

Article

Human Actions Analysis: Templates Generation, Matching and Visualization Applied to Motion Capture of Highly-Skilled Karate Athletes

Tomasz Hachaj ^{1,*} , Marcin Piekarczyk ¹ and Marek R. Ogiela ²

¹ Institute of Computer Science, Pedagogical University of Krakow, 2 Podchorążych Ave, 30-084 Krakow, Poland; marcin.piekarczyk@up.krakow.pl

² AGH University of Science and Technology, Cryptography and Cognitive Informatics Research Group, 30 Mickiewicza Ave, 30-059 Krakow, Poland; mogiela@agh.edu.pl

* Correspondence: tomekhachaj@o2.pl; Tel.: +48-12-662-7845

Received: 1 October 2017; Accepted: 7 November 2017; Published: 10 November 2017

Abstract: The aim of this paper is to propose and evaluate the novel method of template generation, matching, comparing and visualization applied to motion capture (kinematic) analysis. To evaluate our approach, we have used motion capture recordings (MoCap) of two highly-skilled black belt karate athletes consisting of 560 recordings of various karate techniques acquired with wearable sensors. We have evaluated the quality of generated templates; we have validated the matching algorithm that calculates similarities and differences between various MoCap data; and we have examined visualizations of important differences and similarities between MoCap data. We have concluded that our algorithms works the best when we are dealing with relatively short (2–4 s) actions that might be averaged and aligned with the dynamic time warping framework. In practice, the methodology is designed to optimize the performance of some full body techniques performed in various sport disciplines, for example combat sports and martial arts. We can also use this approach to generate templates or to compare the correct performance of techniques between various top sportsmen in order to generate a knowledge base of reference MoCap videos. The motion template generated by our method can be used for action recognition purposes. We have used the DTW classifier with angle-based features to classify various karate kicks. We have performed leave-one-out action recognition for the Shorin-ryu and Oyama karate master separately. In this case, 100% actions were correctly classified. In another experiment, we used templates generated from Oyama master recordings to classify Shorin-ryu master recordings and vice versa. In this experiment, the overall recognition rate was 94.2%, which is a very good result for this type of complex action.

Keywords: motion capture; signal averaging; template generation; kinematic; quaternions; karate; signal processing; dynamic time warping; barycenter averaging

1. Introduction

Human movements and actions analysis is a very interesting and up to date subject of wide research. This is mostly within the scope of interest of biomechanics, sport, medicine and rehabilitation. Human movements can be analyzed by various methods, for example by kinetic modeling [1–3], kinematic [3–8], electromyography [4,6,9–11], gait and posture analysis [12] and pattern recognition [13]. Quite often, those evaluations require statistical analysis of data from several measurements of the same action. Due to this, signal averaging and smoothing are common approaches. Those signal processing techniques are very often used in motion capture (MoCap) for missing sample complementing (estimating missing marker positions), averaging data from many captures of the same activity or removing the high frequency noises (smoothing) of the overall solution [14–19]. Averaging

of several performances of one action allows us to filter out small deviations that are not important in the whole movement. It can also be applied for the generation of action templates that can be compared with other signals of the same activity, in order to detect similarities and differences between them. Often, those types of analyses, for example in sport sciences, are modeled as one-dimensional movement [4,6,20–23] or two-dimensional path analysis problems [24].

There are already papers dealing with the application of DTW in motion capture data analysis and template generation. In [25], the authors generate motions templates (MT), which are a set of relational motion features that describe geometric relations between specified points of a pose. Those MT are then used in the classification process with DTW. After changing the DTW cost function, it might be adapted to operate on higher dimensional signals, for example quaternions [26]. The initial research on using quaternions and DTW for the full-body classification problem (without a description of how to create dedicated MT) is presented in [27]. DTW with quaternion-based distance and the nearest neighbor classification is used in the gait classification problem in [28,29]. In [30], an approach based on DTW and Hermite curve interpolation is used to normalize the length of motion sequences. In [31], the authors propose an algorithm that allows matching two discrete time functions, of which one is the original and the other can be considered a not so accurate copy with modified time, noise and a rotated coordinate system using quaternions, principal components analysis (PCA) and DTW. In [32], the authors use gyroscope sensors and various pattern classification techniques (among them DTW) to assign MoCap to one of the eight classes. In [33], the authors perform analysis of golf swings using a single motion sensor that is comprised of gyroscopes and accelerometers. Applying the PCA to the reference observations of properly performed swings, the PCA components of acceptable swing motion deviations are established. There are also many papers dealing with human action recognition using various acquisition methods and classification approaches, for example [34–36].

