

Article

# Differences in Motion Accuracy of Baduanjin between Novice and Senior Students on Inertial Sensor Measurement Systems

Hai Li <sup>1,2</sup> , Selina Khoo <sup>1</sup>  and Hwa Jen Yap <sup>3,\*</sup> 

<sup>1</sup> Centre for Sport and Exercise Sciences, University of Malaya, Kuala Lumpur 50603, Malaysia; 10000861@njtc.edu.cn (H.L.); selina@um.edu.my (S.K.)

<sup>2</sup> College of Sport, Neijiang Normal University, Neijiang 641112, China

<sup>3</sup> Department of Mechanical Engineering, Faculty of Engineering University of Malaya, Kuala Lumpur 50603, Malaysia

\* Correspondence: hjyap737@um.edu.my; Tel.: +603-79675240

Received: 12 September 2020; Accepted: 30 October 2020; Published: 2 November 2020



**Abstract:** This study aimed to evaluate the motion accuracy of novice and senior students in Baduanjin (a traditional Chinese sport) using an inertial sensor measurement system (IMU). Study participants were nine novice students, 11 senior students, and a teacher. The motion data of all participants were measured three times with the IMU. Using the motions of the teacher as the standard motions, we used dynamic time warping to calculate the distances between the motion data of the students and the teacher to evaluate the motion accuracy of the students. The distances between the motion data of the novice students and the teacher were higher than that between senior students and the teacher ( $p < 0.05$  or  $p < 0.01$ ). These initial results showed that the IMU and the corresponding mathematical methods could effectively distinguish the differences in motion accuracy between novice and senior students of Baduanjin.

**Keywords:** motion capture; inertial sensor measurement systems; motion accuracy; Baduanjin; physical education

## 1. Introduction

Traditional Chinese sport has been a compulsory component of Physical Education (PE) in universities in China since 2002 [1]. Although there are various traditional Chinese sports to choose from, 76.7% of universities taught martial arts in their PE curriculum [2]. In 2016, the Communist Party of China and the Chinese government adopted the ‘Healthy China 2030’ national health plan [3]. In this plan, Baduanjin was identified as a traditional Chinese sport that was promoted and supported by the government. This resulted in increased Baduanjin teaching and research in universities throughout the country [4].

Although universities in China must incorporate traditional Chinese sports into their PE curriculum, there have been problems with its implementation. These include a high student-teacher ratio, uninteresting forms of teaching-learning resources, and an incomplete assessment system. These three problems adversely affected the requirements for teaching quality set by the People’s Republic of China Ministry of Education [5,6]. Although the high student–teacher ratio has been a problem since 2005, it has yet to be resolved [7,8]. Teachers are not able to provide individual guidance to each student because of the large number of students in the class. As a result, teachers cannot correct all the students’ mistakes, and students are not aware of their incorrect movements [9].

In recent years, motion capture (Mocap) has been widely applied in fields such as clinical and sports biomechanics to distinguish between different types of motions or analyze differences between

motions [10,11]. Studies have also applied Mocap in PE for adaptive motion analysis to evaluate the motion quality of learners and feedback the information to assist them in detecting and correcting their inaccurate motions. In the study by Koji Yamada et al. [12], a system based on Mocap was developed for Frisbee learners. Researchers used the Kinect device to obtain 3D motion data of learners during exercise, detect their pre-motion/motion/post-motion, and display the feedback information to improve their motions. The results showed that the system developed by the researchers can effectively improve the motions of learners [12]. Chen et al. [13] applied Kinect in Taichi courses in universities. The 3D data of motions from novice students captured by Kinect were compared with an expert in order to evaluate the quality of motions and students were informed of their results. The research showed that the motion evaluation system on Kinect developed by the researchers accelerated learning by novice Taichi students. More recently, Amani Elaoud et al. [14] used Kinect V2 to obtain red, green, blue and depth (RGB-D) data of motion. They used these data to compare the differences between novice students and experts on the central angles of points that affect throwing performance in handball. These experiments show how researchers have used various categories of Mocap.

Based on different technical characteristics, the application of Mocap in PE can be divided into four categories: optoelectronic system (OMS) [15], electromagnetic system (EMS), image processing systems (IMS), and inertial sensor measurement system (IMU) [16]. In these four Mocap categories, OMS is the most accurate and is considered to be the gold standard in motion capture [16,17]. However, OMS requires a large number of high-precision and high-speed cameras that will inevitably result in issues related to cost, coordination, and manual use [18]. Moreover, OMS cannot capture the movement of objects when the marker is obscured [19]. These deficiencies have limited the practical application of OMS in PE. The advantage of EMS over OMS is that it can measure motion data of a specific point of the body regardless of visual shielding [20]. However, EMS is susceptible to interference from the electromagnetic environment which distorts measurement data [21]. Also, EMS has to be kept within a certain distance from the base station, which limits the use range [22]. IMS has better accuracy compared to EMS and an improved range compared to OMS [16]. Most studies have used low-cost IMS (such as the Kinect device) to capture motion for analyzing motion in PE. However, there are some disadvantages in low-cost IMS, namely low-accuracy, insufficient environment adaptability, and limited range of motion because the Kinect sensor has a small field of vision [16]. High-performance IMS does not have these shortcomings. Generally, high-performance IMS has favorable accuracy and a good measurement range. However, high-performance IMS requires expensive high quality and/or high-speed cameras which has limited its application [16].

Based on the disadvantages of low-cost IMS in its application in PE, applying IMU (a motion capture consisting of an accelerometer, gyroscope, and a magnetometer) in PE may mitigate these application problems [23]. In recent years, the development of technology has reduced the cost of IMU, making it possible to be used in PE. The validity of assessing motion accuracy of IMUs has been confirmed. Poitras et al. [24] confirmed the criterion validity of a commercial IMU system (MVN Awinda system, Xsens) by comparing it to a gold standard optoelectronic system (Vicon). Compared to low-cost IMS, IMU has certain advantages in environmental adaptability and a sufficient range of motion. The IMU does not require any base station to work, which means it is the most mobile of the available motion capture systems [16]. Moreover, IMU can measure high-speed movements and is non-invasive for the user, making it an attractive application for PE [16,25].

However, there are a few issues capturing motions using IMU. First, the IMU sensors are sensitive to metal objects nearby which distort the measurement data [16]. Therefore, participants should wear fewer metal objects when capturing motion using IMU. Fortunately, in traditional Chinese martial arts such as Baduanjin, exercisers should, in principle, wear traditional Chinese costumes without any metal, which minimizes the impact of metals on IMU sensors. Second, a common IMU system, Perception Neuron 2.0, was used in our research which uses data cables to connect all sensors with a transmitter. Although users cannot wear this wired IMU system on their own, it does not affect the

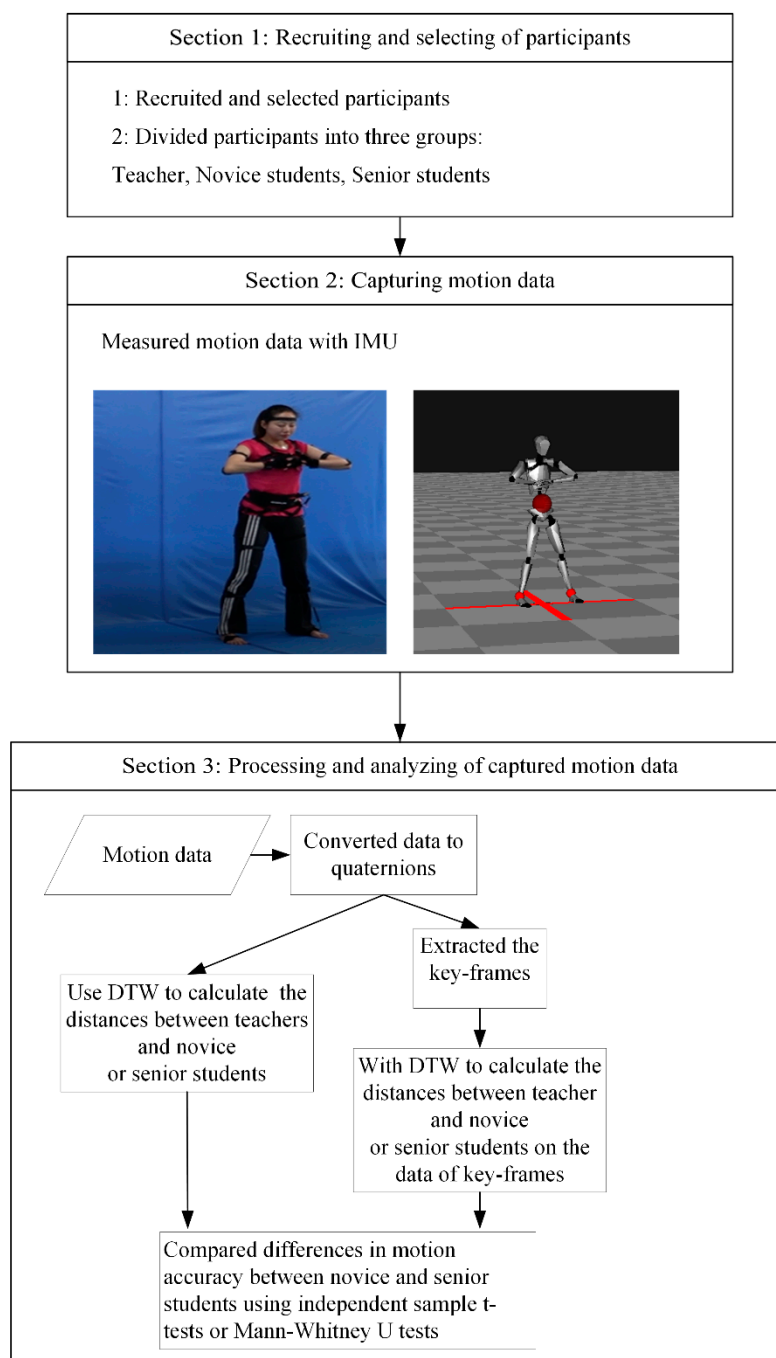
accuracy of the data. Also, the latest IMU overcomes this problem that users can't wear it on their own by having each sensor transmit the data to the external receiving terminal directly [26].

