A flexible motion tracking system based on inertial sensors

Zhonghao Han , Lei Hu, Na Guo, Biao Yang, Hongsheng Liu and Yuhan Wang

School of Mechanical Engineering and Automation, Beihang University, 100191 Beijing, China

Abstract. As a newly emerging human-computer interaction, motion tracking technology offers a way to extract human motion data. This paper presents a series of techniques to improve the flexibility of the motion tracking system based on the inertial measurement units (IMUs). First, we built a most miniatured wireless tracking node by integrating an IMU, a Wi-Fi module and a power supply. Then, the data transfer rate was optimized using an asynchronous query method. Finally, to simplify the setup and make the interchangeability of all nodes possible, we designed a calibration procedure and trained a support vector machine (SVM) model to determine the binding relation between the body segments and the tracking nodes after setup. The evaluations of the whole system justify the effectiveness of proposed methods and demonstrate its advantages compared to other commercial motion tracking system.

1 Introduction

Human motion tracking extracts joints data from human in motion. A whole range of applications become possible due to the motion tracking technique gets cheap, reliable and convenient. These applications can roughly be categorized under two titles: control and analysis [1].

In control applications, the human motion data are used to control something. These things can be physical things like manipulators or other in general exoskeletal parts, which human can control remotely so avoiding physical risks [2]. The thing can also be virtual thing, like applications in the film, game industry, virtual reality and so on [3].

Analysis applications are most used in medical and sports fields. For example, the rehabilitation described in [4] used a full-body motion tracking system to analyse the patient's exercise in order to develop customized treatment strategies.

There are many techniques to implement the motion tracking, such as optical, image-based, inertial measurement units (IMUs) based, mechanical, magnetic, acoustic, and hybrid systems [5-7]. Among these methods, the systems based on IMUs are most promising due to the versatility, robustness and almost unlimited range of motion [8]. Thanks to recent advances in the technology of micro-electromechanical systems, a miniature circuit board can integrate a three-axis accelerometer, a three-axis gyroscope, and a three-axis magnetometer simultaneously [9]. The accelerations, angular velocity, and magnetic data can be fused to estimate the orientation of the IMUs using a complementary filter or extend Kalman filter (EKF) [10-12]. Many motion tracking systems have been developed based on the inertial

sensors, the MVN of the Xsens Technologies B.V. is most outstanding one, which is usually used as gold standard to evaluate the performance of other systems. Its sensor fusion technologies enable a seamless interaction between the physical and the digital world. However, the limitation of commercial tracking systems based on IMUs lies in many aspects other than the high cost.

The first one is the size and weight of the module. There are many efforts in reducing the size of the IMUs [13-14], but this is an endless way because a smaller IMU can usually be mounted on the body segments more easily and causes fewer obstacles to human movement, which means more accurate and convenient.

Other drawbacks depend on the architecture of the motion tracking system. In general, there are two wireless structures based on IMUs. The first one is hub-based. In this architecture, all the tracking nodes connect to a central hub directly or relayed by other nodes. The hub services as a power supply and data exchange center. Due to the centralized design, this architect usually features a higher data exchange rate and easy recharge. But the setup procedure can be onerous even integrating all the tracking nodes into a tracksuit, which generally take more than ten minutes to complete with the help of the others. The entire equipment can be cumbersome when the wire around body hinders the movement of the subject. Besides, this method is only friendly to the whole-body motion capture and is not appropriate when we just want to track the part motion of our body.

The second architecture is hub-free, means that individual measurement module is independent, holding its own battery and communicating with PC station directly. In this method, the nodes usually mount the segments by the strap, so the movement of the subject

would be freer. The number of sensors can be customized according to the applications. However, the data transfer rate remains fixed in these system, means you cannot get a higher data transfer rate by using fewer nodes. Besides, in all current designs, the binding relation between the tracking nodes and the body segments is confirmed before the setup procedure. So, these systems are characterized by a configurable ID in the firmware of node and a label with the name of the body segment it mounts. In this way, the tracking nodes are not interchangeable, which complicates the setup procedure and increases the cost of mass production and maintenance.

