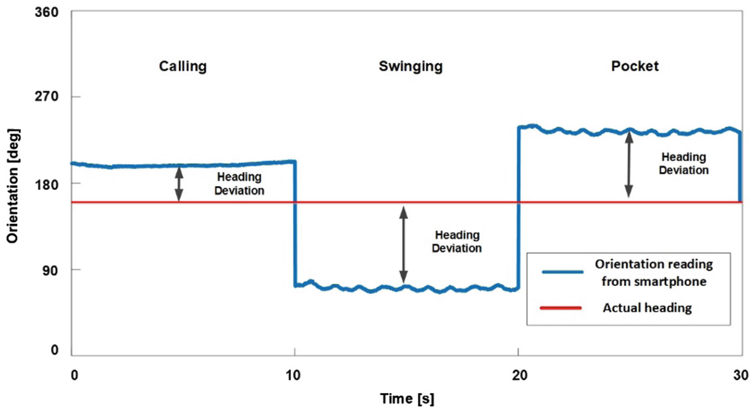

Fig. 8. The heading deviations for Holding and Calling with left-hand and right-hand.

As shown in Fig. 8, the actual heading of pedestrian can be obtained by the orientation readings from smartphone and the heading deviations. It is needed to be highlighted that the deviations are different for making a call with different hands, it is known by the directions of each accelerometer axis and described as Table 1.

Table 1. The directions of accelerometer axes for calling.

Posture	Hand	Directions of accelerometer axes		
		X-axis	Y-axis	Z-axis
Calling	Left-hand	Above and front	Above and back	Right
	Right-hand	Under and back	Above and back	Left

The directions of X-axis in vertical plane are opposite for making a call with left-hand and right-hand. Therefore, it can be determined based on the readings of accelerometer X-axis. The heading estimated by angle deviations are denoted by

$$\theta_1 = \theta_o - \theta_d \tag{7}$$

where θ_o is the orientation information from the smartphone and θ_d is the heading deviation. However, the deviations are varying in the navigation, especially for Calling and Pocket modes. Relative to the initial direction of the phone, once the phone direction has changed a lot, the obtained deviations cannot be adaptive to estimate the actual heading. Therefore, PCA method is utilized to estimate the actual heading. The PCA-based method is based on a fact that the largest variations of horizontal accelerations represent the direction of pedestrian motion, thus horizontal accelerations are needed to calculate. We obtain the horizontal accelerations using the estimated gravity vector.

$$g = -(g_x, g_y, g_z) \tag{8}$$

where g_x, g_y and g_z are the averages of the measurements on the respective axes in the time interval. Then, the vertical components of accelerations and magnetic strengths are estimated by vector dot product.

$$\begin{cases} a_v = \left(\frac{a \cdot g}{g \cdot g}\right) g \\ m_v = \left(\frac{m \cdot g}{g \cdot g}\right) g \end{cases} \tag{9}$$

where $a = (a_x, a_y, a_z)$ and $m = (m_x, m_y, m_z)$, they are the measurement vectors of the accelerometer and magnetometer respectively. Therefore, the horizontal components of accelerations and magnetic strength scan be estimated by moving the vertical components out.

$$\begin{cases} a_h = a - a_v \\ m_h = m - m_v \end{cases} \tag{10}$$

The obtained horizontal acceleration components are used as the input of PCA and the first eigenvector is regarded as the motion vector, which represents the pedestrian's direction and is denoted by

$$v = (v_x, v_y, v_z) \tag{11}$$

Then, the eigen vector v is utilized to calculate the heading θ_2 with the horizontal magnetic vector.

$$\theta_2 = \arccos\left(\frac{v \cdot m_h}{|v||m_h|}\right) \tag{12}$$

By analyzing the experimental data, we find that the headings estimated by the angle deviations are availability when the motion vector v is perpendicular to the abscissa-axis of phone. However, the abscissa-axis is approximate to Z-axis of acceleration for Calling and Swinging. Therefore, a weighted estimation algorithm is presented.

$$\begin{cases} \theta = \theta_1, & \varepsilon - 90^\circ < \delta_1 \\ \theta = (1-\kappa) \cdot \theta_1 + \kappa \cdot \theta_2, & \delta_1 < \varepsilon - 90^\circ < \delta_2 \\ \theta = \theta_2, & \varepsilon - 90^\circ > \delta_2 \end{cases} \tag{13}$$

where θ is the estimated heading by the angel deviations and motion vector obtained by PCA, ε is the angle between the motion vector v and the Z-axis of acceleration, δ_1 and δ_2 are the thresholds for judging the attitude of the smartphone.