Therefore, we propose applying IMU in Baduanjin by developing a system that assesses and records the quality of motions to assist teachers and students in determining inaccurate motions. Using IMU, students can learn Baduanjin independently after class and teachers can evaluate students' progress, which is useful for formative assessment. That may alleviate current problems faced in PE classes in Chinese universities. For this purpose, we explored the feasibility of using an IMU to distinguish the difference in motion accuracy of Baduanjin between novice and senior students.

## 2. Materials and Methods

### 2.1. Overview

This study consists of three sections, namely recruiting and selecting participants, capturing motion data of Baduanjin participants on IMU, and processing and analyzing the motion data. We invited teachers and students from a university in Southwest China to participate in the study. We divided them into three groups—teachers, novice students, and senior students. We captured motion data of all participants on IMU when they practised Baduanjin. The motion data were converted to quaternions and analysed in two different ways. The first way was based on the quaternions of motion, where dynamic time warping (DTW) was used to calculate the distances between the quaternions of the teacher and the two groups of students (novice and senior). The motion accuracy of the students was expressed by distances. DTW is a classic similarity method to solve the time-warping issue in similarity computation of time series [27]. Compared with the other methods, namely the hidden Markov model (HMM) and symbolic aggregate approximation (SAX), the taken time of DTW is shorter [28,29]. Considering that, in the actual teaching, students need to get feedback information and a large number of student data in realtime, we adopted DTW in the study. The second way used the extracted key-frames to calculate distances. Based on the quaternions of key-frames, DTW was used to calculate the distances. Finally, based on data of the distances, an independent sample T-test or Mann–Whitney U test was used to define whether the motions of the two groups of students (novice and senior) were different in motion accuracy (see Figure 1).



**Figure 1.** Flow diagram of the study.

## 2.2. Recruiting and Selecting Participants

In this study, we invited a martial arts PE teacher and undergraduate students to participate in the study. The inclusion criteria for participants are as follows:

**Teacher:** martial arts PE teacher, former national martial arts athlete, with an undergraduate and master's degree in traditional Chinese sports (martial arts specialization), and more than ten years' experience teaching Baduanjin.

**Novice students:** undergraduate students in the university with no experience of Baduanjin, without a disability and no clinical or mental illness.

**Senior students:** undergraduate students in the university who have passed Baduanjin in their PE course, without a disability, and no clinical or mental illness.

Participants read the information sheet that outlined the purpose and procedure of the study. Those who agreed to participate were given the consent form to sign.

### 2.3. Capturing Motion Data of Participants on IMU

Baduanjin is a traditional Chinese martial art for fitness. The speed of motions is relatively slow [30]. We used IMU to capture the motion data of the teacher, novice, and senior students for eight standard motions of Baduanjin as shown in Figure 2.



**Figure 2.** Eight standard motions of Baduanjin.

### 2.3.1. IMU

We used Perception Neuron 2.0, a low-cost IMU developed by Noitom, to capture Baduanjin motion data of participants [31]. This IMU includes 17 inertial sense units and each unit comprised a 3-axis gyroscope, 3-axis accelerate, and 3-axis magnetometer, which measures and records the rotation angle data of 17 position points of human movement. Sers et al. [32] compared the IMU used in our study with a gold standard optoelectronic system (Vicon), and confirmed the IMU's effectiveness in measuring motion accuracy. The supporting software of the IMU, Axis neuron software developed by Noitom, transforms the recorded data into Biovision Hierarchy (BVH) motion files.

### 2.3.2. Capturing Motion Data

Before measuring the motion data, the teacher and senior students practised Baduanjin for 30 min. As the novice students had not learned Baduanjin, they followed the demonstration of the teacher practising Baduanjin for 30 min. After the practice, the motion data of participants were measured by IMU. No feedback was given to students during practice.

## 2.4. Data-Analysis

### 2.4.1. Extracting and Converting Raw Data

The raw data was converted into BVH file by the Axis neuron software. The BVH file is a file format developed by the BVH Company to store skeleton hierarchy information and three-dimensional motion data [33]. The BVH file comprises two parts: one is used to store skeleton hierarchy information and the other to store motion information. The skeleton hierarchy information includes the connection relationship between joint points and the offsets of the child joint points from their parent skeleton points. In the skeleton hierarchy, the first skeleton point is defined as Root. Root is the parent of all other skeleton points in the skeleton hierarchy. Motion information stores the global translation amount and the rotation amount of Root in each frame of the movement. The global translation amount is the position coordinate: X position, Y position, and Z position in the world coordinate system and the rotation amount is the rotation component: X rotation, Y rotation, and Z rotation in the Euler angle [33]. The motion information of other skeleton points is recorded on the rotation amount related to the parent points. The IMU used 17 sensors to measure motion data on 17 points of the body and the recorded order of the rotation amount of each point is Z rotation, Y rotation, and X rotation. The skeleton hierarchy information of BVH on the IMU and the skeleton model are shown in Figure 3.

In the BVH file, the rotation data is recorded on the Euler angle of 17 skeleton points. Some issues with rotation data expressed on the Euler angle (gimbal lock and singularity problems) were overcome using quaternion [34]. Quaternion is a 4-dimensional hyper-complex number, expressing a three-dimensional vector space on real numbers [35]. We used four-tuple notation to represent quaternion as follows:

$$q = [w, x, y, z] \quad (1)$$

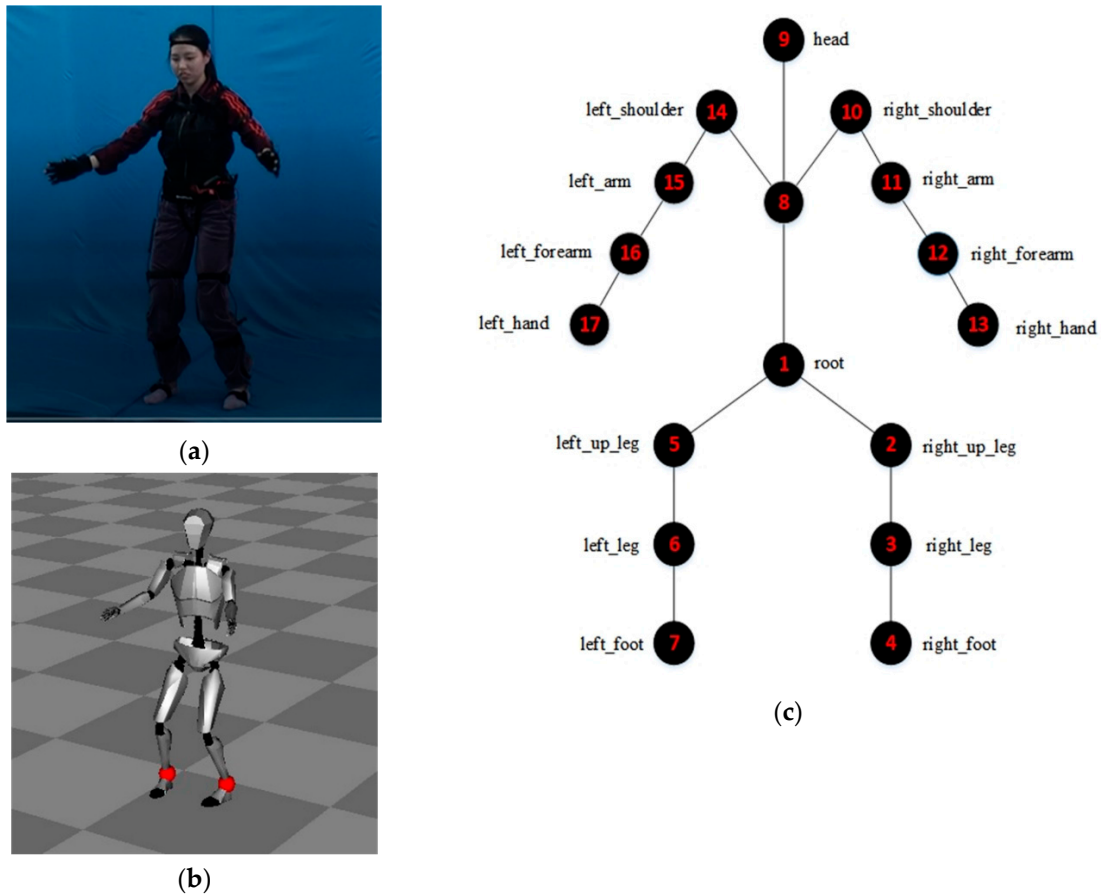
In this quaternion,  $w$  is the scalar component, and  $x, y, z$  are the vectors.

Therefore, the format of the rotation data from BVH files was converted from Euler angle to quaternion. If the order of rotation in Euler angle is  $z, y, x$ , we used  $\alpha, \beta, \gamma$  to represent the rotation angles of the object around  $x, y$ , and  $z$  axes. The corresponding quaternion can be converted as follows:

$$q = \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos(\gamma/2) \\ 0 \\ 0 \\ \sin(\gamma/2) \end{bmatrix} \begin{bmatrix} \cos(\beta/2) \\ 0 \\ \sin(\beta/2) \\ 0 \end{bmatrix} \begin{bmatrix} \cos(\alpha/2) \\ \sin(\alpha/2) \\ 0 \\ 0 \end{bmatrix} \quad (2)$$

$$= \begin{bmatrix} \cos(\gamma/2) \cos(\beta/2) \cos(\alpha/2) + \sin(\gamma/2) \sin(\beta/2) \sin(\alpha/2) \\ \cos(\gamma/2) \cos(\beta/2) \sin(\alpha/2) - \sin(\gamma/2) \sin(\beta/2) \cos(\alpha/2) \\ \cos(\gamma/2) \sin(\beta/2) \cos(\alpha/2) + \sin(\gamma/2) \cos(\beta/2) \sin(\alpha/2) \\ \sin(\gamma/2) \cos(\beta/2) \sin(\alpha/2) - \cos(\gamma/2) \sin(\beta/2) \sin(\alpha/2) \end{bmatrix}$$





**Figure 3.** The skeleton hierarchy information of BVH on the IMU (Perception Neuron 2.0): (a) A participant wearing the IMU (Perception Neuron 2.0) to measure the motion data; (b) The interface of Perception Neuron 2.0; and (c) The skeleton model of BVH file for Perception Neuron 2.0.