Based on this, we propose a flexible motion tracking system based on IMUs, which has features as follows: Firstly, the tracking node is ultra-small with low power consumption and can be easily mounted on the human body. Secondly, the number of the measurement modules used in motion tracking can be arbitrary according to the applications. An asynchronous query method will always optimize the data transfer rate between the tracking nodes and PC station. Finally, the binding relation between the body segments and tracking nodes is determined using a trained support vector machine (SVM) model after setup procedure, which simplifies the setup procedure and makes the interchangeability of all nodes possible.

2 System architecture

2.1 Hardware design

The system in this paper is hub-free (see Figure 1), which means all tracking nodes are independent, containing its own IMU, wireless communication component, and power supply. The IMU focuses on determining the orientation of each individual node, and orientation data are transferred to a PC station via a WI-FI module and wireless router. To maximize the transmission speed, an ethernet cable is recommended between the PC station and router.
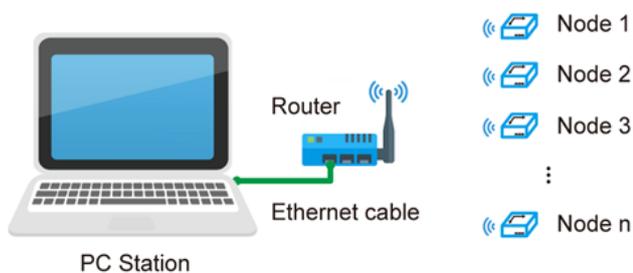

Figure 1. The hub-free architecture of the motion tracking system

The IMU is the core of the tracking node, for that it determines the accuracy of the system directly. One IMU usually consists of a three-axis accelerometer, a three-axis gyroscope, and a three-axis magnetometer. The orientation of the IMU can be estimated from the accelerations, angular velocity and magnetic data using a complementary filter or EKF. In this paper, JY901 made

by Wit-Motion was chosen as the integrated IMU. Consisting of an MPU9250 and a 32-bit high performance MCU, the JY901 features a miniaturized $15.24 \times 15.24 \times 2mm$ package size. The output data can be customized in a range of acceleration, angular velocity, magnetic field component, Euler angle and quaternion fused by EKF. The output frequency can be configured between 0.1Hz and 1kHz, which satisfies the requirements of most situations. The other specifications of JY901 are shown in Table 1.

Table 1. The specifications of JY901

Name	Value
Dynamic angle error	0.05°
Accelerator range	±16g
Angular velocity	±2000°/s
Angle	±180°
Operating voltage	3.3V
Operating current	10mA
Interface	UART, IIC

In the motion tracking field, the ZigBee and Wi-Fi protocol are usually used for wireless data communication [15]. The former provides a low-power network connection with a great deal of flexibility [16]. The theoretical maximum transfer rate is 250 Kbps [17], but in the actual test, this number is close to only 25Kbps, which may not be suitable for data-intense whole-body motion tracking. Though the power consuming of Wi-Fi protocol is relative high, the data transfer rate can be improved significantly [18]. In this system, an EPS-12S is chosen for the purpose of high-speed communication. This Wi-Fi module integrates a low power 32-bit MCU, a radio frequency (RF) component, and several communication interfaces.

When connecting the UART interface of EPS-12S to the JY901, the orientation data can be transferred to the PC station using a wireless method. To save the space, the two components were attached on the offside of the printed circle board (see Figure 2). The highly integrated design features an ultra-small size of 28×32×10mm (weight: 14g) which is the smallest tracking node in all these systems proposed in [3], [13-14]. The whole power consumption is only 330mW.
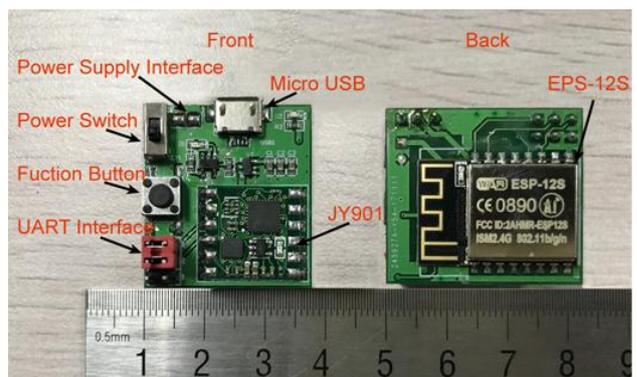

Figure 2. The structure of the tracking node

2.2 Software design

In the original design, all tracking nodes sent the data with a fixed interval time plus a small random delay but the data transfer rate reduced significantly when more modules join the network. Besides, due to the fixed transmit frequency, we cannot get a higher data transfer rate by using fewer nodes. So, an asynchronous query strategy was adopted to guarantee the data transfer rate. In this method, the EPS-12S is configured to send the orientation data from the JY901 after receiving one byte of data. For the PC station, there are three sub-program (thread).