In the case that $\varepsilon - 90^\circ < \delta_1$, that is, the v is nearly perpendicular to the Z-axis of phone and the angle deviations can be credible, so the θ_1 obtained by angle deviations is taken as the final estimated heading. In the case that $\varepsilon - 90^\circ > \delta_2$, the angle deviations cannot be used and the estimated heading is equal to θ_2 obtained by the motion vector. For $\delta_1 < \varepsilon - 90^\circ < \delta_2$, the heading θ is estimated by the weighted average operation of θ_1 and θ_2, the weighted coefficient are calculated by

$$\kappa = \frac{(\varepsilon - 90 - \delta_1)}{\delta_2 - \delta_1} \tag{14}$$

3 Experiments and Results

3.1 Experiment Field and Setup

In this section, the posture recognition and heading estimation experiments are introduced to verify the performance of the proposed methods. The experimental site is in an office building. The handheld terminal is Xiaomi MIX2 smartphone and the sampling frequency is set to 100 Hz.

3.2 Posture Recognition Experiment

We selected 10 testers to verify the posture recognition performance and the subjects move with different postures. The total data collection duration of each participant is about 10 min. The size of the sliding window for feature extraction is set to 256 samples with 50% overlap. In this experiment, 60% of the feature data are chosen for training dataset and 40% are chosen for testing dataset. The instances for testing are listed in Table 2.

Table 2. The instances of postures for testing.

	Holding	Calling	Swinging	Pocket
Walking	500	500	500	500

As shown in Table 3, the lowest accuracy of the posture recognition is 96.4% and 2.7% instances are mistaken as Swinging, this is because the phone rotates in both postures and the periodicities of the motion are similar. The highest accuracy is

Table 3. Confusion matrix for phone poses recognition.

	Holding	Calling	Swinging	Pocket
Holding	**97.3%**	1.1%	0.2%	1.4%
Calling	1.7%	**96.6%**	0.5%	1.2%
Swinging	0.2%	0.5%	**98.2%**	1.1%
Pocket	0.6%	0.3%	2.7%	**96.4%**

Swinging and can be to 98.2%, and only 0.2% instances are misclassified as Holding, this is because the postures of Holding and Calling are more stable, Swinging is significantly different from the two postures. The average accuracy is 97.1%, which can satisfy the pattern recognition in the process of pedestrian navigation.

To prove that the proposed algorithm can effectively improve the recognition accuracy, we introduce the SVM algorithm and carry out the comparative experiments. As shown in Fig. 9, the recognition accuracy of our proposed method is 8.4% higher than that of SVM method on average.

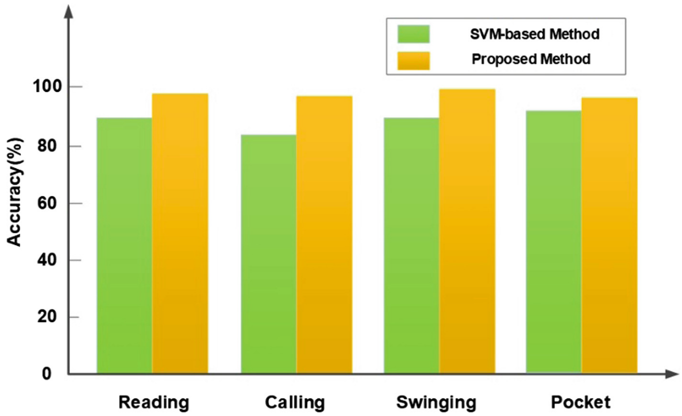

Fig. 9. The compare for the recognition accuracies of SVM and proposed method.

3.3 Heading Estimation Experiment

To verify the performance of the proposed heading estimation method, the PCA-based method and the angle deviation method are also introduced. For the comparison experiment, the smartphone is in Swinging posture. The results of the experiments are illustrated in Fig. 8. The 50% heading estimation errors for the proposed method, PCA-based and angle deviation method are 6.7°, 7.8° and 11.8°, respectively. The 75% absolute estimation errors of heading are 11.4°, 13.6° and 22.1°, respectively, The average error is 6.2°. The proposed method combines the motion axis estimated by PCA and the angle deviation, the final pedestrian's heading are obtained by the weighted average operation. Therefore, we can obviously see that the performance of the proposed method is much better (Fig. 10).

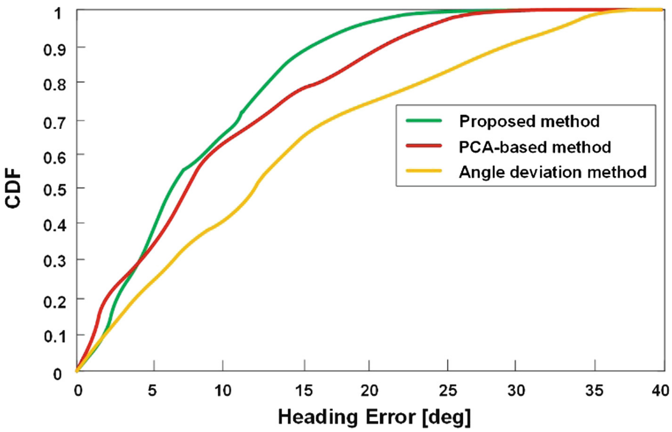

Fig. 10. The cumulative error distribution of heading estimation with the proposed method, PCA-based method and the angle deviation method.