#### 2.4.2. Extracting Key-Frames

After extracting the motion data, we used key-frames extraction to reduce the motion data. Due to the limited storage and bandwidth capacity available to users, the large amount of motion data collected on Mocap may restrict its application [36]. Key-frames extraction, which extracts a small number of representative key-frames from a long motion sequence, is widely used in motion analysis. This technology can reduce the data amount, which facilitates data storage and subsequent data analysis [36,37].

##### Extraction of Key-Frames on Inter-Frame Pitch

We used the distance between quaternions to evaluate the inter-frame pitch between frames and set a threshold of inter-frame pitch to extract key-frames [38]. The method is based on the rotation data of each skeleton point which is represented as a quaternion and uses a simple form to evaluate the distance between two quaternions. The inter-frame pitch between the two frames is assessed by the sum of the distances between the quaternions of every point. The process is constructed with three sections: calculating the distance between quaternions, calculating the inter-frame pitch between frames, and extracting key-frames on the set threshold of inter-frame pitch.

##### 1. The distance between quaternions

To evaluate the distance between two quaternions, the conjugate quaternion  $q^*$  of a quaternion is defined as follows:

$$q^* = [w, -x, -y, -z] \quad (3)$$

and the quaternion norm  $\|q\|$  is defined as follows:

$$\|q\| = \sqrt{w^2 + x^2 + y^2 + z^2} \quad (4)$$

then:

$$\|q\|^2 = qq^* = w^2 + x^2 + y^2 + z^2 \quad (5)$$

when a quaternion norm  $\|q\|$  is 1, which means:

$$w^2 + x^2 + y^2 + z^2 = 1 \quad (6)$$

the quaternion is a unit quaternion. A quaternion is converted to a unit quaternion by dividing it by its norm.

From the definitions of conjugate quaternion, quaternion norm, and unit quaternion, we can define the inverse of a quaternion ( $q^{-1}$ ) as follows [39]:

$$q^{-1} = \frac{1}{\|q\|} = \frac{1}{\|q\|^2} q^*, \quad \|q\| \neq 0 \quad (7)$$

According to Shunyi et al. [38], if there are two quaternions:  $q_1, q_2$  are unit quaternions and:

$$q_1 q_2^{-1} = [w, x, y, z] \quad (8)$$

the distance between the quaternions  $q_1$  and  $q_2$  is:

$$d(q_1, q_2) = \arccos w \quad (9)$$

Therefore, we converted the rotation of a skeleton point based on Euler angles into quaternion, then normalized and converted the quaternion into unit quaternion, and finally calculated the difference between any two quaternions of the point according to Equation (9).

## 2. Calculation of Inter-Frame Pitch between Two Frames

We used the sum of the differences between the quaternions at 17 skeleton points to evaluate the inter-frame pitch between two frames. The human motion represented by the BVH file are discrete-time vectors, which is the same after conversion to quaternions [38]. The weightage for different points needs to be taken into account when calculating the inter-frame pitch due to the tree-structure (parent-child) of the BVH format. Referring to the methods used in previous research [38,40], and the relationship structure between the skeleton points on the IMU in this study (see Figure 3), we assigned the weightage values of the 17 skeleton points as shown in Table 1.

If  $t_1$  and  $t_2$  are the two frames in a sequence of frames, we defined the inter-frame pitch between two frames:  $t_1$  and  $t_2$  as the following equation:

$$D(t_1, t_2) = \sum_{i=1}^n W_i d(q_i(t_1), q_i(t_2)) \quad (10)$$

In Equation (10),  $n$  represents the total number of skeleton points ( $n = 17$ ),  $W_i$  represents the weightage of each skeleton point (shown in Table 1), and  $q_i$  represents the quaternions of each skeleton point.

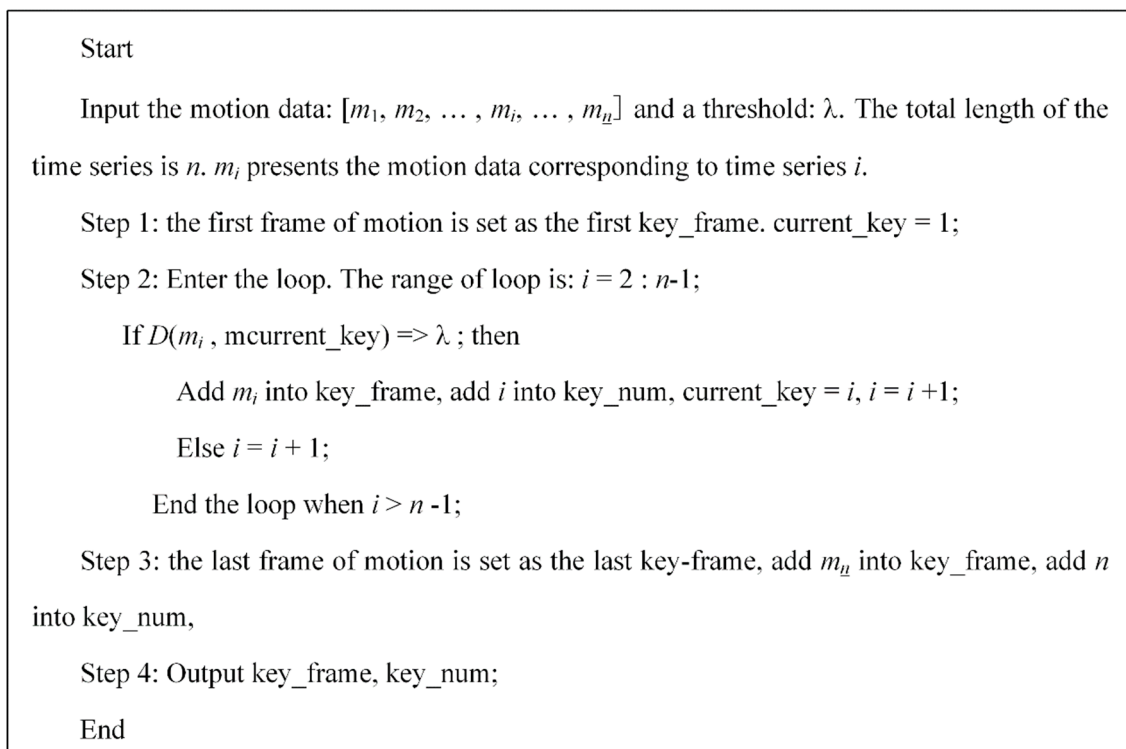


**Table 1.** The weightage of the 17 skeleton points.

Point	Weightage
Hip	16
Right up leg	8
Right leg	4
Right foot	2
Left up leg	8
Left leg	4
Left foot	2
Spine	8
Head	4
Right shoulder	4
Right arm	2
Right forearm	1
Right hand	0.5
Left shoulder	4
Left arm	2
Left forearm	1
Left hand	0.5

### 3. Key-frames extraction on the set threshold of inter-frame pitch

Based on the inter-frame pitch between two frames, we set: `key_frame` as an array to store the quaternion corresponding to the key-frames of motion; `key_num` as a set of vector to store the serial number corresponding to a key-frame; `key_num1` presents the time series number corresponding to the first key-frame; `current_key` as the last frame in the set of `key_num`.  $\lambda$  is a preset threshold value of inter-frame pitch which is mainly determined based on the demand for a compression rate of frames. The algorithm steps are shown in Figure 4.

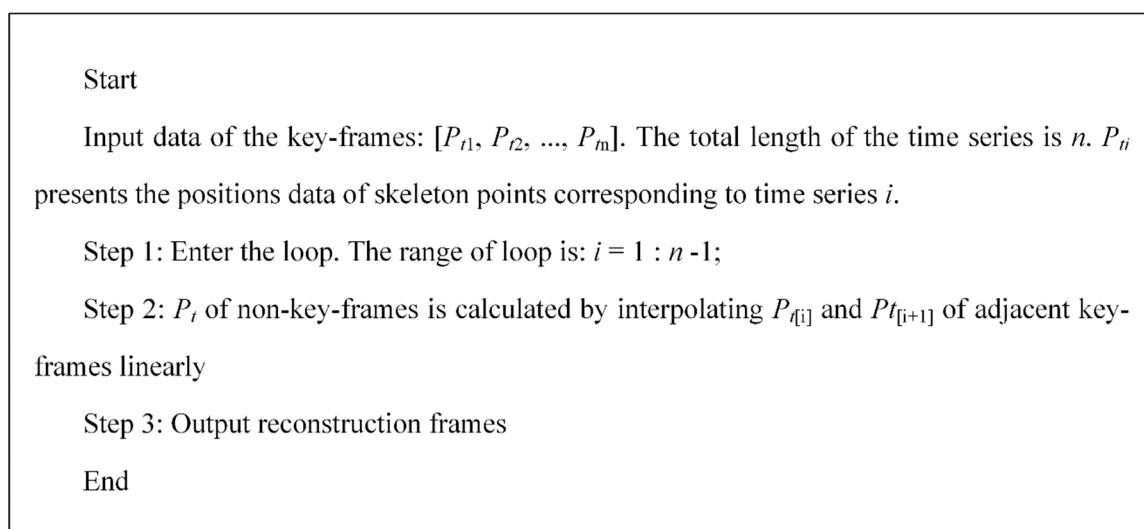
**Figure 4.** The algorithm steps of key-frames extraction.

#### 4. Motion reconstruction error

The purpose of motion reconstruction is to rebuild the same number of frames as the original frames based on interpolation reconstruction of non-key-frames between adjacent key-frames [38,41]. First, individually, the position coordinates (in the world coordinate system) of points were calculated on the point hierarchy and relative rotation angle between the points in the BVH file. Second, given that  $p_{t_1}$  and  $p_{t_2}$  are the positions of a point of adjacent key-frames in time  $t_1$  and  $t_2$ , then  $p_t$  (representing the position of a point of non-key-frame in time  $t$ ) is calculated by linear interpolation between  $p_{t_1}$  and  $p_{t_2}$  as follows [41]:

$$\begin{aligned} p_t &= u(t)p_{t_1} + (1 - u(t))p_{t_2}, \\ u(t) &= \frac{t_2 - t}{t_2 - t_1}, \\ t_1 &< t < t_2 \end{aligned} \quad (11)$$

The algorithm steps of motion reconstruction are shown in Figure 5.