The first is the device detecting program. When the main program starts, this program scans all the devices that connect to the router and builds the inquiry table (see Figure 3). One term of the table includes a state flag, an IP address, and a quaternion data buffer. In the process of data collecting, the program deletes the items of devices that don't return data for a period and adds the new devices that connect to the router dynamically. This program executes with a frequency of 1Hz.

The second program is the data collecting program. This program loops through the inquiry table and send one-byte data to the tracking nodes designated by the IP address with an interval of 1ms. In most instances of motion tracking, the real-time of the data is important than reliability, so the UDP protocol was adopted to maximize the data transfer rate. If the program receives one data, it will check the flag of the corresponding inquiry table term. The return quaternion data are saved in the data buffer.

The third program is the main program. This program displays the animation model after processing the quaternion in the data buffer with an update frequency of 30Hz.
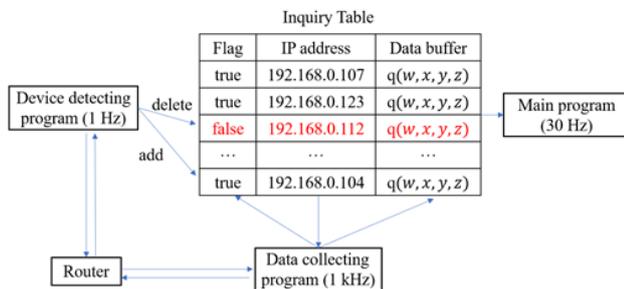

Figure 3. The program architecture and inquiry table

3 Calibration method

3.1 Theoretical analysis

For one specific IMU, the quaternion when subjects pose into a standard posture (T-Pose in this paper) is Q_P and the quaternion changes to Q_C after calibration gesture. Then, the transform quaternion from Q_P to Q_C in the world coordinate system (JY901 uses an "East North Up" (ENU) coordinate system) can be formulated as:

$$Q_T = Q_C \otimes Q_P^* \qquad (1)$$

where $\otimes$ denotes a quaternion multiplication and $*$ the complex conjugate of the quaternion.

The result can be expressed as:

$$Q_T = (w_T, x_T, y_T, z_T) \qquad (2)$$

Transfer the quaternion to axis-angle representation:

$$\theta = 2 \arccos w_T$$

$$e = (\frac{x_T}{\sin\frac{\theta}{2}}, \frac{y_T}{\sin\frac{\theta}{2}}, \frac{z_T}{\sin\frac{\theta}{2}}), \text{ if } \theta=0, \ e=(0,0,0) \qquad (3)$$

The rotation vector θ is given by:

$$\theta = \theta e = (\frac{2x_T \arccos w_T}{\sin(\arccos w_T)}, \frac{2y_T \arccos w_T}{\sin(\arccos w_T)}, \frac{2z_T \arccos w_T}{\sin(\arccos w_T)}) \qquad (4)$$

We represent rotation vector as:

$$\theta = (R_x, R_y, R_z) \qquad (5)$$

In the process of calibration, the subjects always keep perpendicular to the ground but can face any horizontal orientations. That means the rotation vector θ will rotate around the Z axis for a specific calibration gesture. However, Figure 4 shows that $R_x^2 + R_y^2$ remains constant when rotation vector θ rotates. As a consequence, a transform quaternion of an IMU for one calibration gesture can be determined by two critical components R_z and $R_x^2 + R_y^2$.
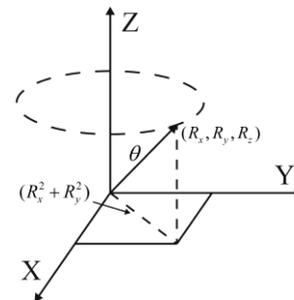

Figure 4. The $x^2 + y^2$ remains constant when rotation vector θ rotates around Z axis

3.2 Design calibration procedure

There are three guidelines for designing calibration positions:

(1) The calibration gestures should be easily performed that the subjects can make a standard pose just following the guide of tutorial or program.