The direct inspiration for this research was the fact that to our best knowledge, there were no published papers that deal with the problem of the three-dimensional whole body MoCap data averaging that comes from multiple recordings of the same activity. Due to this fact, we did not find an appropriate approach to which to compare the proposed method. However, we have evaluated our approach in the role of template generation and matching for inertial sensor-based action recognition. What is more, our averaged data can be visualized using a kinematic model as the one presented in Figure 1 and presented directly to the expert. The aim of this paper is to introduce and evaluate the novel method of template generation, matching, comparing and visualization applied to motion capture (kinematic) analysis of motion capture data, which is a valuable extension of already published works [37]. The algorithm in [37] is not aware of quaternion dual mapping, and due to this fact, some three-dimensional rotations represented with quaternions could not be wrapped correctly. The algorithm proposed in this paper fixes this issue. Furthermore, this paper introduces mathematical fundamentals of the averaging and smoothing algorithm, while the previous one proposes only one possible implementation. We also present the methodology of DTW-matched results visualization in 3D (DTW signal mappings; see Algorithm 3, Figure 3), which is especially useful when one can manually rotate and observe this mapping from various angles. Furthermore, we introduce a very useful DTW signal distance plot (see Section 2.5 and Figure 4) that allows identifying the corresponding parts of actions that are the least and the most similar to each other. Those two methods of DTW algorithm results visualization are far more usable than the popular DTW signal alignment plot (see Figure 5) because our plots inform us about real distances between aligned samples rather than indices of alignment samples. Indices are not very relevant if we want to compare two actions using the DTW framework. The evaluation of the proposed methodology is performed on a very large dataset that contains 28 various karate techniques performed by two elite karate masters. Each technique is repeated 10 times: the dataset contains in total 560 recordings. Half of this dataset can be downloaded from the website of our project [38]. The rest of it will be published in the same place after finishing the ongoing research grant.

Our methodology can be nicely applied to martial arts action analysis, because the practice of combat training requires sequential repetition of prearranged movements until their performance becomes nearly automatic. Athletes optimize their skills by repeating the template movements presented by the coach. The computer-aided analysis of this process is one of the possible applications of the methodology presented in this paper. What is more important, combat sports and MoCap analysis are influential on athletes' health, injuries and rehabilitation [39,40]. Furthermore, data generated by our approach can be visualized in a virtual reality system that can be used as an innovative training procedure for athletes [41]. The proposed algorithms are composed of some established approaches; however, applying them according to our research idea resulted in valuable theoretical and practical results.

In the following sections, we will present the computational problem we are dealing with, the dataset and the proposed methodology.

2. Materials and Methods

In this section, we will describe our dataset, kinematic chain model and algorithms for template generation, matching, comparing and visualization applied to motion capture data.

2.1. Dataset and Kinematic Model

The MoCap data were recorded with a Shadow 2.0 wireless motion capture system. It consisted of 17 inertial measurement units that contain: 3-axis accelerometer, gyroscope and magnetometer. The information about range and resolution is presented in Table 1. The tracking frequency was set to 100 Hz with 0.5 degrees static accuracy and 2 degrees dynamic accuracy. This motion capture system has already been used and cited in scientific research; for example, [42,43]. Data measured by inertial sensors of our costume are processed by the software provided by the hardware manufacturer to calculate the above-mentioned MoCap signal values according to the hierarchical model in Figure 1. Each body joint of the model holds information about local rotations (orientation) of the so-called "bone" attached to it, relative to the parent joint. The root (hips) joint holds information about global rotation of the whole body. Those rotations are measured in Euler angles (this is the standard output format of the software). The calibration pose is a T-pose. In order to calculate the global reference frame, the MoCap system we used, similarly to other popular inertial solutions [44], acquires information regarding the location of the experiment and estimates the magnetic field of the Earth. These improve the determination of the global reference frame, enabling the correction of the angle between the gravity vector and magnetic field. After the calibration procedure, persons taking part in the recording might not face the same direction. As the body direction, we assume the rotation of the hip joint around the y (vertical) axis.