**Figure 5.** The algorithm steps of motion reconstruction.

In this study, we used the position error of the human posture to calculate the reconstruction error between the reconstructed frames and the original frames [38]. Assuming  $m_1$  is the original motion sequence,  $m_2$  is the reconstruction motion sequence from the key-frames, the reconstruction error  $E(m_1, m_2)$  is evaluated as [42]:

$$E(m_1, m_2) = \frac{1}{n} \sum_{i=1}^n D(p_1^i - p_2^i) \quad (12)$$

The distance of human posture is used to measure the position error of human posture:

$$D(p_1^i - p_2^i) = \sum_{k=1}^m \|p_{1,k}^i - p_{2,k}^i\|^2 \quad (13)$$

In this equation,  $m$  represents the total number of skeleton points,  $p_{1,k}^i$  is the position of  $k$  point in  $i$  frame of the original motion sequence, and  $p_{2,k}^i$  is the position of  $k$  point in  $i$  frame in the reconstruction sequence.

## Extraction of Key-Frames on Clustering

A problem with the key-frames extraction on inter-frame pitch is that the compression rate of the key-frames with the same inter-frame threshold for different actions may vary considerably [40]. As the eight motions of Baduanjin are quite different, the key-frames extraction on the inter-frame pitch may cause some motions to be compressed too much, and some motions not compressed enough. Therefore, we also chose another way to extract key-frames on clustering. This method was used for key-frames with the pre-set compression rate [43].

### 1. K-means clustering algorithm

K-means clustering algorithm is an iterative partition clustering algorithm. In this key-frame extraction method, we used the K-means clustering algorithm to cluster the 3D coordinates  $([x, y, z])$  of the skeleton points in the original frame. Assuming that the total length of the original frames is  $N$ ,  $i$  represents the  $i$  frame in  $N$ .  $p^i$  is the vectors of the 3D coordinate positions of all relevant skeleton points of the  $i$  frame in the original frames. Therefore, the vectors collection of the 3D coordinate data of every point of original frames is  $(p^1, p^2, \dots, p^i), p^i \in R^N$ . According to the K-means clustering algorithm, the data of skeleton points ( $R^N$ ) in the frames is clustered into  $K$  ( $K \leq N$ ) clusters as follows [44]:

Step 1: Randomly select  $K$  cluster centroids from  $R^N$  are  $u_1, u_2 \dots u_K$ ;

Step 2: Repeat the following process to get convergence.

For the  $p^i$  corresponding to one frame, we calculated the distances from each cluster centroid ( $u_j, j \in K$ ) and classified it into the class corresponding to the minimum distance [45]:

$$D = \operatorname{argmin} \sum_{i=1}^N \sum_{j=1}^K \|p^i - u_j\|^2 \quad (14)$$

In this equation,  $D$  represents the minimum distance between the cluster centroid and the centre of  $p^i$ , and when  $D$  is the smallest,  $p^i$  is classified into class  $j$ .

For each class  $j$ , the cluster centroid ( $u_j$ ) of that class was recalculated:

$$u_j = \frac{\sum_{i=1}^N r_{ij} p^i}{\sum_{i=1}^N r_{ij}} \quad (15)$$

In this equation,  $r_{ij}$  indicates that when  $p^i$  is classified as  $j$ , it is 1; otherwise, it is 0.

### 2. Key-frames extraction

Using the above k-means clustering algorithm, we extracted  $K$  cluster centroids from the original frame. Each cluster is clustered from the 3D coordinates of the 17 points in the original frames. Therefore, one cluster centroid is constructed with 51 ( $17 \times 3$ ) vectors. Based on these cluster centroids, we extracted the key-frames by calculating the Euclidean distance between the cluster centroid of each point and the corresponding point coordinates in the original frames. The steps to extract key-frames are as follows:

Start

Input the 3D coordinate data of every point of the original frames:

$$\begin{aligned} & (p^1, p^2 \dots p^i), \quad p^i \in R^N; \\ & p^i = (p_1^i, p_2^i \dots p_j^i), \quad j = 17; \\ & p_j^i = [x_j^i, y_j^i, z_j^i] \end{aligned} \quad (16)$$

and the number of key-frames to be extracted is  $K$ ;

Step 1: Using the k-means clustering algorithm to calculate cluster centroids of the  $K$  clusters are expressed as:

$$\begin{aligned} u_m &= (u_{m1}, u_{m2} \dots u_{mj}), m \in (1, 2, 3 \dots K), j = 17; \\ u_{mj} &= [x_{mj}, y_{mj}, z_{mj}] \end{aligned} \quad (17)$$

Step 2: Calculate the Euclidean distance of 3D coordinates between each point of the cluster and the corresponding point of the original frames:

$$\begin{aligned} C_m &= \min(u_m, p^i) \\ &= \sum_{j=1}^{17} \min(dis(u_{mj}, p_j^i)); \\ dis(u_{mj}, p_j^i) &= \|u_{mj} - p_j^i\|^2 \end{aligned} \quad (18)$$

$\min(dis(u_{mj}, p_j^i))$  means that after calculating the distances between  $m$  cluster and all original frames, the  $j$  point of  $p_i$  which value of  $dis(u_{mj}, p_j^i)$  is minimum is recorded as 1; otherwise, it is recorded as 0.  $i$  of  $p^i$  corresponding to the maximum value of  $C_m$  is a sequence of key-frames.

Step 3: Sequences of key-frames are arranged from small to large after extraction. If the first frame and the last frame in the original frames are not included in the key-frames, the first frame and the last frame must be added into key-frames.

End

In this key-frames extraction, the number of key-frames can be preset. The key-frames of the corresponding compression rate is obtained by presetting the compression rate as follows [42]:

$$K = c\_rate * N \quad (19)$$

where  $K$  is the number of key-frames to be extracted,  $c\_rate$  is the compression rate of the key-frame to be obtained, and  $N$  is the total number of original frames.

After extracting key-frames, we continued with the ways to motion reconstruction and evaluate reconstruction error as described above.

#### 2.4.3. Evaluate Motion the Accuracy of Motions Data

In this study, we referred to previous studies [13,46] to evaluate the motion accuracy of student motions by assessing the differences between students' motions and teacher's motions. Due to the difference in speed between individual movements, different time series were considered when assessing the difference between two motions. We chose DTW, a well-established method, to account for different time series to evaluate the difference in the motions between teachers and students [47]. Since DTW compares the other methods, i.e., HMM and SAX, without a training stage, the taken time is shorter. First, the derived quaternions were normalized in unit length of a quaternion:  $q = [w, x, y, z]$  can be described as:  $\|q\| = 1$  and  $w^2 + x^2 + y^2 + z^2 = 1$ . Therefore, three components ( $x, y, z$ ) out of the four components ( $w, x, y, z$ ) of the quaternions can be used to represent the rotations of the skeleton points over a temporal domain. Then, we used DTW to evaluate the difference between two sequences of motions on the skeleton points. First, we assessed the difference between two motions on a single skeleton point. For example, there are two motion data on quaternions for a skeleton point from a teacher and a student, one from the teacher:  $q_{tea}(t)$ , one from a student:  $q_{stu}(t)$ . The length of the two sequences of quaternions are  $n$  and  $m$ :

$$\begin{aligned} q_{tea}(t) &= q_{tea}(1), q_{tea}(2), \dots, q_{tea}(i), \dots, q_{tea}(n) \\ q_{stu}(t) &= q_{stu}(1), q_{stu}(2), \dots, q_{stu}(j), \dots, q_{stu}(m) \end{aligned} \quad (20)$$

The vector in the quaternion arrays consists of three components ( $x, y, z$ ) of quaternions. A distance matrix ( $n \times m$ ) is constructed to align the quaternions of two sequences. The elements ( $i, j$ ) in the matrix represent the Euclidean distance:  $dis(q_{tea}(i), q_{stu}(j))$  between the two points  $q_{tea}(i)$  and  $q_{stu}(j)$ :

$$dis(q_{tea}(i), q_{stu}(j)) = |q_{tea}(i) - q_{stu}(j)|^2 \quad (21)$$

In the distance matrix, many paths are from the upper-left corner to the lower-right corner of the distance matrix. We used  $\Phi k$  to represent any point on these paths:  $\Phi k = (\Phi_{tea}(k), \Phi_{stu}(k))$  where:

$\Phi_{tea}(k)$ : the value of  $k$  is  $1, 2, \dots, n$ ,

$\Phi_{stu}(k)$ : the value of  $k$  is  $1, 2, \dots, m$ ,

$\Phi k$ , the value of  $k$  is  $1, 2, \dots, T, (T = n \times m)$

We found a suitable path as the warping path, where the cumulative distance of path is the smallest of all paths [39]:

$$DTW(q_{tea}(t), q_{stu}(t)) = \min \sum_{k=1}^T dis(\Phi_{tea}(k), \Phi_{stu}(k)) \quad (22)$$

Then, the distance of  $DTW(q_{tea}(t), q_{stu}(t))$  is obtained through dynamic programming as follows [47]:

$$\begin{aligned} DTW(q_{tea}(t), q_{stu}(t)) &= f(n, m); \\ f(0, 0) &= 0; \\ f(0, 1) &= f(1, 0) = \infty; \\ f(i, j) &= dis(q_{tea}(i), q_{stu}(j)) + \min\{f(i-1, j), f(i, j-1), f(i-1, j-1)\}, (i = 1, 2, \dots, n; j = 1, 2, \dots, m) \end{aligned} \quad (23)$$

To prevent the wrong matching by excessive time warping, the warping path was constrained near the diagonal of the matrix by setting the global warping window for DTW [48,49]. In this study, the global warping window is set as 10 percent of the entire window span:  $0.1 \times \max(n, m)$ . The cumulative distance of the warping path represents the difference of rotation between teacher and student on the skeleton points is shown in Equation (22). Then, the macro difference between students' motions and teacher's motions was evaluated by taking the average of the cumulative distances of all the skeleton points as follows:

$$D(m_{tea}, m_{stu}) = \frac{\sum_{i=1}^n DTW(q_{tea}^i, q_{stu}^i)}{n} \quad (24)$$

In this equation,  $m_{tea}$  represents the teacher motion sequence;  $m_{stu}$  represents the students' motion sequence,  $q^i$  is the vectors of the quaternion of  $i$  skeleton point in the two motion sequences, and the total number of skeleton points is  $n$ .