(2) For one calibration gesture, the R_z or $R_x^2 + R_y^2$ components of each node should be different enough to guarantee the nodes can be classified rightly.

(3) Each single calibration gesture should identify as many nodes as possible.

Based on this, we have designed a whole calibration procedure consists of four gestures. Figure 5 shows the three views of each gesture.

Figure 5a) is the T-Pose calibration, in which the subject keeps upright with arms horizontally and thumbs forward. This pose, determining the initial quaternion of the calibration, is the basic pose of all the calibrations.

The gesture in Figure 5b) is used to determine the binding relation between the tracking nodes and upper limbs including the head. In this gesture, the subject lowers head and puts hand together toward the front. The angle between forearm and arm should be greater than 90° (145° is recommended for the accuracy of recognition). All the upper limbs should keep horizontal.

Figure 5c) and d) show the gestures to recognize the nodes bound to the left lower limb and right lower limb respectively. The two gestures are mirror symmetry. The thigh and foot should be as horizontal as possible, and the calf keeps inward. The whole leg turns out.
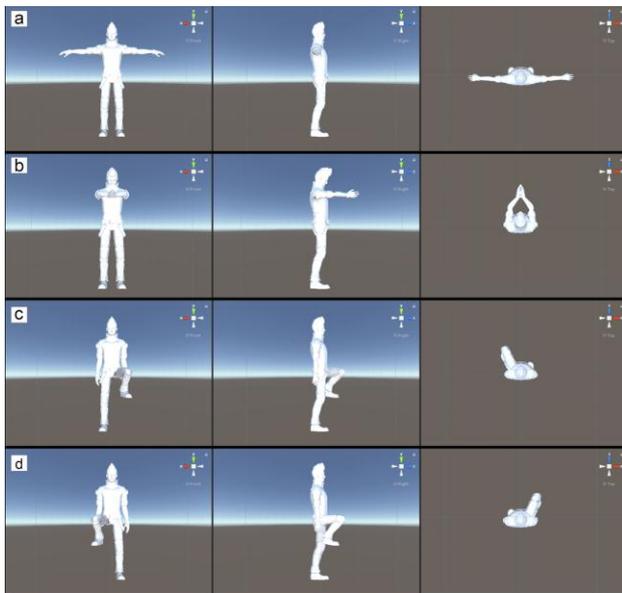

Figure 5. Calibration procedure

3.3 Training SVM model

During the model training phase, three subjects were chosen to perform 3-4 calibration procedures respectively. Figure 6 shows the points extracted from the transform quaternion of all nodes with $R_x^2 + R_y^2$ as abscissa and R_z as ordinate.

Note all these points are two-dimensional linear separable. An SVM was used to classify these points [19].

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j))$$

$$s.t. \sum_{i=1}^{n} \alpha_i y_i = 0 \ , \quad \alpha_i \geq 0, \ i = 1, 2, ..., n \quad (6)$$

where $\boldsymbol{x}$ is the vector represent of the points, y is the classification label, the kernel function we used is

$\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i^T \cdot \boldsymbol{x}_j$ and α_i can be found by using SVC learning algorithm [20].

Then, the dividing line $w^* \boldsymbol{x} + b^* = 0$ can be got by:

$$w^* = \sum_{i=1}^{N} \alpha_i^* y_i \boldsymbol{x}_i \quad (7)$$

$$b^* = y_i - \sum_{i=1}^{N} \alpha_i^* y_i (\boldsymbol{x}_i^T \cdot \boldsymbol{x}_j) \quad (8)$$

The classification decision function is given by:

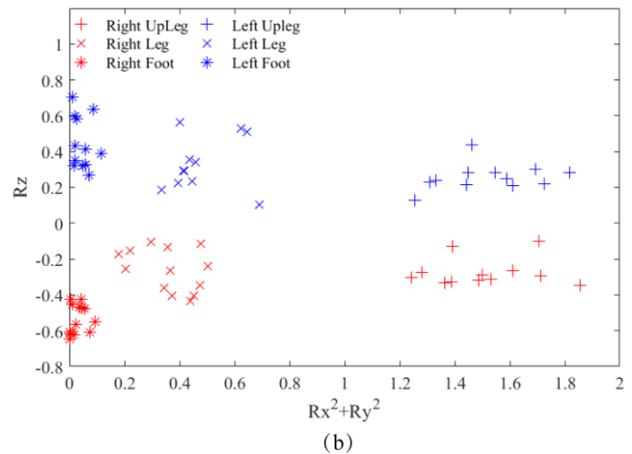$$f(\boldsymbol{x}) = sign(w^* \boldsymbol{x} + b^*) \quad (9)$$
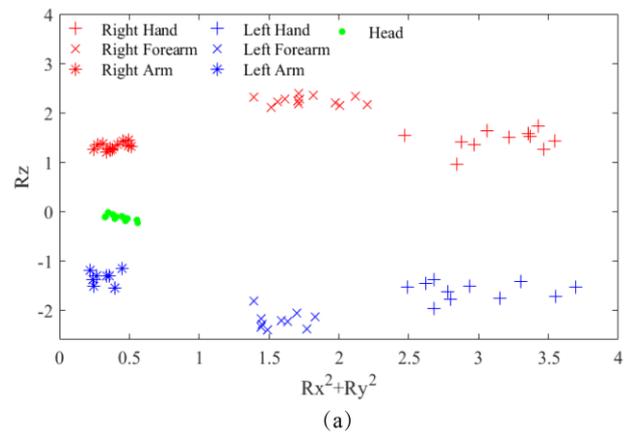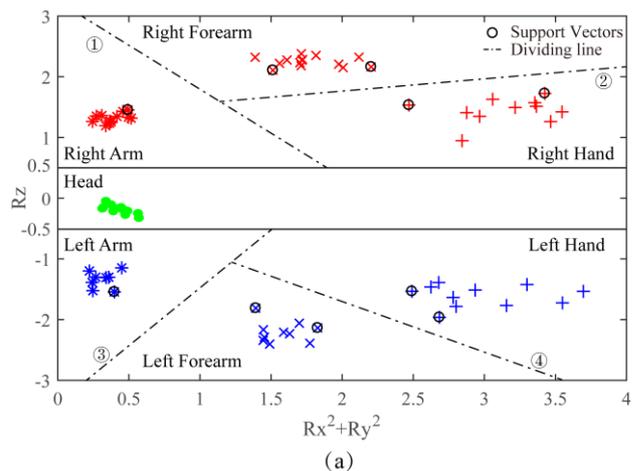

(a)

(b)

Figure 6. a) Extracted points of upper limbs and head; b) Extracted points from lower limbs
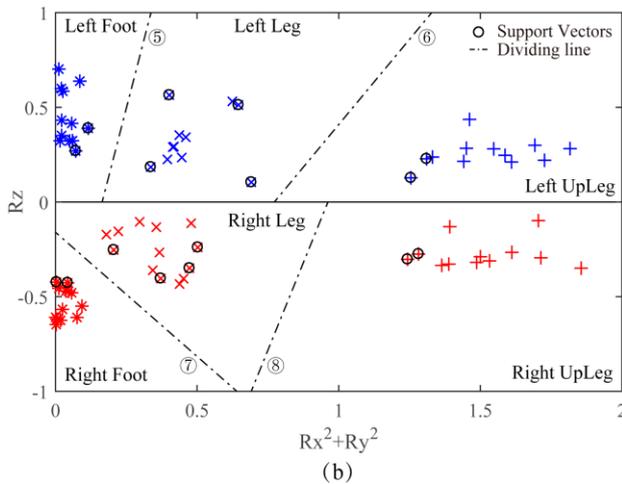

(a)

(b)

Figure 7. a) The SVM classification results of upper limbs; b) The SVM classification results of lower limbs

According to the distribution of points, we built two lines to separate the data of each limb, and the head is distinguished by the limitation of z. Compared to the traditional one versus one method for multi-classification problem, this can guarantee one point will belong to one specific class. Figure 7a) and b) show the eight dividing lines of all body segments. Table 2 shows the parameters of each dividing line represented by $y = kx + b$, which can be used as a reference.

Table 2. The parameters of each separated lines in SVM model

	L1	L2	L3	L4	L 5	L6	L7	L8
k	-1.49	0.20	1.91	-0.83	5.80	1.80	-0.91	3.69
b	3.25	1.38	-3.38	-0.04	-0.96	-1.39	-0.24	-3.55

4 Experimental evaluation

For the evaluation, the system was built following the architecture presented in the Section II. The router is TL-WR886N made by TP-Link and the network card of the PC station is RTL8168/8111/8112 Gigabit Ethernet Controller made by Realtek. An ethernet cable is used to connect the PC station to the router.