Table 1. This table presents information about the inertial sensors of our motion capture (MoCap) system.

Sensor	Full-Scale Range	Resolution
Accelerometer	$\pm 2, 4, 8, 16$ g	0.18 mg
Gyroscope	± 4000 °/s	0.12 °/s
Magnetometer	± 8 gauss	0.25 mG

Numerous studies have reported successful use of systems based on accelerometers or gyroscopes in sports and medical applications [45,46]. Among the biggest advantages of this type of solution is the portability of the solution and the possibility to use it in indoor and outdoor environment. What is more, the obtained data do not require further manual post-processing, which is quite common in visual MoCap solutions, for example due to marker occlusion and confusion. Research also proved that it is possible to obtain comparable accuracy of inertial sensors with respect to the optical MoCap

system [47,48]. The disadvantage is the lack of direct position tracking (this is calculated by double integration of accelerometer data) [49], which might cause some amount of drift. Despite this, in some circumstances, researchers prefer inertial technology over optical [50]. All of the mentioned factors might influence the final accuracy of tracking; however, the quality of our MoCap system seems to be suitable to evaluate our proposed methodology on karate data. Nevertheless, the proposed template generation algorithm works in the same way regardless of the motion capturing technique.

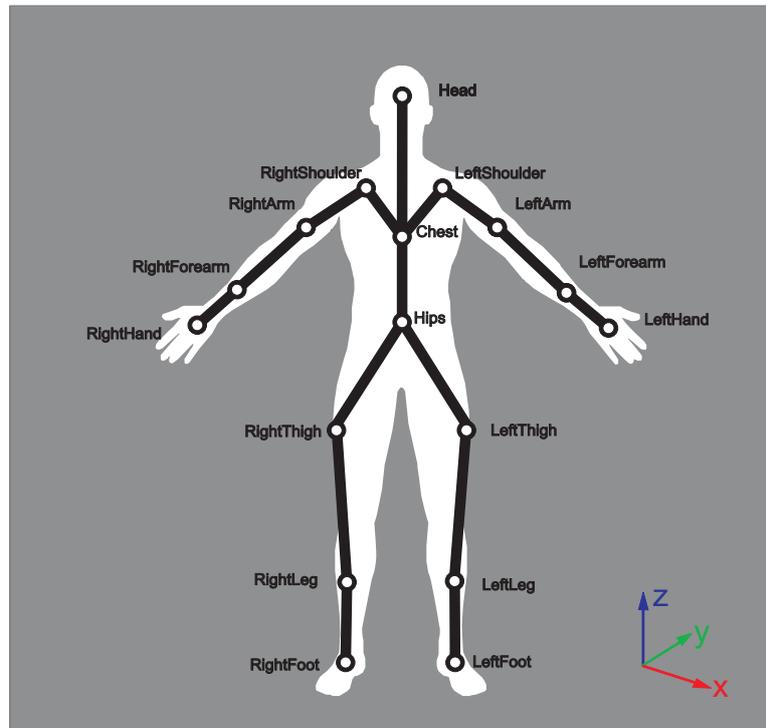


Figure 1. The hierarchical kinematic model that we used in our modeling. The root node of the model is the hips joint. The local coordinate system of each joint is right-handed. In the bottom left part of the image is the global right-handed coordinate system.

