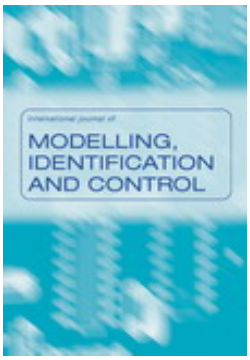


International Journal of Modelling, Identification and Control

ISSN online: 1746-6180 - ISSN print: 1746-6172
https://www.inderscience.com/ijmic

HMM-based IMU data processing for arm gesture classification and motion tracking

Danping Wang, Jina Wang, Yang Liu, Xianming Meng

DOI: 10.1504/IJMIC.2023.10053831

Article History:

Received:	25 November 2021
Last revised:	17 January 2022
Accepted:	04 February 2022
Published online:	03 February 2023

HMM-based IMU data processing for arm gesture classification and motion tracking

Danping Wang*

School of Information Engineering,
Shenyang University,
Shenyang, 110044, China
Email: 2026678321@qq.com
*Corresponding author

Jina Wang

Liaoning Province's Construction and Engineering Center for
Advanced Equipment Manufacturing Base,
Shenyang, 110001, China
and
Liaoning Information Security and Software Testing and Certification Center,
Shenyang, 110001, China
Email: wangjina12@163.com

Yang Liu

School of Information Engineering,
Shenyang University,
Shenyang, 110044, China
Email: liuyang@pku.edu.cn

Xianming Meng

Wuxi Forward Technology Co., Ltd,
Wuxi, 214191, China
Email: mengxm@126.com

Abstract: This paper investigates inertial measurement unit (IMU) data processing methods for human gesture classification and arm motion tracking in wireless body sensor network (WBSN). The method is adopted that consists of two main stages. In the training stage, the supervised learning method is adopted to obtain the HMM model and the Viterbi algorithm is used to obtain the optimal hidden state sequence in the testing stage. HMM also complements the intuitional evaluation for arm motion recovery. We take advantage of the twists and exponential maps to recover the arm motion process. In addition, visual tracking device-VICON is utilised to validate the accuracy of the inertial tracking system. The experimental results show that the HMM algorithm gesture classifier achieves up to 96.63% accuracy on five commonly used arm gestures and visual assisted tracking outcomes verify the robustness and feasibility of the IMU tracking device.

Keywords: gesture classification; arm motion tracking; inertial measurement unit; computer network; vision motion tracking.

Reference to this paper should be made as follows: Wang, D., Wang, J., Liu, Y. and Meng, X. (2023) 'HMM-based IMU data processing for arm gesture classification and motion tracking', *Int. J. Modelling, Identification and Control*, Vol. 42, No. 1, pp.54–63.

Biographical notes: Danping Wang received her PhD from the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China, in 2019. Currently, she is a Teacher at Shenyang University, Shenyang, China. Her current research interests include swarm intelligence and their applications in design and optimisation of intelligent transportation systems.

Jina Wang received her MS degree in Software Engineering from Shenyang University of Technology, in 2011. Currently, she is a Senior Engineer of Liaoning Province's Construction and Engineering Center for Advanced Equipment Manufacturing Base (Liaoning Information Security and Software Testing and Certification Center). Her research interests include software testing, software engineering cost evaluation, etc.

Yang Liu received her PhD from the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China,in 2017. Currently, she is a Teacher at Shenyang University, Shenyang, China. Her current research interests include artificial intelligence, evolutionary computation, swarm intelligence, and their applications in design and optimisation of intelligent transportation systems, and RFID systems.

Xianming Meng received his BS degree in Material Forming and Control Engineering from Shenyang University of Technology, Shenyang, China in 2003 and MS degree in Computer Software and Theory from Shenyang University of Technology, Shenyang, China in 2006. Currently, he is an Associate Researcher at the Department of Information, Wuxi Forward Technology Co., Ltd. His research interests include issues related to RFID, industrial IoT, big data.

1 Introduction

Gesture classification is an important application in pattern recognition and pertains to recognising the meaningful expressions of motion by a human, involving the hands, arm, face, head, and body (Sushmita and Tinku, 2007). Human motion tracking or Mocap is currently a booming research field and being applied to a variety of areas such as biomedical gait analysis, lost motor function recovery, effective athlete training, virtual reality, human-computer interface, to name only a few. As the proliferation of MEMS technology has inspired the sensor-based research, quite a lot of literatures on gesture recognition and motion tracking have been published.

In gesture classification, most of the previous literatures have been focused on computer vision techniques. But the vision-based recognition performance is largely degraded when the lighting condition is bad or the fixed camera's angle is not appropriate. By contrast, accelerometer-based gesture techniques are widely used in mobile or wearable devices like Apple iPhone or Wiimote. Usually, the typical gestures are hand and arm gestures (Markos et al., 2010; Geng et al., 2019; Mace et al., 2013; Yin et al., 2014; Xu et al., 2012; Zhang et al., 2011; Liu et al., 2009; Wu et al., 2009; Zhu et al., 2019), head and face gestures (Morency and Darrell, 2006) and body gestures (Cutler and Turk, 1998). Among these gesture categories, arm gestures are relatively more meaningful and its degree of freedom is larger than the other limbs (e.g., the lower limbs). In addition, the number of the accelerometer-integrated IMUs mounted on the arm is more than that attached or held by the hand, which provide more accurate and reliable gesture information. A few researches have been conducted on arm gesture recognition. The feature is selected using the discrete cosine transformation (DCT) and principal component analysis (PCA) to reduce the number of features and increase the probability of correct classification results. Yet this method is applicable as far as three gestures are concerned. For more arm gestures, PCA or DCT is not appropriate for feature extraction as the optimal number of the features is not well determined. Another gesture recognition method uses dynamic time warping and affinity propagation that creates the exemplars in the testing stage. The candidate traces are projected to the lower dimensional subspace and compared with the features stored in the established gesture dictionary. In this paper, we mount IMUs on the right arm joints and the readings from the IMU accelerometers are all taken into account (compared and selected) in arm gesture classification.