The data transfer rate using asynchronous inquiry method was investigated when different numbers of nodes access the network. A specific program was designed which can find out how many frames of data received from each node in a given period. In the experiment, the max number of tracking nodes was 17, a standard value for whole body motion tracking. The average distance between the nodes and router was about 4m. The data collection program executed with a fixed frequency of 1kHz. As a comparison, the relation between the receive frequency and transmit frequency using auto sending method was also investigated when there were 17 tracking nodes in the network.

To construct the data set for analysis, the tracking nodes were snugly strapped to the subjects' body using straps. Then, the subjects performed the calibration procedure described in the Section III under the tips of the motion tracking program. Three subjects each performed 3 calibrations and tried to face the different

direction in each calibration. The calibration time and nodes transformation quaternion were recorded.

5 Result

Figure 8a) shows relation the between data transfer rate and the number of nodes in the network using asynchronous inquiry method. The receive frequency is the average number of received frame data of all joined nodes in a period. The receive frequency is 815Hz when only one node in the network and the transfer rate decreases when more node joined the network. The average receive frequency of communicating with 17 nodes is 49 Hz. Figure 8b) shows the relation between the receive frequency and transmit frequency of 17 nodes using auto sending strategy, the maxim receive frequency is 32Hz when each node sends data with a transmit frequency of 40Hz. The results confirm the effectiveness of the asynchronous inquiry method.
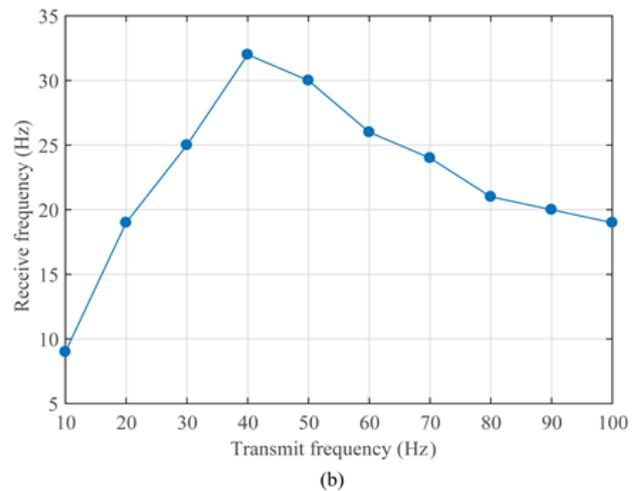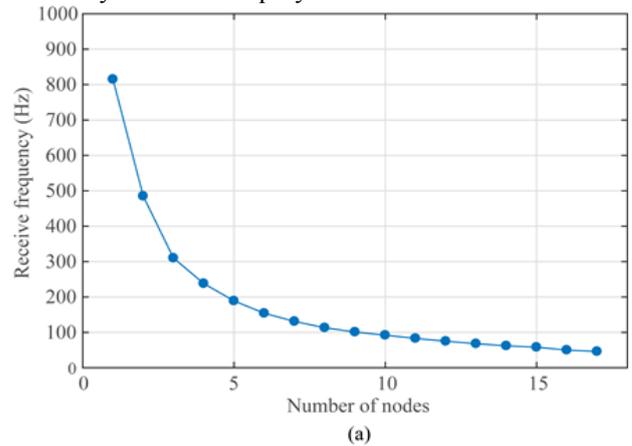

(a)

(b)

Figure 8. a) The receive frequency according to the nodes number; b) The receive frequency according to transmit frequency