Finally, data of the differences were analysed using IBM SPSS Statistics 25.0 to assess if there were significant differences in the motion accuracy of the two groups of students (novice and senior students) on the whole and each point. We used the independent sample T-test on data with normal distribution and the Mann-Whitney U test on data with non-normal distribution.

### 3. Results

#### 3.1. Demographic Characteristics of Participants

We recruited 21 participants for this study, including a martial arts teacher, nine undergraduate students who have not learned Baduanjin (novice students), and 11 undergraduate students who had completed the Baduanjin course (senior students). All participants gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the University of Malaya Research Ethics Committee (UM.TNC2/UMREC-558). The demographic characteristics of the students are shown in

Table 2. For each mean duration of the eight motions shown in Table 3, we measured all participants three times with IMU, resulting in 63 motion data.

**Table 2.** Demographic characteristics of the students.

Group	Gender	Age (years) (mean $\pm$ SD)	Height (cm) (mean $\pm$ SD)	Weight (kg) (mean $\pm$ SD)
Novice Student	Male: 5	18.60 $\pm$ 0.55	169.40 $\pm$ 3.91	58.20 $\pm$ 4.60
	Female: 4	18.75 $\pm$ 0.96	161.25 $\pm$ 3.59	48.25 $\pm$ 2.06
Senior Student	Male: 4	20.25 $\pm$ 0.50	170.75 $\pm$ 4.11	60.50 $\pm$ 4.80
	Female: 7	20.00 $\pm$ 0.82	161.43 $\pm$ 3.84	48.57 $\pm$ 3.74

**Table 3.** Mean duration of Baduanjin.

Motion	Valid	Mean duration (s) (mean $\pm$ SD)
Motion 1	63	12.24 $\pm$ 2.49
Motion 2	63	21.52 $\pm$ 4.19
Motion 3	63	16.75 $\pm$ 4.14
Motion 4	63	14.79 $\pm$ 3.85
Motion 5	63	19.18 $\pm$ 5.25
Motion 6	63	16.95 $\pm$ 4.09
Motion 7	63	13.10 $\pm$ 3.93
Motion 8	63	1.51 $\pm$ 0.25

### 3.2. Differences in Motion Accuracy between Novice and Senior Students on Original Frames

Algorithms explained in the data analysis section were coded with Matlab R2018b. Independent sample T-tests and Mann–Whitney U tests were used to assess the differences in motion accuracy of novice and senior students.

Before assessing macro differences, we assessed the normality of original frames data using the Shapiro–Wilk test (see Table 4).

**Table 4.** Normality of data of groups using the Shapiro–Wilk test.

Motion	Group	Statistic	df	Sig.
1	Novice Student	0.788	27	0.000
	Senior Student	0.753	33	0.000
2	Novice Student	0.962	27	0.408
	Senior Student	0.973	33	0.575
3	Novice Student	0.956	27	0.299
	Senior Student	0.948	33	0.120
4	Novice Student	0.963	27	0.428
	Senior Student	0.979	33	0.758
5	Novice Student	0.963	27	0.428
	Senior Student	0.979	33	0.758
6	Novice Student	0.789	27	0.000
	Senior Student	0.852	33	0.000
7	Novice Student	0.881	27	0.005
	Senior Student	0.878	33	0.001
8	Novice Student	0.932	27	0.078
	Senior Student	0.890	33	0.003



From Table 4, we can see that the data of the groups on Motions 2, 3, and 4 were normally distributed ( $p > 0.05$ ), whereas the others were not. Therefore, we assessed the differences in motion accuracy of Motions 2, 3, and 4 between novice and senior students using independent sample T-tests (see Table 5). The differences in the motion accuracy of other motions between novice and senior students were assessed using Mann–Whitney U tests (see Table 6).

**Table 5.** Differences in motion accuracy between novice and senior students on original frames (using the independent sample T-test).

Motion	Group	N <sup>1</sup>	Mean <sup>2</sup>	Std. Deviation	F	Sig.	t	Sig. <sup>3</sup>
2	Novice Student	27	640.76	74.38	2.289	0.136	4.275	0.000
	Senior Student	33	565.72	61.64			4.195	0.000
3	Novice Student	27	543.46	78.92	4.879	0.031	5.085	0.000
	Senior Student	33	455.75	54.30			4.903	0.000
4	Novice Student	27	536.45	41.44	0.061	0.806	5.805	0.000
	Senior Student	33	468.66	47.70			5.888	0.000

<sup>1</sup> Number of motions; <sup>2</sup> Mean of differences in motion between teacher and students; <sup>3</sup> 2-tailed.

**Table 6.** Differences in motion accuracy between Novice students and senior students on original frames (using the Mann-Whitney U test).

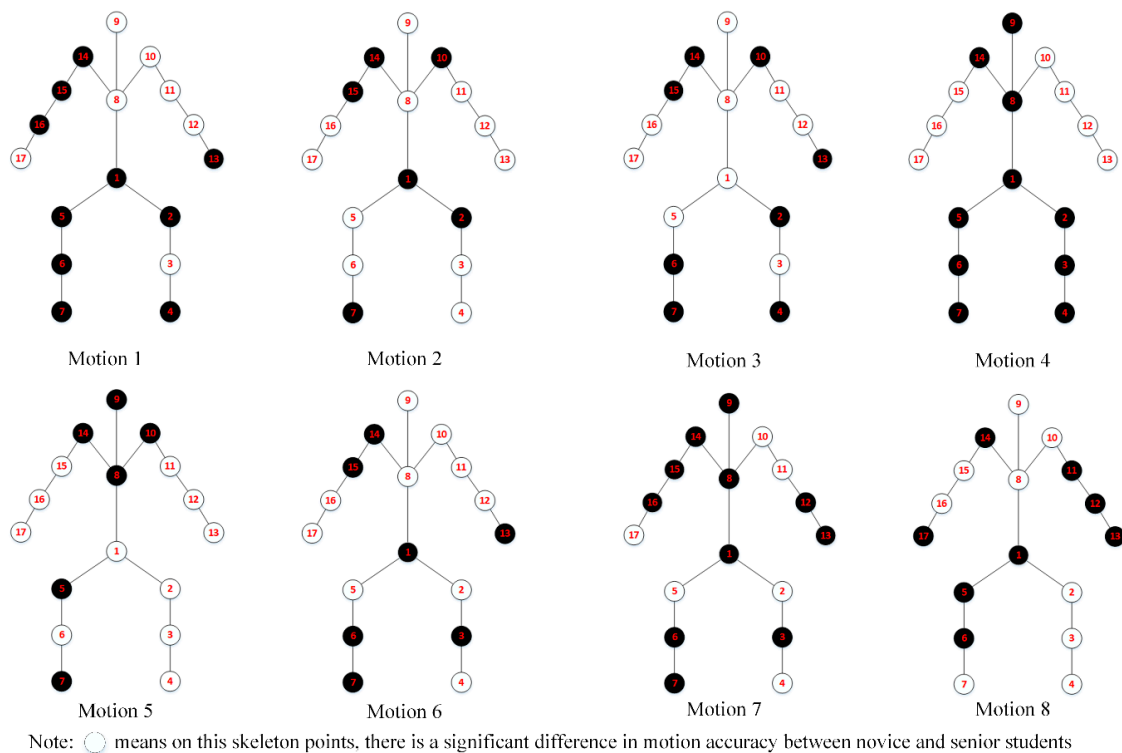
Motion	Group	N <sup>1</sup>	Mean Rank	Sum of Ranks	M-W U <sup>2</sup>	Wilcoxon W	Z	Asymp. Sig. <sup>3</sup>
1	Novice Student	27	38.52	1040.00	229.00	790.00	−3.217	0.001
	Senior Student	33	23.94	790.00				
5	Novice Student	27	41.96	1133.00	136.00	697.00	−4.599	0.000
	Senior Student	33	21.12	697.00				
6	Novice Student	27	35.93	970.00	299.00	860.00	−2.177	0.029
	Senior Student	33	26.06	860.00				
7	Novice Student	27	37.41	1010.00	259.00	820.00	−2.771	0.000
	Senior Student	33	24.85	820.00				
8	Novice Student	27	42.19	1139.00	130.00	691.00	4.688	0.000
	Senior Student	33	20.94	691.00				

<sup>1</sup> Number of motions; <sup>2</sup> Mann-Whitney U; <sup>3</sup> 2-tailed.

From Tables 5 and 6, we can see significant differences ( $p < 0.05$  or  $p < 0.01$ ) in motion accuracy of all eight motions between novice and senior students. The differences in motion accuracy between the teacher and senior students were lower than the differences in motion accuracy between the teacher and novice students.

We also evaluated the difference in motion accuracy on each skeleton point between novice and senior students (Figure 6).

From Figure 6, we found that out of the 17 points on eight motions of Baduanjin, there were significant differences in the motion accuracy between novice and senior students for some points. For example, in Motion 1, there were significant differences in motion accuracy between the two groups at the head and neck (points 8 and 9) and the right upper limb (points 10, 11, and 12).



**Figure 6.** Differences in motion accuracy of points between novice and senior students on original frames.

### 3.3. Differences in Motion Accuracy between Novice and Senior Students on Key-Frames

#### 3.3.1. Compression Rate and Reconstruction Error of Two Different Key-Frames Extraction Methods

Motion accuracy is assessed based on key-frames. In this study, we chose two methods to extract key-frames. In the key-frames extraction method on inter-frame pitch, we selected different thresholds (0.1, 0.5, 1.0, 1.5, 2.0) to extract key-frames and evaluated the compression rate and the reconstruction error of corresponding key-frames on different thresholds. The results are shown in Table 7.