Motion tracking is categorised as visual tracking and non-visual tracking (Hadjidj et al., 2013). Visual motion tracking system can localise the human's body or limbs by means of combining the data such as the coordinates from a certain number of cameras that are installed at different perspectives in a lab. The popular marker-based system such as VICON or QUALYSIS performs. However, visual tracking fails to accommodate the human motions in free living environments, though its capture accuracy is within 1.5 mm. Besides, high-resolution in-studio commercial capture system requires the external sources like calibrated cameras and specified workstation for the computational load is dramatically heavy, which is not appropriate for tracking the people in daily life, especially for the patients in rehabilitation. The mechanical assistance of robotic rehabilitators in improving the arm movement ability justifies its tracking's stability and effectiveness (Salah et al., 2020). This kind of tracking device fixed on the body measures the acceleration, velocity and force, which facilitates those patients with severe disabilities and unable to perform the normal daily exercises in everyday life. Yet the mechanical device is quite cumbersome and inhibits the personal movements. Sensors have been used in motion tracking to avoid the problems inherently existing in visual tracking, e.g., line-of-sight problems and self-occlusion. Ultrasound tracking, magnetic tracking and inertial tracking are all classified as sensor-based tracking methods. Ultrasound tracking has characteristics of low cost and compactness, but its intrinsic vulnerability to occlusion and inaccuracy limits its wide applications. Similarly, the ferromagnetic materials severely affect the measurement on earth magnetic direction and therefore limit the magnetic tracking in daily life. By contrast, compactable and low-cost micro electromechanical system (MEMS) is relatively competent for human motion tracking. The advantages of these wearable tracking devices are:

1 source free and wearable

2 low power consumption and long battery span

3 free from the self-occlusion and unobtrusiveness, which inherently exist in vision tracking

4 low cost and long-term tracking.

Although its drawbacks on drift problem caused by noise and offsets are considerable, the corresponding methods are proposed, such as the quaternion-based filter and the adaptive-gain orientation filter, which have improved gyroscope output efficiently (Tian et al., 2013; Chen et al., 2011). IMU positioning for indoor pedestrian, which has proposed the ZUPT algorithm to correct the walking velocity that effectively reduces the accumulated drift errors in accelerometers (Hadjidj et al., 2013).

In this paper, we firstly propose an HMM-based arm gesture classification method based on three-axis accelerometer data. In HMM modelling, we adopt the supervised learning instead of the widely used Baum-Welch method, which is largely susceptible to the initial points and apt to plunge into local minimum. HMM also complements the intuitional evaluation for arm motion recovery for fear that the recovery result is not convincing. Secondly, a quaternion-based arm motion tracking method is utilised. To the best of our knowledge, there's no literature on human motion tracking methodology taking gesture recognition into consideration. Personally speaking, motion tracking requires the data collected by IMU being imported to the corresponding computer, which is impossible for the user without any PC. But the gesture recognition sequences can be calculated by the micro-controller unit, stored and displayed on the IMU board that potentially facilitates the user to refer to their personal motion data in any specified time.

Main contributions of this paper are as follows:

1 HMM applied in gesture classification using supervised learning method in modelling and to evaluate the arm motion tracking outcome with the HMM optimal hidden sequences to achieve the multi-functional goals for the users, especially for the patients in rehabilitation. In supervised learning, the training data has both features and labels. Through training, the machine can find the connection between features and labels on its own, so that it can determine the labels when faced with data that has only features but no labels.

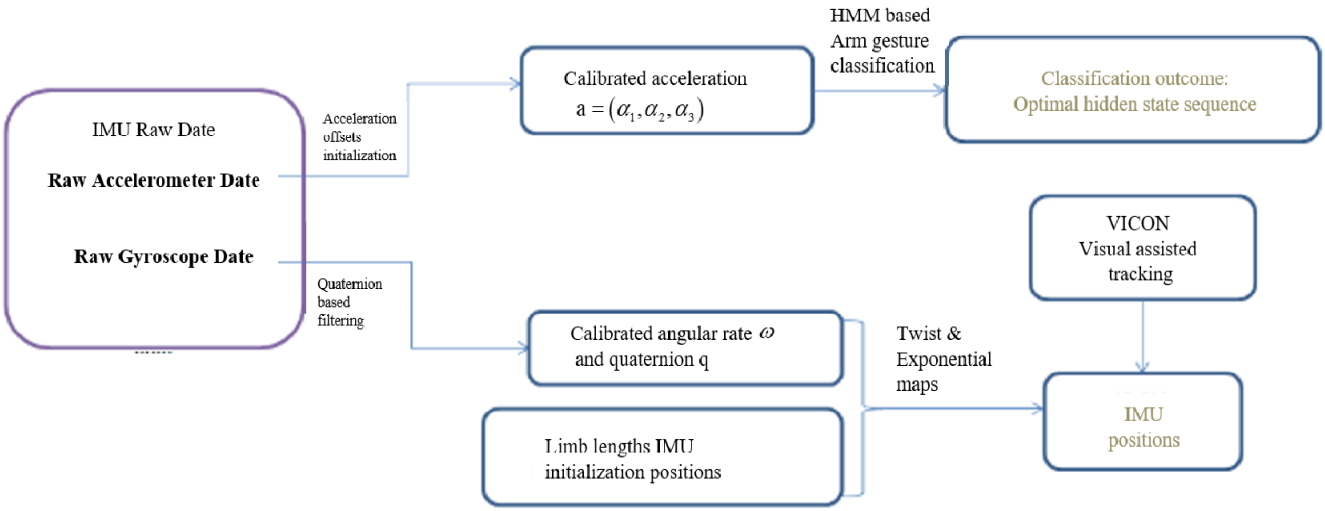
2 Unlike the traditional wearable system that uses both accelerations and angular rates to estimate the human joints positions and orientations in which the double integration of accelerations would largely exaggerate the sensor noises on position estimation, the proposed method uses merely the angular rate and predetermined arm lengths to fulfil the arm motion tracking.

3 The 'golden standard' visual tracking device – VICON is utilised to assist and verify the IMU tracking results by directly comparing the joint positions rather than the joint angles.