Figure 9 shows the classification results of test data. Although the points distribution of a specific gesture is relatively decentralized, the SVM can distinguish these points with 100% accuracy. But we can find some points (for example, left foot) are close to the separated line, so, the model can be optimized by training more data further. For the reason that tracking nodes can be deployed arbitrarily, the average setup time is reduced to 6 minutes,

and the average calibration time is 20.3s (see Table 3). These results verify the feasibility and reliability of the calibration method and the trained SVM model. Furthermore, a comparison of the main features of proposed motion tracking system with other commercial systems is shown in Table 4.
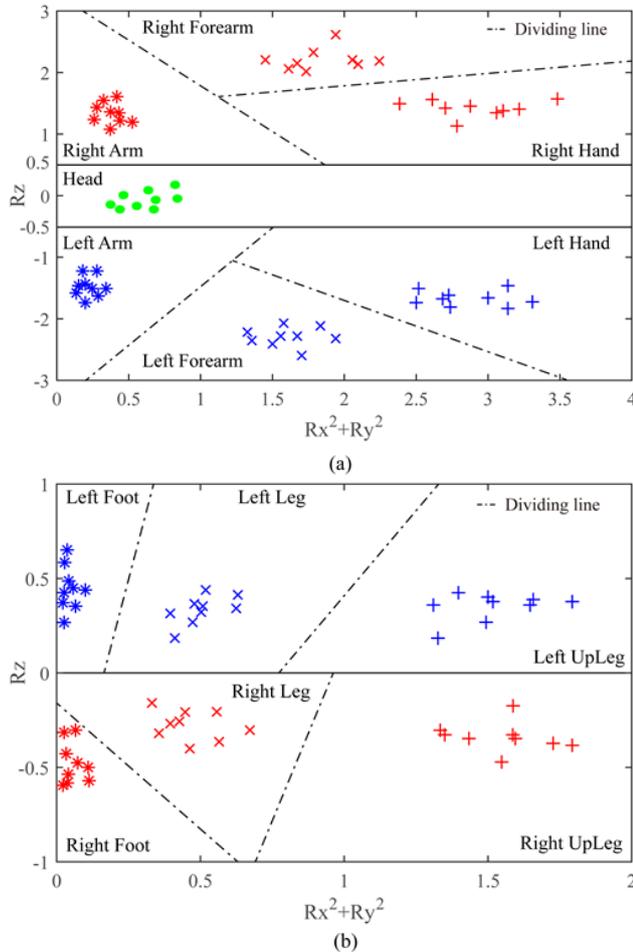
Way of fixed	Tracking Suit	Strap	Strap
Setup Time	10 minutes	8 minutes	6 minutes
Node Interchangeabilit	No	No	Yes
Sample Rate	240 (fixed)	60 (fixed)	49 (17 nodes)
Hardware Price	$15380	$12820	$2600

Figure 9. a) The classification results of the upper limbs and head using SVM model; b) The classification results of the lower limbs using SVM model.

Table 3. The time of setup and calibration

	Setup(mins)	**Cali. 1(s)**	**Cali. 2(s)**	**Cali. 3(s)**
s1	5.6	19	23	21
s2	5.8	18	19	17
s3	5.6	19	23	21

Table 4. Comparison among MVN AWINDA, MVN link and system in this paper

Parameters	**MVN Link**	**MVN Awinda**	**This Paper**
Architecture	Hub-based	Hub-free	Hub-free
Node Size	33×26×10mm	47×30×13mm	32×28×10mm
Battery life	10 hours	6 hours	7 hours
Node number	17 (fixed)	17 (fixed)	17 (adjustable)

6 Conclusion

In this paper, we have presented a wireless motion tracking system based on a hub-free architecture. An integrated IMU, a Wi-Fi module and a power supply made up a most simplified tracking node which features a minimize size and significant low power consumption. An asynchronous query strategy was proposed to monitor the network connection and optimize the data transfer rate between PC station and tracking nodes. Besides, given the particularity of calibration procedure, a transformation quaternion of one tracking node can be mapped to a single two-dimensional point. Base on this, we designed a specific calibration procedure and trained an SVM model using pre-collected data to classify the tracking nodes.

With all these techniques, this motion tracking system is extremely flexible. Due to the miniature size and weight, the tracking nodes can be easily mounted on the human body which reduces the resistance and improves the accuracy. The number of nodes used in one tracking task can be customized or adjusted dynamically according to the real applications, and the data transfer rate will always remain optimized. The experiments demonstrated that the average transfer rate for 17 tracking nodes is 49 Hz, which is high enough for some demanding applications. Moreover, compared to the commercial systems, the setup time was reduced to 6mins by a factor of the arbitrary arrangement of nodes. After calibration procedure, the trained SVM model can map the body segments and the nodes mounting on with 100% accuracy. Last but not least, considering individual node in this system is identical and burned same firmware, the cost of mass production and maintenance will be reduced significantly.