**Table 7.** Compression rate and reconstruction error of corresponding key-frames on inter-frame pitch.

Threshold	Index	Motion							
		1	2	3	4	5	6	7	8
0.1	Rate <sup>1</sup>	60.45	80.57	55.84	62.00	86.84	76.58	58.83	67.30
	Error <sup>2</sup>	0.059	0.028	0.062	0.046	0.017	0.042	0.022	0.068
0.5	Rate	15.28	27.36	14.02	16.87	36.75	25.12	15.29	20.16
	Error	0.447	0.364	0.455	0.378	0.330	0.448	0.474	0.612
1.0	Rate	7.74	14.79	7.09	8.77	20.72	13.39	8.03	9.97
	Error	1.031	0.904	1.110	0.971	0.799	1.021	1.164	1.969
1.5	Rate	5.14	10.04	4.70	5.93	14.44	9.05	5.57	6.36
	Error	1.811	1.590	1.967	1.719	1.346	1.692	2.099	3.971
2.0	Rate	3.81	7.58	3.49	4.48	11.12	6.80	4.35	4.50
	Error	2.712	2.362	3.021	2.586	1.966	2.474	3.203	6.936

<sup>1</sup> Compression rate (%); <sup>2</sup> Reconstruction error.

Table 7 shows significant differences in the compression rates of the different motions extracted under the same threshold. We can see when the threshold value is set to 1 for obtaining key-frames using the inter-frame pitch, there was a difference in average compression rates ranging from 7.08% to 20.78% for the eight motions of Baduanjin. Moreover, when the threshold value increased, the number

of key-frames decreased, which decreased the compression rate. However, the error of motion reconstruction also increased. Based on the data in Table 7, it can be seen that in the five preset values, the compression rate and reconstruction error of the extracted key-frames are relatively reasonable when the threshold is 1. In the other key-frames extraction method on clustering, we chose different compression rates (5, 10, 15, 20, 25) to extract key-frames and evaluate the reconstruction error on different key-frames. The results are shown in Table 8.

**Table 8.** Reconstruction error of corresponding key-frames on clustering.

Rate (%) <sup>1</sup>	Reconstruction Error							
	Motion 1	Motion 2	Motion 3	Motion 4	Motion 5	Motion 6	Motion 7	Motion 8
5	4.528	7.185	3.875	3.430	8.790	6.823	3.886	6.206
10	1.484	2.428	1.268	0.997	2.927	2.359	1.281	2.216
15	0.851	1.244	0.638	0.547	1.485	1.286	0.650	1.342
20	0.498	0.757	0.401	0.353	0.971	0.797	0.415	0.769
25	0.366	0.518	0.281	0.235	0.700	0.569	0.297	0.531

<sup>1</sup> Compression rate (%) of key-frames.

From Table 8, we can see that as the compression rate increases, the error of motion reconstruction decreases. When the compression rate increased from 5% to 15%, the reconstruction error dropped sharply. But when the compression ratio increased from 15% to 25%, the reconstruction error decrease tended to be smooth. It can be seen that, in the five preset values, the compression rate and reconstruction error of the extracted key frames were relatively reasonable when the preset compression rate is 15%.

### 3.3.2. Differences in Motion Accuracy on Key-Frames

The differences in motion accuracy on key-frames between novice and senior students are shown in Tables 9 and 10.

**Table 9.** Differences in motion accuracy on the key-frames on inter-frame pitch between novice and senior students.

Threshold	p Value							
	Motion 1	Motion 2	Motion 3	Motion 4	Motion 5	Motion 6	Motion 7	Motion 8
0.1	0.000	0.000	0.000	0.000	0.000	0.006	0.075 <sup>1</sup>	0.000
0.5	0.001	0.000	0.001	0.005	0.000	0.004	0.122 <sup>1</sup>	0.000
1.0	0.001	0.000	0.001	0.017	0.004	0.004	0.122 <sup>1</sup>	0.000
1.5	0.001	0.000	0.002	0.050 <sup>1</sup>	0.008	0.004	0.141 <sup>1</sup>	0.000
2.0	0.001	0.001	0.004	0.112 <sup>1</sup>	0.018	0.004	0.176 <sup>1</sup>	0.000

<sup>1</sup>  $p \geq 0.05$ .

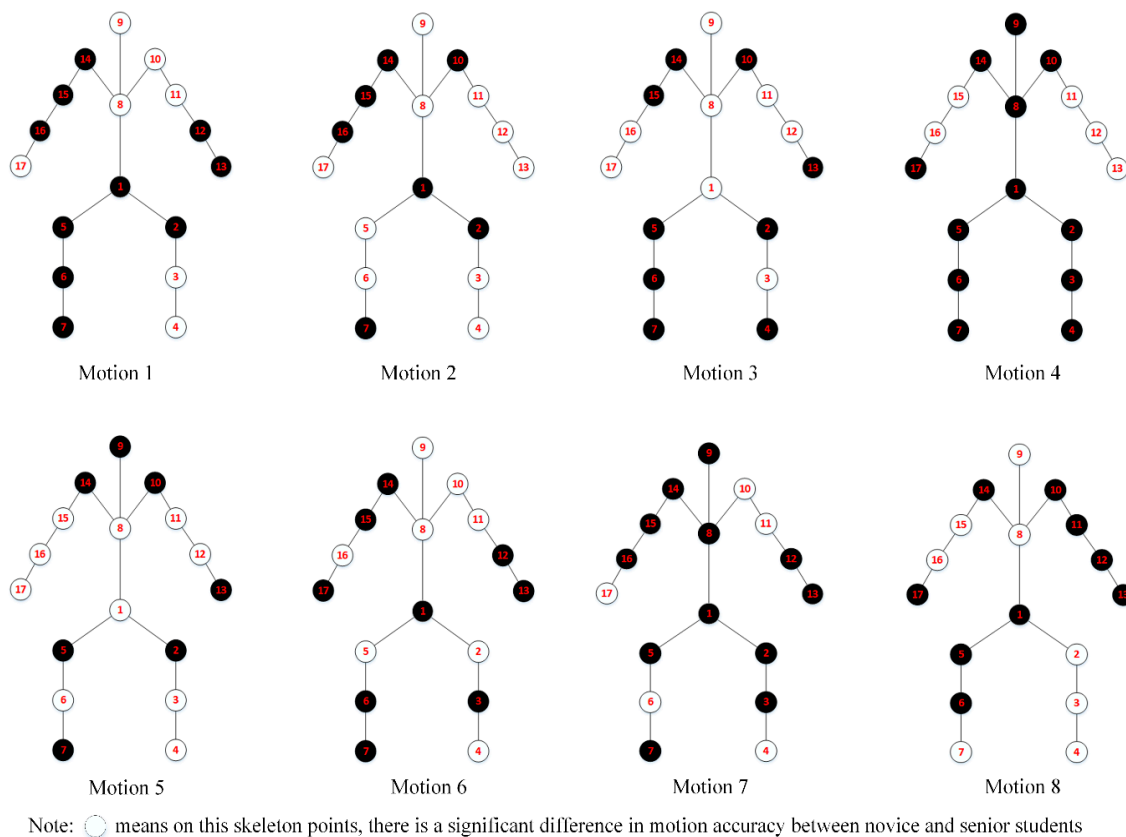
**Table 10.** Differences in motion accuracy on the key-frames on clustering between novice and senior students.

Rate (%) <sup>1</sup>	p Value							
	Motion 1	Motion 2	Motion 3	Motion 4	Motion 5	Motion 6	Motion 7	Motion 8
5	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.000
10	0.000	0.000	0.000	0.000	0.000	0.024	0.003	0.000
15	0.000	0.000	0.000	0.000	0.000	0.020	0.004	0.000
20	0.000	0.000	0.000	0.000	0.000	0.031	0.004	0.000
25	0.000	0.000	0.000	0.000	0.000	0.024	0.004	0.000

<sup>1</sup> Compression rate (%) of key-frames.

From the results of the key-frames on clustering, the motion accuracy of the eight motions of novice and senior students were significantly different. This result is consistent with the result based on the original frames. However, on the key-frames of inter-frame pitch on five different thresholds, there was no significant difference in motion accuracy between the two groups in Motion 7.

The differences in motion accuracy of points between the two groups on key-frames were also evaluated. Figure 7 shows the results on the key-frames of inter-frame pitch when the setting threshold = 1.

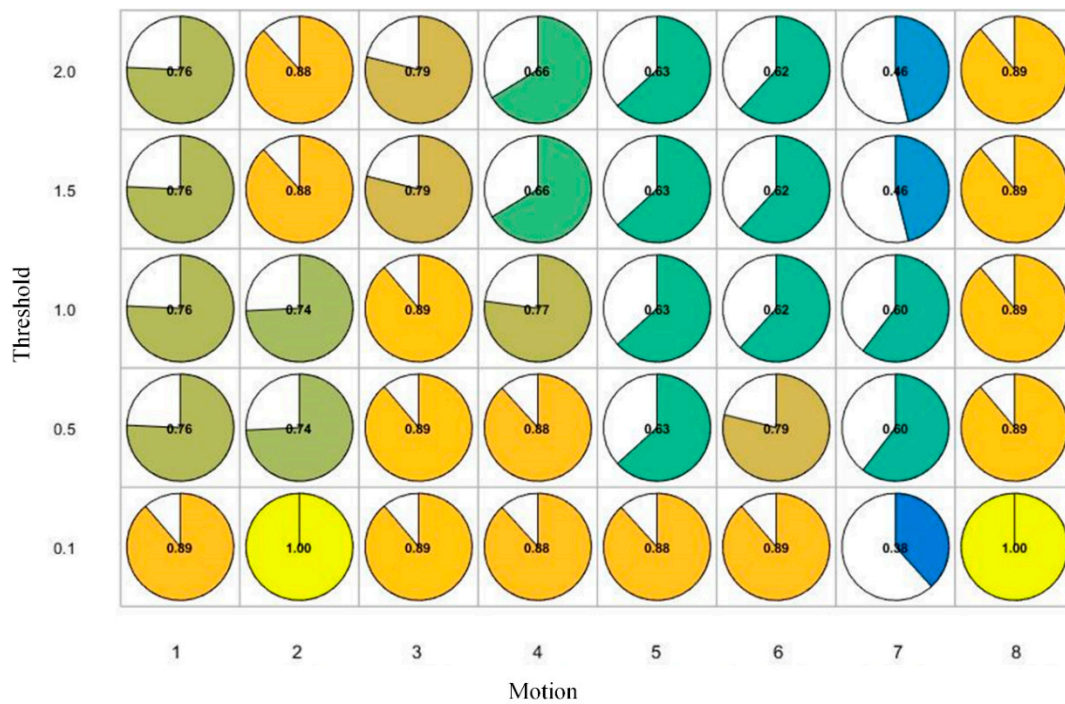


**Figure 7.** Differences in motion accuracy of points between novice and senior students on the key-frames of inter-frames pitch (Threshold = 1).