The organisation of this paper is as follows: Section 2 formulates the problem and introduces a general overview of the proposed gesture recognition and motion tracking system. Section 3 describes the Hidden Markov Model (HMM) algorithm to classify the arm motions. Section 4 describes the motion reconstruction with twists and exponential maps. The accurate visual-assisted tracking method is also introduced to validate the IMU tracking system. Section 5 shows the HMM-based motion classification outcome and tracking result. Section 6 concludes the whole paper.

The main framework of the recognition and tracking methods is shown in Figure 1.

Figure 1 The main framework of the proposed arm gesture classification and motion tracking



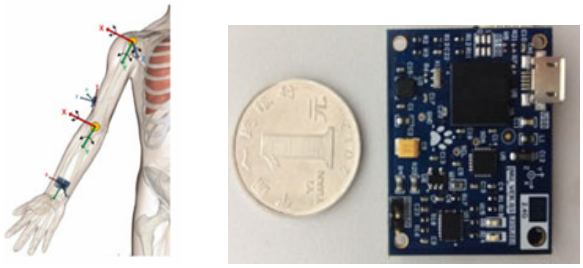
Firstly, the acceleration offset of the raw acceleration data is initialised, and the raw gyroscope data is filtered based on quaternions. Secondly, the resulting acceleration is calibrated，and optimal hidden state sequence can be obtained by HMM-based arm gesture classification. Then, twist and exponential maps can be used by calibrated angular rate ω, quaternion q, limb lengths and IMU initialisation positions to recover the arm motion，Finally, the feasibility of the IMU tracking system can be verified by VICON visual assisted tracking.

2 Problem formulation and system overview

2.1 Arm kinematic modelling

Human arm is modelled as an articulated multi-link chain consisting of three rigid body parts, three joints and 7 DoFs as is shown in Figure 2(a). Shoulder is described as a ball that has 3 DoFs. Elbow is modelled as two hinge joints with 2 DoFs. Wrist is modelled as an ellipsoid joint with 2 DoFs.

Figure 2 IMU placements and board, (a) IMU placements (b) IMU board (see online version for colours)



(a) (b)

Our proposed gesture classification and motion tracking system employs IMUs mounted on the arm joints, (i.e., wrist, elbow, shoulder) shown in Figure 2(a). IMU is intended to collect the instant acceleration α, angular velocity ω, magnetic field intensity m and quaternion q.

Before conducting the IMU tracking, we need to measure the radius of rotation around each degree of freedom and to load the accelerometer offsets in the IMU initialisation step. The gyroscope on the IMU is capable of capturing the link circular motion around each axis. Thus, the initial global position of the relative joint is determined on condition that the radius is given. The shoulder joint and elbow joint are taken as an example, the elbow angular rate relative to the shoulder joint can be recorded by the IMU mounted near the elbow. Once the radius of the upper arm length in three axes X-Y-Z is obtained, elbow joint position will be determined. Similarly, repetitive procedure is suitable for the wrist joint position relative to the elbow joint.

2.2 System overview

The hardware, as is shown in Figure 2(b) includes a module MPU6500 integrated with a MEMS triaxial accelerometer, triaxial gyroscope and triaxial magnetometer. An embedded processor MSP4300 and a Bluetooth module are also designed on this IMU board. The board size is 35 mm × 30 mm × 8 mm. The sensor signals are sampled at 40 Hz and interfaced to the computer by bluetooth.

3 HMM in gesture recognition

Markov process is a general stochastic process characteristic of the well-known statistical Markov model. The ubiquitous case is the first-order discrete Markov chain. The probability of the current state is only related to the predecessor state, i.e., $\Pr(q_{t+1} = S_i|q_i = S_j, q_{t-1} = S_k, \ldots, q = S_m) = \Pr(q_{t+1} = S_i|q_t = S_j) = a_{ji}$, where q_t is the state at the time instant t that has N states: $S_1, S_2, \ldots, S_N$. This conditional probability is independent of time t, i.e., $\Pr(q_{t+1} = S_i|q_t = S_j) = \Pr(q_{t+\Delta t+1} = S_i|q_{t+\Delta t} = S_j)$. The parameter a_{ij} is an element of the state transition probability matrix A, which has the property that a_{ij} is no smaller than zero and $\sum_{j=1}^{N} a_{ij} = 1$. for they obey standard stochastic constraints and each element a_{ij} corresponds to a specific physical event. As for the discussed problems of interest, it can be used to represent the probability of the coupled arm gestures.

HMM is a doubly stochastic process, in which the observable state sequence is generated from a hidden Markov sequence (Nguyen et al., 2010). It is depicted in Figure 3.

The outputs of the triaxial accelerometers are three feature curves parts of which are shown in Figure 6. The red one represents the X-axis accelerometer output. Similarly, the green one and the blue one represent the Y-axis and Z-axis accelerations respectively. Traditional methods in selecting the features of acceleration curves are FFT with an appropriate window size. Usually, the window size is determined by the length of the motion time which largely differs from person to person. Even the same person performs the identical motion with the movement time space discrepant. Thus, the manual segmentation approach is applied to improve the gesture recognition accuracy. The features of the commonly used five motions are determined intuitionally by the listed curves. The flow chart for HMM-based motion classification is shown in Figure 4.