Acknowledgement

References

1. O T. Moeslund, A. Hilton and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," Comput. Vis. Image Und., vol. **104**, no. 2-3, pp. 90-126 (2006)
2. S. Won, F. Golnaraghi and W. Melek, "A Fastening Tool Tracking System Using an IMU and a Position

Sensor With Kalman Filters and a Fuzzy Expert System," IEEE Trans. Ind. Electron., vol. **56**, no. 5, pp. 1782-1792 (2009)

3. D. Roetenberg, H. Luinge and P. Slycke, "Xsens MVN: full 6DOF human motion tracking using miniature inertial sensors," Xsens Motion Technologies BV, Tech. Rep. (2009)

4. S. Komada, Y. Hashimoto, N. Okuyama, T. Hisada and J. Hirai, "Development of a Biofeedback Therapeutic-Exercise-Supporting Manipulator," IEEE Trans. Ind. Electron.vol. **56**, no. 10, pp. 3914-3920 (2009)

5. R. Zhu and Z. Zhou, "A Real-Time Articulated Human Motion Tracking Using Tri-Axis Inertial/Magnetic Sensors Package," IEEE Trans. Neur. Sys. Reh., vol. **12**, no. 2, pp. 295-302 (2004)

6. Y. Nakamura, K. Yamane, Y. Fujita and I. Suzuki, "Somatosensory computation for man-machine interface from motion-capture data and musculoskeletal human model," IEEE Trans. Robot., vol. **21**, no. 1, pp. 58-66 (2005)

7. A. Sabatini, "Estimating Three-Dimensional Orientation of Human Body Parts by Inertial/Magnetic Sensing," Sensors, vol. **11**, no. 12, pp. 1489-1525 (2011)

8. J. K. Lee and E. Park, "A Fast Quaternion-Based Orientation Optimizer via Virtual Rotation for Human Motion Tracking," IEEE Trans. Biomed. Eng., vol. **56**, no. 5, pp. 1574-1582 (2009)

9. A. Kim and M. Golnaraghi, "A quaternion-based orientation estimation algorithm using an inertial measurement unit," in PLANS., pp. 268-272 (2004)

10. SP. Tseng, WL. Li, CY. Sheng, JW. Hsu and CS. Chen. "Motion and attitude estimation using inertial measurements with complementary filter," in ASCC., Kaohsiung, Taiwan, pp. 863-868 (2011)

11. AM. Sabatini, "Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing," IEEE Trans. Biomed. Eng., vol. **53**, no. 7, pp. 1346-1356 (2006)

12. X. Yun and ER. Bachmann, "Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking," IEEE Trans. Robot., vol. **22**, no. 6, pp. 1216-1227 (2006)

13. H. Harms, O. Amft, R. Winkler, J. Schumm, M. Kusserow and G. Tröster, "Ethos: Miniature orientation sensor for wearable human motion analysis." Sensors, pp. 1037-1042 (2010)

14. C. Brigante, N. Abbate, A. Basile, A. Faulisi and S. Sessa, "Towards Miniaturization of a MEMS-Based Wearable Motion Capture System," IEEE Trans. Ind. Electron., vol. **58**, no. 8, pp. 3234-3241 (2011)

15. N. Abbate, A. Basile, C. Brigante and A. Faulisi, "Development of a MEMS based wearable motion capture system," in HSI, Catania, pp. 255-259 (2009)

16. C. Cifuentes, A. Braidot, L. Rodríguez, M. Frisoli, A. Santiago and A. Frizera, "Development of a wearable ZigBee sensor system for upper limb rehabilitation robotics," in BioRob, Rome, pp. 1989-1994 (2012)

17. P. Kinney, "Zigbee technology: Wireless control that simply works," Communications design conference. vol. **2**, pp. 1-7 (2003)

18. J. Lee, Y. Su and C. Shen, "A Comparative Study of Wireless Protocols: Bluetooth, UWB, ZigBee, and Wi-Fi," in IECON, pp. 46-51 (2007)

19. J. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," Neural Processing Letters, vol. **9**, no. 3, pp. 293-300 (1999)

20. VN. Vapnik, and V. Vlamimir, Statistical learning theory, New York: Wiley, pp123-129.0 (1998)