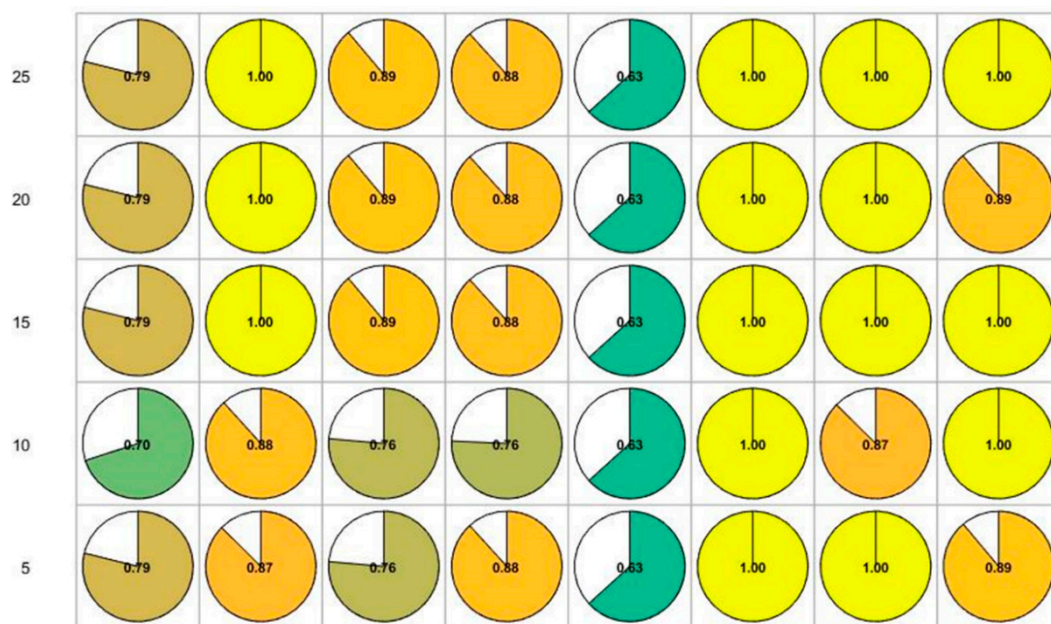
From Figures 6 and 7, we find that there was a difference between the results on the original frame and the key-frames on the inter-frame pitch. When there was a significant difference in motion accuracy between the two groups, we set the point to 1, otherwise, it was 0. Then, we evaluated the correlation between the results on the original frames and different key-frames on the Kendall correlation coefficient test (see Figures 8 and 9).

The results for key-frame extraction on inter-frame pitch show that when the threshold value was 0.1, the result of the differences in motion accuracy on the key-frames was highly correlated with the result based on the original frame (Kendall coefficient of points in each motion is higher than 0.7 except for Motion 7). However, when the threshold was 0.1, the compression rates of the key-frames were higher. As shown in Table 7, when the threshold was 0.1, the compression rate of each motion exceeded 50%. For key-frames extraction on clustering, there is a high correlation when the compression rate is 0.1. The Kendall coefficient of points in each motion is higher than 0.7 except for Motion 5, where the coefficient was 0.63.

We also tested the mean processing time for using DTW to calculate the distances between motions on original frames and key-frames (Table 11).



**Figure 8.** The Kendall coefficient of differences between skeleton points based on two difference methods (on the original frames and the key-frames on inter-frame pitch).



**Figure 9.** The Kendall coefficient of differences between skeleton points based on two difference methods (on the original frames and the key-frames on clustering).

**Table 11.** The mean processing time on the original frames and the key-frames.

Motions	The Mean Processing Time (s)		
	Original Frames	Key-Frames <sup>1</sup>	Key-Frames <sup>2</sup>
1	1.891	0.021	0.042
2	5.960	0.195	0.180
3	3.674	0.026	0.078
4	4.439	0.027	0.069
5	5.515	0.218	0.117
6	4.055	0.053	0.069
7	2.209	0.028	0.043
8	0.145	0.013	0.017

<sup>1</sup> Key-frames on inter-frames pitch (Threshold = 1); <sup>2</sup> Key-frames on clustering (compression rate = 15%).

From Table 11, the processing time on the key-frames is lower than original frames. Therefore, using key-frames can effectively decrease data processing time.

#### 4. Discussion

When using mathematical methods, the macro differences between the motion data of novice students and the teacher were higher than the distances between the motion data of senior students and the teacher on eight motions of Baduanjin. Because the motion data of the experimental analysis are the rotation data of specific skeleton points measured by the IMU, if the teacher's motions were taken as the standard, the results show that the motions of senior students were closer to the standard motions. Therefore, IMU can effectively distinguish the differences in motion accuracy in Baduanjin between novice and senior students.

When using the original frames to evaluate the differences at 17 skeleton points in eight motions between novice and senior students, the results show the differences in motion accuracy between the two groups on skeleton points varied for the different motions. For Motion 1, the differences between the two groups were mainly concentrated on the head-spine segment and upper limbs, especially the right upper limb. The differences mean that the motion errors of novice students relative to senior students were mainly concentrated on these joints. The results are consistent with the common motion errors described in the official book: "When holding the palms up, the head is not raised enough, or the arms are not raised enough" [30]. However, for Motion 4, the common motion errors are described in the official book as: "Rotating head and arm are insufficient" [30]. The description shows that the main errors occur in the head-spine and bilateral upper limbs. However, significant differences of skeleton points were at bilateral upper limbs but not head-spine. This difference may be related to the small number of participants in this study.

In this study, we also used two methods to extract key-frames. The raw data can be effectively compressed to decrease the data storage space using extracting key-frames [40,41]. The repetitiveness of action exercises in the teaching process will generate an extremely large amount of raw data. From the results, both key-frames extraction methods can effectively compress the raw data. We also found that the data processing speed could be accelerated on key-frames. However, the compression rates of key-frames on different motions when using key-frames on inter-frame pitch were different. We found that the differences in skeleton points on the key-frames on inter-frame pitch were not consistent with the results on the original frames. However, there was high consistency between the results on the key-frames on clustering and the results on the original frames, especially when the compression rate was 15%. Therefore, we can use key-frames to replace the original frames to evaluate motion accuracy of Baduanjin in order to decrease data storage space and processing time.

However, the small number of participants in our study limits the application of the results. As the participants were from a university in China, the results might only be suitable for university students



in China because different populations have variations in anatomical characteristics, physiological characteristics, and athletic ability.

Based on our results, IMU can effectively distinguish the difference in the motion accuracy of Baduanjin between novice and senior students. Therefore, in the following work, we can develop a system using IMU to evaluate the motion quality of students and provide feedback to teachers and students. Thus, it would be able to assist teachers in correcting errors in the motions of students immediately.

## 5. Conclusions

These initial results show that, based on the original frames, the IMU and the corresponding mathematical methods can effectively distinguish the motion accuracy of all eight motions of Baduanjin between novice and senior students. Furthermore, the IMU can identify the differences between the novice and senior students on the specific skeleton points of the eight motions of Baduanjin. The results regarding key-frames on clustering were highly correlated with the results of the original frames, which means, to a certain extent, that key-frames can replace the original frame to decrease the data storage space and processing time.

**Author Contributions:** Conceptualization, H.L, S.K., and H.J.Y.; Methodology, H.L, S.K., and H.J.Y.; Validation, H.L, S.K., and H.J.Y.; Data curation, H.L. and H.J.Y.; Writing—original draft preparation, H.L.; Writing—review and editing, S.K. and H.J.Y.; Supervision, S.K. and H.J.Y.; Funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Neijiang Normal University, Grant no: YLZY201912-1/-15, and the University of Malaya Geran Penyelidikan Fakulti, Grant no: GPF041A-2019.

**Acknowledgments:** We thank all participants and Neijiang Normal University for providing the facilities for the experiment.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. The Guiding Outline of Physical Education in National Colleges and Universities. Available online: [http://www.moe.gov.cn/s78/A10/moe\\_918/tnull\\_8465.html](http://www.moe.gov.cn/s78/A10/moe_918/tnull_8465.html) (accessed on 6 August 2020).
2. Several Opinions on Comprehensively Improving the Quality of Higher Education. Available online: [http://www.moe.gov.cn/srcsite/A08/s7056/201203/t20120316\\_146673.html](http://www.moe.gov.cn/srcsite/A08/s7056/201203/t20120316_146673.html) (accessed on 7 August 2020).
3. Healthy China 2030. Available online: [http://www.gov.cn/zhengce/2016-10/25/content\\_5124174.htm](http://www.gov.cn/zhengce/2016-10/25/content_5124174.htm) (accessed on 3 May 2020).
4. Li, F.Y.; Chen, X.H.; Liu, Y.C.; Zhang, Z.; GUO, L.; Huang, Z.L. Research progress on the teaching status of fitness Qigong Ba Duan Jin. *China Med. Herald* **2018**, *15*, 63–66. [CrossRef]
5. Yan, Y.C.; Xia, Z.Q. Present situation and development countermeasures of traditional Wushu in colleges and universities in Hubei province. *J. Jishou Univ. (Soc. Sci. Ed.)* **2013**, *34*, 167–169. [CrossRef]
6. Yang, J.Y.; Qiu, P.X. Introspection of Wushu elective course teaching content reform in Zhejiang University of Technology. *J. Phys. Ed.* **2015**, *22*, 92–97. [CrossRef]
7. Lu, C.Y.; Song, L.; Shan, Z.S.; Wang, Z.Q. Present situation of physical education teachers and sports equipments after college expansion in Shandong province. *J. Tianjin Univ. Sport* **2006**, *1*, 26. [CrossRef]
8. Notice of Provincial Education Department on the Results of the Fifth Batch of Public Physical Education Courses in Colleges and Universities. Available online: <https://wenku.baidu.com/view/bb4c0d8f876fb84ae45c3b3567ec102de2bddf83.html> (accessed on 7 August 2020).
9. Zhan, Y.Y. Exploring a New System of Martial Arts Teaching Content in Common Universities in Shanghai. Master's Thesis, East China Normal University, Shanghai, China, 2015.
10. Xavier, R.-L.; Hakim, M.; Christian, L.; André, P. Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis. *Med. Biol. Eng. Comput.* **2017**, *55*, 609–619. [CrossRef]