Figure 3 Hidden state sequence and observation sequence in HMM (see online version for colours)

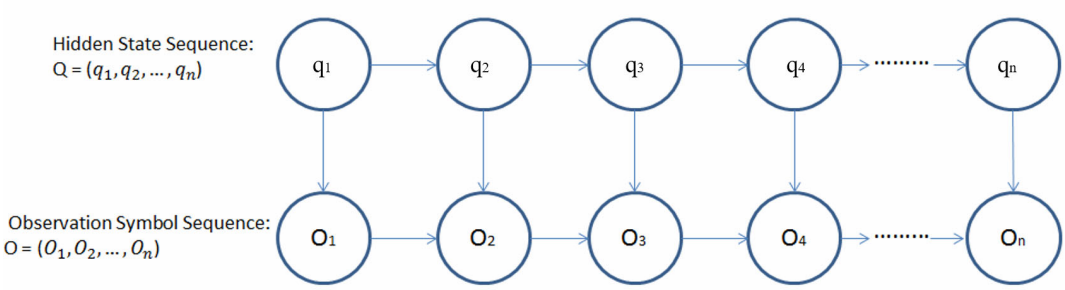
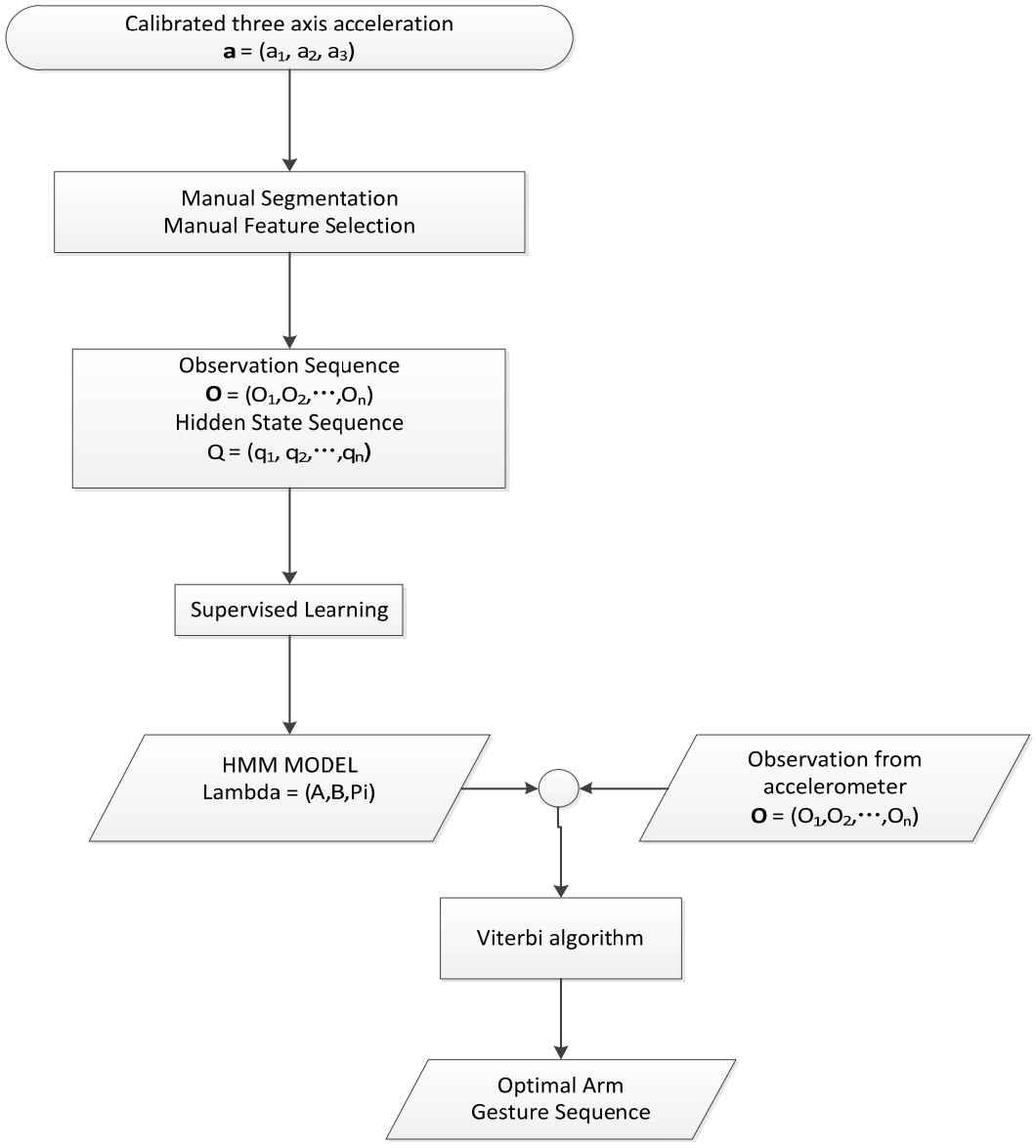


Figure 4 HMM-based arm gesture classification flow chart



3.1 HMM modelling

The critical issue in HMM is modelling, to be specific, learning. HMM learning method is dichotomised into supervised learning and unsupervised learning. Baum-Welch algorithm is a pervasively used unsupervised learning method but it suffers from the local maxima and is vulnerable to the initial values. In contrast, the supervised learning method is independent of the initial estimation as well as the local maximum though it requires the repetitive work in labelling the observation symbols. Thus, the supervised learning is adopted to determine the model parameter λ at the cost of pre-labelling work. The step-in determining λ is as follows:

1 State transition probability matrix A = $[a_{ij}]_{N\times N}$ determination

$$a_{ij} = \frac{A_{ij}}{\sum_{j=1}^{N} A_{ij}} \tag{1}$$

where A_{ij} is the number of times that the hidden state transits from state S_i to S_j.

2 Observation probability matrix $B = [b_{ij}]_{N\times N}$ determination

$$b_{ij} = \frac{B_{jk}}{\sum_{k=1}^{M} B_{jk}} \tag{2}$$

where B_{jk} is the number of times that the observation symbol is k and the hidden state is S_j.

3 Initial probability distribution $\pi = [\pi_i]_{N\times 1}$, where

$$\pi_i = \Pr(q_0 = S_i) \tag{3}$$

3.2 Optimal hidden state sequence

The second critical issue in HMM, however, is decoding, i.e., given the model parameter λ and the output sequence O, how could we deduce the 'optimal' hidden state sequence that best fits the model as well as the observation sequence. Compared with the approximation algorithm, the case where the transition state equals to zero isn't taken into account, will result in the impossible hidden state sequence. Henceforth, the Viterbi algorithm in our gesture recognition is applied to find the best hidden state sequence.