4 Conclusions and Future Works

In this paper, the posture recognition and heading estimation method based on machine learning is proposed. The data are all from the MEMS sensors embedded in a smartphone. According to the analysis of accelerometer and gyroscope signals, the representative feature information is extracted, including the statistic feature and the frequency-domain feature. By comparing the characteristics of various features in different postures, a DT model for recognizing postures is designed. Besides, considering the attitude of the smartphone in different postures, we present an improved heading estimation method, which combines the PCA-based method and the angle deviation method innovatively. The results of the experiments show that the average accuracy of posture recognition is nearly 97.1%, which can satisfy the pattern recognition in the process of pedestrian navigation. The average error of heading estimation is 6.2°. The result of the comparison experiment shows that the performance of the proposed heading estimation method is better than the single PCA-based and angle deviation method. With the methods of this paper, the performance of the indoor pedestrian positioning will also be improved.

For the future works, we plan to research more advanced machine learning algorithms for the mode recognition of the pedestrian and smartphone. We also hope to apply the motion mode recognition to more fields of LBS.

Acknowledgments. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant HEUCF180801, and in part by the National Key Research and Development Plan of China under Grant 2016YFB0502100 and Grant 2016YFB0502103.

References

1. He, X., Jin, R., Dai, H.: Leveraging spatial diversity for privacy-aware location based services in mobile networks. IEEE Trans. Inf. Forensics Secur. **70**(99), 1 (2018)
2. Perrone, G., Vallan, A.: GNSS - Global Navigation Satellite Systems (2008)
3. Kaplan, E.D.: Understanding GPS: principles and application. J. Atmos. Solar Terr. Phys. **59**(5), 598–599 (1996)
4. Torressospedra, J., Montoliu, R., Mendozasilva, G. M., Belmonte, O., Rambla, D., Huerta, J.: Providing databases for different indoor positioning technologies: pros and cons of magnetic field and Wi-Fi based positioning. Mob. Inf. Syst. 1–22 (2016)
5. Mingchi, L.U., Wang, S., Yunke, L.I., Yuanfa, J.I., Sun, X., Deng, G.: Bluetooth location algorithm based on feature matching and distance weighting. J. Comput. Appl. 45–54 (2018)
6. Wang, B., Liu, X., Yu, B., Jia, R., Gan, X.: Pedestrian dead reckoning based on motion mode recognition using a smartphone. Sensors **18**(6), 1811 (2018)
7. Forouzannezhad, P., Jafargholi, A., Jahanbakhshi, A.: Multiband compact antenna for near-field and far-field RFID and wireless portable applications. IET Microw. Antenna. Propag. **11**(4), 535–541 (2017)
8. Zhang, C., Kuhn, M.J., Merkl, B.C., Fathy, A.E., Mahfouz, M.R.: Real-time noncoherent UWB positioning radar with millimeter range accuracy: theory and experiment. IEEE Trans. Microw. Theory Tech. **58**(1), 9–20 (2010)
9. Qi, J., Liu, G.P.: A robust high-accuracy ultrasound indoor positioning system based on a wireless sensor network. Sensors **17**(11), 2554 (2017)
10. Kuo, Y.S., Pannuto, P., Dutta, P.: Demo: Luxapose: indoor positioning with mobile phones and visible light. In: International Conference on Mobile Computing & Networking, pp. 5–11 (2014)
11. Gu, F., Khoshelham, K., Valaee, S., Shang, J., Zhang, R.: Locomotion activity recognition using stacked denoising autoencoders. IEEE Internet Things J. **5**(3), 2085–2093 (2018)
12. Jian, K.W., Liang, D., Xiao, W.: Real-time physical activity classification and tracking using wearble sensors. In: International Conference on Information, pp. 1–6 (2008)
13. Shin, B., Kim, C., Kim, J., Lee, S., Kee, C., Lee, T.: Motion recognition-based 3D pedestrian navigation system using smartphone. IEEE Sens. J. **16**(18), 6977–6989 (2016)
14. Ling, P., Jingbin, L., Robert, G., Yuwei, C., Heidi, K., Ruizhi, C.: Using LS-SVM based motion recognition for smartphone indoor wireless positioning. Sensors **12**(5), 6155–6175 (2012)
15. Rautaray, S.S., Agrawal, A.: Vision based hand gesture recognition for human computer interaction: a survey. Artif. Intell. Rev. **43**(1), 1–54 (2015)
16. Deng, Z., Si, W., Qu, Z., Xin, L., Na, Z.: Heading estimation fusing inertial sensors and landmarks for indoor navigation using a smartphone in the pocket. Eurasip J. Wirel. Comm. Netw. **2017**(1), 160 (2017)