11. Pielka, M.; Janik, M.A.; Machnik, G.; Janik, P.; Polak, I.; Sobota, G.; Marszałek, W.; Wróbel, Z. A Rehabilitation System for Monitoring Torso Movements using an Inertial Sensor. In Proceedings of the Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), Poznan, Poland, 18–20 September 2019; pp. 158–163.
12. Yamaoka, K.; Uehara, M.; Shima, T.; Tamura, Y. Feedback of flying disc throw with Kinect and its evaluation. *Procedia Comput. Sci.* **2013**, *22*, 912–920. [[CrossRef](#)]
13. Chen, X.M.; Chen, Z.B.; Li, Y.; He, T.Y.; Hou, J.H.; Liu, S.; He, Y. ImmerTai: Immersive motion learning in VR environments. *J. Vis. Commun. Image Represent.* **2019**, *58*, 416–427. [[CrossRef](#)]
14. Elaoud, A.; Barhoumi, W.; Zagrouba, E.; Agrebi, B. Skeleton-based comparison of throwing motion for handball players. *J. Ambient Intell. Humaniz. Comput.* **2019**, *11*, 419–431. [[CrossRef](#)]
15. Thomsen, A.S.S.; Bach-Holm, D.; Kjørbo, H.; Højgaard-Olsen, K.; Subhi, Y.; Saleh, G.M.; Park, Y.S.; la Cour, M.; Konge, L. Operating room performance improves after proficiency-based virtual reality cataract surgery training. *Ophthalmology* **2017**, *124*, 524–531. [[CrossRef](#)]
16. Van der Kruk, E.; Reijne, M.M. Accuracy of human motion capture systems for sport applications; state-of-the-art review. *Eur. J. Sport Sci.* **2018**, *18*, 806–819. [[CrossRef](#)]
17. Corazza, S.; Mündermann, L.; Gambaretto, E.; Ferrigno, G.; Andriacchi, T.P. Markerless motion capture through visual hull, articulated icp and subject specific model generation. *Int. J. Comput. Vis.* **2010**, *87*, 156. [[CrossRef](#)]
18. Spörri, J.; Schiefermüller, C.; Müller, E. Collecting kinematic data on a ski track with optoelectronic stereophotogrammetry: A methodological study assessing the feasibility of bringing the biomechanics lab to the field. *PLoS ONE* **2016**, *11*. [[CrossRef](#)] [[PubMed](#)]
19. Panjkota, A.; Stancic, I.; Supuk, T. Outline of a Qualitative Analysis for the Human Motion in Case of Ergometer Rowing. In Proceedings of the 9th WSEAS International Conference on Simulation, Modelling and Optimization, Iwate Prefectural University, Iwate, Japan, 4–6 October 2009; pp. 182–186.
20. Schepers, H.M.; Veltink, P.H. Stochastic magnetic measurement model for relative position and orientation estimation. *Meas. Sci. Technol.* **2010**, *21*, 065801. [[CrossRef](#)]
21. Day, J.S.; Dumas, G.A.; Murdoch, D.J. Evaluation of a long-range transmitter for use with a magnetic tracking device in motion analysis. *J. Biomech.* **1998**, *31*, 957–961. [[CrossRef](#)]
22. Schuler, N.; Bey, M.; Shearn, J.; Butler, D. Evaluation of an electromagnetic position tracking device for measuring in vivo, dynamic joint kinematics. *J. Biomech.* **2005**, *38*, 2113–2117. [[CrossRef](#)] [[PubMed](#)]
23. Woodman, O.J. *An Introduction to Inertial Navigation*; Computer Laboratory, University of Cambridge: Cambridge, UK, 2007.
24. Poitras, I.; Biemann, M.; Campeau-Lecours, A.; Mercier, C.; Bouyer, L.J.; Roy, J.-S. Validity of Wearable Sensors at the Shoulder Joint: Combining Wireless Electromyography Sensors and Inertial Measurement Units to Perform Physical Workplace Assessments. *Sensors* **2019**, *19*, 1885. [[CrossRef](#)]
25. McGinnis, R.S. Advancing Applications of IMUs in Sports Training and Biomechanics. Ph.D. Thesis, University of Michigan, Ann Arbor, MI, USA, 2013.
26. Advanced Wireless Motion Capture System: Perception Neuron PRO. Available online: <https://www.noitom.com.cn/perception-neuron-pro.html> (accessed on 5 August 2020).
27. Myers, C.S.; Rabiner, L.R. A comparative study of several dynamic time warping algorithms for speech recognition. *Bell Labs Tech. J.* **2013**, *60*, 1389–1409. [[CrossRef](#)]
28. Carmona, J.M.; Climent, J. A Performance Evaluation of HMM and DTW for Gesture Recognition. In *Proceedings of CIARP 2012: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 236–243.
29. Badiane, M.; O'Reilly, M.; Cunningham, P. Kernel Methods for Time Series Classification and Regression. In Proceedings of the 26th AIAI Irish Conference on Artificial Intelligence and Cognitive Science, Trinity College Dublin, Dublin, Ireland, 6–7 December 2018.
30. Health Qigong Management Center of the State General Administration of Sports. *Health Qigong Baduanjin*; People's Sports Press: Beijing, China, 2018; Volume 3, pp. 215–217.
31. Perception Neuron 2.0. Available online: <https://www.noitom.com.cn/perception-neuron-2-0.html> (accessed on 20 August 2020).
32. Sers, R.; Forrester, S.E.; Moss, E.; Ward, S.; Zecca, M. Validity of the Perception Neuron inertial motion capture system for upper body motion analysis. *Measurement* **2019**, *149*, 107024. [[CrossRef](#)]

33. Dai, H.; Cai, B.; Song, J.; Zhang, D.Y. Skeletal Animation Based on BVH Motion Data. In Proceedings of the 2nd International Conference on Information Engineering and Computer Science, Wuhan, China, 25–26 December 2020; pp. 1–4.
34. Srivastava, R.; Sinha, P. Hand Movements and Gestures Characterization Using Quaternion Dynamic Time Warping Technique. *IEEE Sens. J.* **2016**, *16*, 1333–1341. [[CrossRef](#)]
35. Mukundan, R. Quaternions: From Classical Mechanics to Computer Graphics, and Beyond. In Proceedings of the 7th Asian Technology Conference in Mathematics, Melaka, Malaysia, 17–21 December 2002; pp. 97–105.
36. Kim, M.-H.; Chau, L.-P.; Siu, W.-C. Keyframe Selection for Motion Capture using Motion Activity Analysis. In Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Seoul, Korea, 20–23 May 2012; pp. 612–615.
37. Yan, C.; Qiang, W.; He, X.J. *Multimedia Analysis, Processing and Communications*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 535–542.
38. Li, S.Y.; Hou, J.; Gan, L.Y. Extraction of motion key-frame based on inter-frame pitch. *Comput. Eng.* **2015**, *41*, 242–247. [[CrossRef](#)]
39. Jablonski, B. Quaternion Dynamic Time Warping. *IEEE Trans. Signal Process.* **2012**, *60*, 1174–1183. [[CrossRef](#)]
40. Yan, Y. An Adaptive Key Frame Extraction Method Based on Quaternion. Master's Thesis, Xiamen University, Xiamen, Fujian, China, 2011.
41. Liu, X.M.; Hao, A.M.; Zhao, D. Optimization-based key frame extraction for motion capture animation. *Vis. Comput.* **2013**, *29*, 85–95. [[CrossRef](#)]
42. Zhang, Y.; Cao, J. 3D Human Motion Key-Frames Extraction Based on Asynchronous Learning Factor PSO. In Proceedings of the Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), Qinhuangdao, China, 18–20 September 2015; pp. 1617–1620.
43. Shi, X.B.; Liu, S.P.; Zhang, D.Y. Human action recognition method based on key frames. *J. Syst. Simul.* **2015**, *27*, 2401–2408. [[CrossRef](#)]
44. Sun, S.M.; Zhang, J.M.; Sun, C.M. Key Frame Extraction Based on Improved K-means Algorithm. *Comput. Eng.* **2012**, *38*, 169–172. [[CrossRef](#)]
45. Zhao, H.; Xuan, S.B. Optimization and Behavior Identification of Keyframes in Human Action Video. *J. Gr.* **2018**, *36*, 463–469. [[CrossRef](#)]
46. Alexiadis, D.S.; Daras, P. Quaternionic signal processing techniques for automatic evaluation of dance performances from MoCap data. *IEEE Trans. Multimed.* **2014**, *16*, 1391–1406. [[CrossRef](#)]
47. Keogh, E.; Ratanamahatana, C.A. Exact indexing of dynamic time warping. *Knowl. Inf. Syst.* **2005**, *7*, 358–386. [[CrossRef](#)]
48. Zhang, W.J.; Wang, J.J.; Zhang, X.; Zhang, K.; Ren, Y. A Novel Cardiac Arrhythmia Detection Method Relying on Improved DTW Method. In Proceedings of the IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 25–26 March 2017; pp. 862–867.
49. Chen, Q.; Hu, G.Y.; Gu, F.L.; Xiang, P. Learning Optimal Warping Window Size of DTW for Time Series Classification. In Proceedings of the 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA), Montreal, QC, Canada, 2–5 July 2012; pp. 1272–1277.

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).