4 Arm motion tracking

Liu et al. (2009) pointed out that a rigid body can be moved from any one place to any other by a movement, which consists of rotation and transition. Rigid motion has the elegant property that preserves the distance between two points and the angle between the vectors. In this paper, arm is modelled as a kinematic chain and is characteristic of 7 DoFs. Shoulder is described as a socket joint with 3 DoFs. Elbow is modelled as two hinge joints with non-intersecting axes. Wrist is modelled as an ellipsoid joint with 2 DoFs.

4.1 IMU configuration

In our arm motion system, three coordinate frames are considered. These are shoulder joint coordinate frames, elbow joint coordinate frame and wrist joint coordinate frame. The shoulder joint is assigned to be the base frame linking a succession of upper arm and forearm movements. The instantaneous position and orientation of the wrist and elbow joint, relative to the shoulder inertial frame, can be represented by the pair $\{P_{ab}, R_{ab}\}$, where P_{ab} is the vector from the origin of shoulder coordinate frame to wrist/elbow frame, R_{ab} is the rotation of wrist/elbow frame relative to the shoulder frame.

Three IMUs are mounted on the right arm near the joints based on the consideration that IMU will affect the normal movements of the arm if being placed right on the joint centre. Besides, we must focus on how to lessen the effect on soft tissue artefact and try to improve the relative accuracy of the deduced joint position. Therefore, one of the IMUs is positioned over the distal humerus to capture the upper arm movement, another is attached to the distal flat surface of radius and ulna, adjacent to the joint elbow. The third is mounted on the right of shoulder as it slightly affects the shoulder joint movements.

4.2 Twists and exponential maps

As we have mentioned above, the pair $\{P_{ab}, R_{ab}\}$ describes the relationship between the inertial frame B and inertial frame A. More precisely, let q_a and q_b be the coordinates of point Q in the frame A and B respectively. Then q_a and q_b satisfy $q_b = P_{ab} + R_{ab}q_a$.

The homogeneous representation of the above equation is:

$$\begin{bmatrix} q_a \\ 1 \end{bmatrix} = \begin{bmatrix} R_{ab} & P_{ab} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} q_b \\ 1 \end{bmatrix} = \overline{g}_{ab}\overline{q}_b, \quad \overline{g}_{ab} \in se(3) \tag{4}$$

where $P_{ab} \in \mathbb{R}^3$ is the vector that directs the origin of frame B relatively to the origin of frame A; $R_{ab} \in \mathbb{R}^{3\times 3}$ is an orthogonal matrix that represents the rotation of frame B relative to frame A.

The commonly used frame rotation expression is quaternion. Unlike the Euler angle, quaternion doesn't suffer from singularities. Besides, quaternion is computationally efficient than the conventional rotational matrix representation.

Quaternion generalises a four-dimensional vector $q = q_0 + q_{1i} + q_{2j} + q_{3k} = (q_0, \vec{q})$, where q_0 is the scalar part, q_1, q_2, q_3, are the weighted values on the basis elements denoted as $\vec{I}, \vec{j}, \vec{k}$. In the coordinate rotation representation,

$$q = \left(\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\cos\omega_x, \sin\frac{\theta}{2}\cos\omega_y, \sin\frac{\theta}{2}\cos\omega_z\right),$$

(cos ω_x, cos ω_y, cos ω_z) represent the quaternion axis; θ symbolises the angle around this axis.

As for every homogeneous matrix $\overline{g}_{ab}$, there always exists a corresponding twist in the tangent space $se(3)$, and we define

$$se(3) = \left\{(v, \hat{w}) : v \in \mathbb{R}^3, \hat{w} \in so(3)\right\} \tag{5}$$

$$so(3) = \left\{S \in R^{3\times 3} : S^T = -S\right\} \tag{6}$$

Let $\hat{\xi} = \begin{bmatrix} v \\ \omega \end{bmatrix}^{\wedge} = \begin{bmatrix} \hat{w} & v \\ \mathbf{0} & 0 \end{bmatrix} \in se(3),$, where $v = -\omega \times q_g$ represents the velocity of the point and $\hat{w}$ is a skew-symmetric matrix. Define $\omega = [\omega_1, \omega_2, \omega_3]$, then we get

$$\hat{w} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \tag{7}$$

The exponential product of $\hat{\xi}\theta$ is an element that belongs to $se(3)$.

$$e^{\hat{\xi}\theta} = \begin{cases} \begin{bmatrix} e^{\hat{w}\theta} & (I - e^{\hat{w}\theta})(\hat{w}v + \omega\omega^T v\theta) \\ 0 & 1 \end{bmatrix} & \omega \neq 0 \\ e^{\hat{\xi}\theta} = \begin{bmatrix} I & v\theta \\ 0 & 1 \end{bmatrix} & \omega = 0 \end{cases} \tag{8}$$

If we define $g_{global,n}(0)$ as the initial configuration of a rigid body in reference n relative to the global frame, then the final configuration is given by $g_{global,n}(\theta) = e^{\hat{\xi}\theta} g_{global,n}(0)$, where $e^{\hat{\xi}\theta}$ can be calculated by the Rodrigues' formula:

$$e^{\hat{w}\theta} = I + \hat{w}\sin\theta + \hat{w}^2(1 - \cos\theta) \tag{9}$$

4.3 Coordinate transformation

In order to verify the IMU motion tracking accuracy, we take advantage of the 'golden standard' visual tracking device: the arm motion capture can be assisted by VICON as this system tracking error is less than 1.5 mm. The experiment is conducted by keeping IMU tracking records as well as the VICON tracking records frame by frame at the same time. The problem of interests is to transform the joint position in VICON coordinate frame to the corresponding position in IMU coordinate frame. For this reason, the coordinate transformation is conducted, in which the scale factor K_0, the rotation matrix R and coordinate translation parameters (X_0, Y_0, Z_0) need to be determined.

Three points P_1, P_2, P_3 are denoted as the shoulder joint, elbow joint and wrist joint respectively. The IMU coordinate frame is defined as O-XYZ. VICON coordinate frame is defined as O'-$X'Y'Z'$. (x_i, y_i, z_i) i = 1, 2, 3 represents coordinate of P_i in O-XYZ frame. In a similar way, (x_i', y_i', z_i') represents coordinate of P_i in O'-$X'Y'Z'$.

The difference of coordinate length basis gives rise to the parameter k_0. Here we define

$$S_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \tag{10}$$

$$S_{ij}' = \sqrt{(x_i' - x_j')^2 + (y_i' - y_j')^2 + (z_i' - z_j')^2} \tag{11}$$

where S_{ij} is the distance of P_i and P_j in O-XYZ frame and S_{ij}' is the distance of P_i and P_j in O'-$X'Y'Z'$ frame. Based on the above definition, we could determine the scale factor k_0 as follows:

$$k_0 \frac{s_{12} + s_{13} + s_{23}}{s_{12}' + s_{13}' + s_{23}'} \tag{12}$$

4.3.1 Transformation matrix R

A new coordinate frame P_1-XYZ is defined in which P_1 is the origin and vector $\overrightarrow{P_1P_2}$ lies on the X axis. After normalisation, we obtain $\overrightarrow{P_1P_4} = \frac{\overrightarrow{P_1P_2}}{\|\overrightarrow{P_1P_2}\|} = (a_1, b_1, c_1)$. As we can see from Figure 5, the product of $\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3}$ lies on the Z axis and after normalisation, $\overrightarrow{P_1P_6} = (a_2, b_2, c_2)$ is derived. Similarly, the cross product of $\overrightarrow{P_1P_6}$ and $\overrightarrow{P_1P_4}$ renders the normalised vector $\overrightarrow{P_1P_5} = (a_3, b_3, c_3)$, since $\overrightarrow{P_1P_6}$ and $\overrightarrow{P_1P_4}$ are all normalised, the cross product is still normalised. Conspicuously, $\overrightarrow{P_1P_5}$ lies on the Y axis in P_1-XYZ coordinate frame. In P_1-XYZ coordinate frame, $\overrightarrow{P_1P_4}$, $\overrightarrow{P_1P_5}$, $\overrightarrow{P_1P_6}$ are quantity of the form: $\vec{i}$, $\vec{j}$, $\vec{k}$. While in O-XYZ coordinate frame, $\overrightarrow{P_1P_4}$, $\overrightarrow{P_1P_5}$, $\overrightarrow{P_1P_6}$ are directional cosines: (a_1, b_1, c_1), (a_3, b_3, c_3), (a_2, b_2, c_2). Hence, we obtain the rotation matrix P_{op1} between frame P_1-XYZ and O-XYZ.

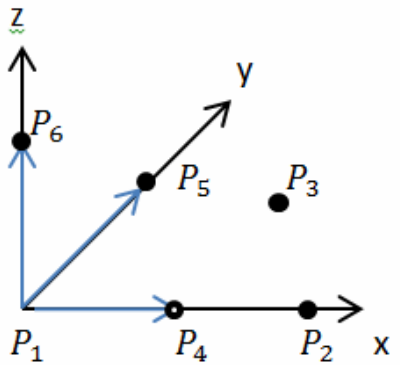
$$P_{op1} = \begin{bmatrix} a_1 & a_3 & a_2 \\ b_1 & b_3 & b_2 \\ c_1 & c_3 & c_2 \end{bmatrix} \tag{13}$$

By means of this transformation stated above, we can also obtain the transformation matrix $R_{o'p1}$ between P_1-$X'Y'Z'$ and O'-$X'Y'Z'$.

Finally, the transformation matrix between the frame O-XYZ and frame O'-$X'Y'Z'$ is

$$R = P_{op1} \times R_{o'p1}^{-1} \tag{14}$$

Figure 5 P_1-XYZ frame (see online version for colours)



4.3.2 Translation candidates (X_0, Y_0, Z_0)

Given the above derived scale factor k_0 and matrix R, we can easily deduce the translation parameters. The point P is arbitrarily chosen of which (x_p, y_p, z_p) and (x_p', y_p', z_p') are the coordinates in O-XYZ and O'-$X'Y'Z'$ frames respectively. Then we get

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} x_p \\ y_p \\ z_p \end{pmatrix} - k_0 R \begin{pmatrix} x'_p \\ y'_p \\ z'_p \end{pmatrix} \quad (15)$$

5 Experiment result and analysis

5.1 HMM classifier for arm motion

In our experiment, five commonly used symbolic gestures are designed:

S1 arm extending horizontally

S2 arm waving up and down

S3 arm waving left and right

S4 forearm bending over

S5 arm rotating inside and out.

There are seven observation symbols corresponding to these states, we listed as O_1 to O_7. These observation symbols are the selected feature curves of three-axis accelerometer output.

Seven experiments are conducted with the sample size ranging from 40 to 160 respectively. The accuracy of the tests is shown in Table 1.

The overall accuracy reaches up to 92%. As we can see from Table 1, the accuracy increases while the sample size enlarges and this outcome shows that the more the sample size is, the more accurate and reliable the hidden state sequence will be. The best accuracy the proposed system yields are 96.63% that outperforms the discrete HMMs, which is 89.9%. And our system's performance reaches the same level compared with the uWave.

5.2 Arm motion tracking

The proposed system recovers the arm motion process by using the twist and exponential maps. The parameters determined before conducting the experiment are the lengths of forearm and lower arm. Besides, the initial positions of IMU mounted adjacent to the wrist, elbow and shoulder are also measured. The unique data to be processed is quaternion. Here we use MATAB to process the IMU data and use VICON studio to process the VICON data in recovering the arm motion process.

Figure 6 Observation symbols in HMM (see online version for colours)

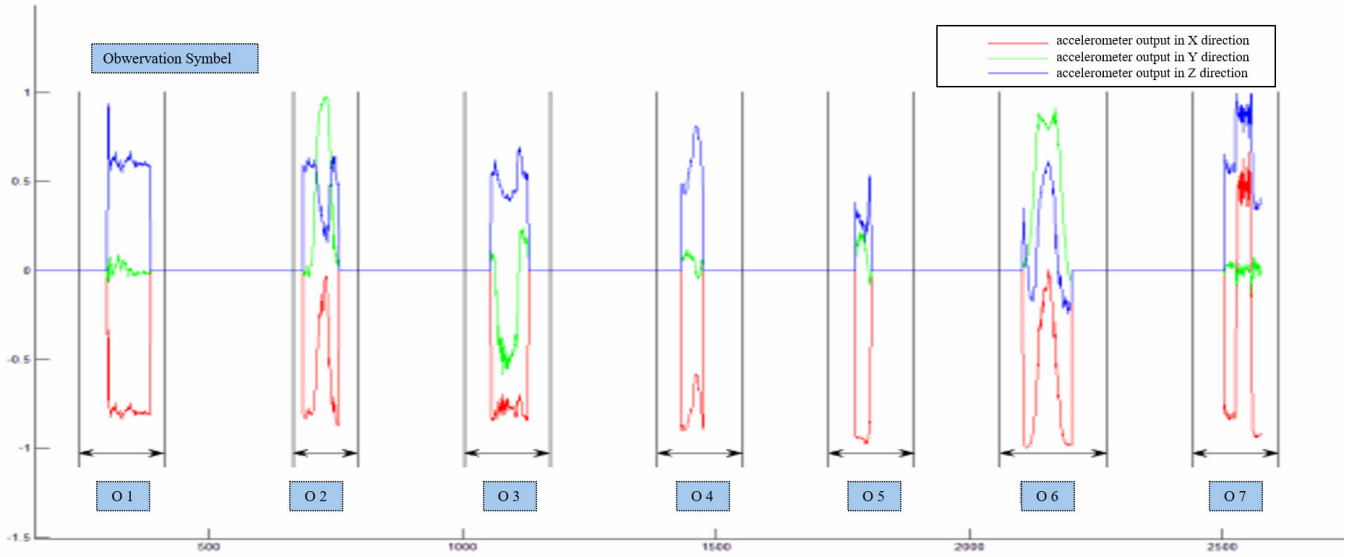


Table 1 HMM classifier accuracy

Sample size	*40*	*60*	*80*	*100*	*120*	*140*	*160*
Accuracy	85%	88%	92.5%	95.24%	93.41%	94.29%	94.29%

Table 2 Comparison of the classification performance

Classification method	*Number of gestures*	*Accuracy %*
Proposed method	5	96.63%
Discrete HMMs in Akl et al. (2011)	5	89.7%
uWave	8	98.6%

Figure 7 Triaxial wrist joint error (see online version for colours)

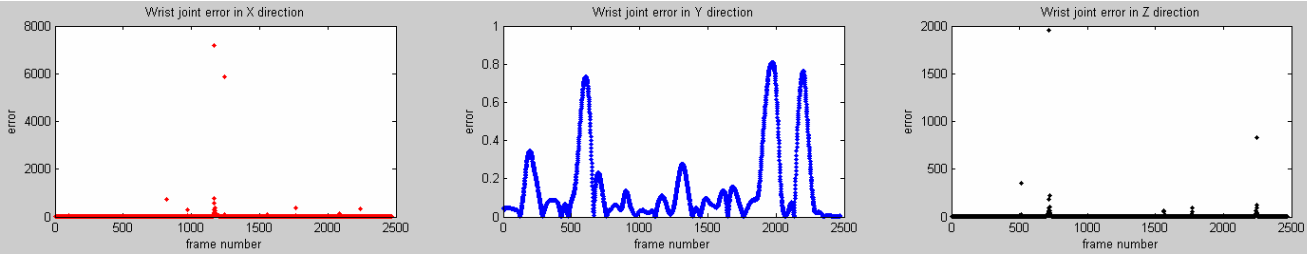


Figure 8 Triaxial elbow joint error (see online version for colours)

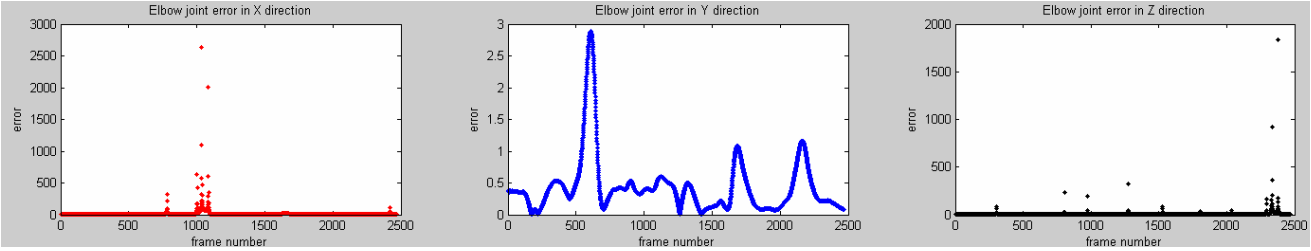
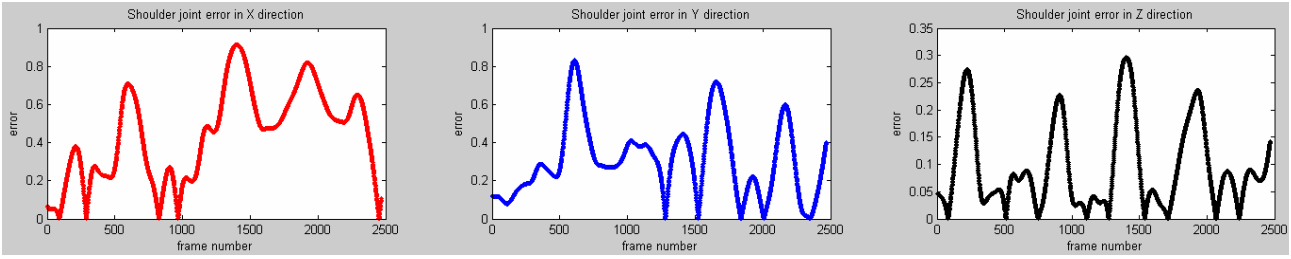


Figure 9 Triaxial shoulder joint error (see online version for colours)



5.3 Tracking accuracy test

For verifying our wearable IMU tracking system feasibility, we keep VICON visual tracking records while conducting the inertial tracking simultaneously. The problem we encountered is the frame alignment between IMU coordinate and VICON coordinate. Due to that reason, a local minimisation solution is applied.

$$\min_{i,j}\left(\sum_{k=1}^{3}\omega_k \left\|P_k(i)-P_k(j)\right\|_2\right), \sum_{i=1}^{3}\omega_i = 1 \tag{16}$$

where k = 1, 2, 3 represent the shoulder, elbow, wrist joint respectively. ω_k represents the weighted value on the 2-norm distance between the $P_k(i)$ and $P_k(j)$ of which i represents the frame number in IMU coordinate and j represents the frame number in VICON coordinate. In this experiment, $\omega_1 = 0.6$, $\omega_2 = 0.3$, $\omega_3 = 0.1$.

Figures 7, 8 and 9 depict the arm joint errors in three axes. From these three figures, we can deduce that wrist errors are pretty unpredictable in X and Z directions. Some of the deflection points are due to the sensor noise and the soft tissue artefact.

6 Conclusions and discussion

In conclusion, we've proposed an IMU data processing system for arm gesture classification and arm motion tracking. In HMM-based arm gesture classification, five commonly used arm gestures are selected and HMM is established in the supervised learning, which is good to avoid the influence of initial points and not easy to fall into local minima. Afterwards, Viterbi algorithm is introduced to deduce the optimal hidden state sequences. In addition to classification, HMM also complements the intuitional evaluation for arm motion recovery. In arm motion tracking, twists and exponential maps are applied to recover the arm motion using the quaternions, pre-measured upper-limb length and designated initial IMU positions. VICON visual tracking device is utilised to validate the feasibility of the IMU tracking system.

One drawback in the tracking system is the impracticability in firmly fixing the IMU board at the designated positions，which will give rise to the soft tissue artefact and noise from the sensor output. The future goal is to obtain an IMU working system that can work in any position. Another issue of interest is the IMU accumulated errors caused by gyroscope, which will pose tremendous effect on tracking. A potential solution to this issue will be meaningful in real-time accurate motion tracking and positioning. Our research group will dig on that challenging issue afterwards.

Acknowledgements

This work was supported by the Scientific Research Project of Liaoning Province under Grants No. 2017104005 and Liaoning Province Innovative Talent Support Program No. LR201861.

References

Akl, A., Feng, C. and Valaee, S. (2011) 'A novel accelerometer-based gesture gecognition system', *IEEE Transactions on Signal Processing*, Vol. 59, No. 12, pp.6197–6205.

Chen, X., Hu, S., Shao, Z.Z. and Tan, J. (2011) 'Pedestrian positioning with physical activity classification for indoors', *IEEE International Conference on Robotics and Automation*, pp.1311–1316.

Cutler, R. and Turk, M. (1998) 'View-based interpretation of real-time optical flow for gesture recognition', *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp.416–421.

Geng, X., Zhu, Q., Liu, T. and Na, J. (2019) 'U-model based predictive control for nonlinear processes with input delay', *Journal of Process Control*, Vol. 75, pp.156–170.

Hadjidj, A., Souil, M., Bouabdallah, A., Challal, Y. and Owen, H. (2013) 'Wireless sensor networks for rehabilitation applications: challenges and opportunities', *Journal of Network and Computer Applications*, Vol. 36, No. 1, pp.1–15.

Liu, J., Wang, Z., Zhong, L., Wickramasuriya, J. and Vasudevan, V. (2009) 'uWave: accelerometer-based personalized gesture recognition and its applications', *IEEE International Conference on Pervasive Computing and Communications*, pp.1–9.

Mace, D., Gao, W. and Coskun, A.K. (2013) 'Improving accuracy and practicality of accelerometer-based hand gesture recognition', *Proceedings of the 2nd IUI Workshop on Interacting with Smart Objects*, pp.45–49.

Markos, S., Haris, B. and Panos, T. (2010) 'Gesture recognition based on arm tracking for human-robot interaction', *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.5424–5429.

Morency, L.P. and Darrell, T. (2006) 'Head gesture recognition in intelligent interfaces: the role of context in improving recognition', *Proceedings of the 11th International Conference on Intelligent User Interfaces*, January, pp.32–38.

Nguyen, H.Q.P., Kang, H.J. and Suh, Y.S. (2010) 'Vision-inertial tracking algorithm with a known object's geometric model', *Advanced Intelligent Computing Theories and Applications, Lecture Notes in Computer Science*, Vol.6216, pp.450-456.

Salah, N., Kais, B. and Hassen, M. (2020) 'Fractal, chaos and neural networks in path generation of mobile robot', *International Journal of Modelling, Identification and Control*, Vol. 34, No. 1, pp.41–50.

Sushmita, M. and Tinku, A. (2007) 'Gesture recognition: a survey', *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 37, No. 3, pp.311–324.

Tian, Y., Wei, H. and Tan, J. (2013) 'An adaptive-gain complementary filter for real-time human motion tracking with MARG sensors in free-living environments', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 21, No. 2, pp.254–264.

Wu, J., Pan, G., Zhang, D., Qi, G. and Li, S. (2009) 'Gesture recognition with a 3-D accelerometer', *Ubiquitous Intelligence and Computing. Lecture Notes in Computer Science*, Vol. 5585, pp.25–38.

Xu, R., Zhou, S. and Li, W.J. (2012) 'MEMS accelerometer based nonspecific-user hand gesture recognition', *IEEE Sensors Journal*, Vol. 12, No. 5, pp.1166–117.

Yin, L., Dong, M. and Duan, Y. (2014) 'A high-performance training-free approach for hand gesture recognition with accelerometer', *Multimedia Tools and Applications*, Vol. 72, No. 1, pp.843–864.

Zhang, X., Chen, X., Li, Y., Lantz, V., Wang, K. and Yang, J. (2011) 'A framework for hand gesture recognition based on accelerometer and EMG sensors', *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, Vol. 41, No. 6, pp.1064–1076.

Zhu, Q., Weicun, Z., Zhang, J. and Sun, B. (2019) 'U-neural network-enhanced control of nonlinear dynamic systems', *Neurocomputing*, Vol. 352, No. 4, pp.12–21.