We have obtained data from two highly-skilled black belt karate athletes. One of them is a multiple Oyama karate knockdown champion, while the other is a Shorin-ryu kata champion. Both of them performed several karate techniques, and every technique was repeated 10 times. The Oyama karate master performed the following group of techniques:

- stances: kiba dachi, zenkutsu dachi, kokutsu dachi;
- blocks: gedan barai, jodan uke, soto uke, uchi uke;
- kicks: hiza geri, mae geri, mawashi geri, yoko geri;
- punches: furi uchi, shita uchi, tsuki.

All techniques were performed both on left and right side (leg or hand), giving in total 28 techniques.

The Shorin-ryu karate master performed the following group of techniques:

- stances: naihanchi dachi, shiko dachi, zenkutsu dachi, uki ashi dachi;
- blocks: age uke, soto uke, uchi uke, gedan barai;
- elbow strikes: age empi, mae empi;
- punches: choku tsuki;
- kicks: hiza geri, mae geri, mawashi geri, yoko geri.

The detailed movement descriptions with illustrations can be found, for example, in [51].

All techniques besides the fully-symmetrical ones like *naihanchi dachi* and *shiko dachi* were performed both on the left and right side (leg or hand), giving in total 28 techniques. All recordings are from 2–4 s in length. We have performed sensor calibration before each motion capture session using software provided by the hardware manufacturer. Because we acquired 10 samples of each karate technique of both athletes, we finally gathered 560 recordings. We have chosen those particular techniques to be recorded because they are commonly used during karate training. To precisely describe the movements, we used the kinematic chain model (see Figure 1), the root of which is the hips joint. The descriptions below show where sensors are placed on the subject.

- Head – located approximately between left and right temple, over eyebrows.
- Right/left shoulder: placed in the middle of the left scapula.
- Right/left arm: placed between the acromioclavicular joint and lateral epicondyle in a proportion 1/3–2/3 closer to the shoulder.
- Right/left forearm: placed between the lateral epicondyle and wrist in a proportion 1/3–2/3 closer to wrist.
- Right/left hand: approximately over the capitate carpal bone and third metacarpal bone.
- Chest: approximately on the xiphoid process of the sternum.
- Hips: between the left and right anterior superior iliac spine.
- Right/left thigh: over the lower lateral 1/3 surface of the thigh, just below the swing of the hand.
- Right/left leg: placed over the lower 1/3 of the shank.
- Right/left foot: placed on tarsal bones superior.

We utilize only angle-based notation of the chain to make the motion model invariant to body proportions. The algorithm does not use dual quaternions because we do not apply translations in our model. Ideally, all of the considered actions that are performed properly would be performed entirely identically, and the analyzed data would be equal among all observations [33]; however, there are always some deviations. The averaging was performed only on recordings that consisted of correctly performed actions. The decision whether the action was correct or not has been made by an expert.

2.2. Rotation Signal Averaging

MoCap data acquired by our system are a set of time-varying signals. We have recalculated Euler angle representation to quaternions. To average a set of time-varying signals, we have used the dynamic time warping barycenter averaging (DBA) heuristic algorithm [52] that exploits the DTW similarity measure. The cost function in DTW for quaternions is defined as [28]:

$$cf(x, y) = 1 - |x \circ y| \quad (1)$$

where x, y are normalized quaternions, \circ represents a dot product and $||$ is an absolute value. The absolute value is necessary because the unit quaternions q and $-q$ correspond to a single 3D rotation.

There are several other functions that have been proposed to measure distance in a quaternion space, for example:

- the length of a geodesic line connecting two quaternions [26]:

$$dg(x, y) = 2 \cdot \min\left(\left\| \log(x^{-1} \cdot y) \right\|, \left\| \log(y^{-1} \cdot x) \right\| \right) \quad (2)$$

where $|||$ is a quaternion norm.

- the distance in tangent space, into which quaternions are transformed by the logarithm operator [28]:

$$dtan(x, y) = \|\log(x) - \log(y)\| \quad (3)$$

- corresponding to rotation angle concatenation of rotations q_1 and inverse rotation to q_2 [28]:

$$drot(x, y) = re(x \cdot \bar{y}) \quad (4)$$

where \bar{y} is a conjugate quaternion.

Based on the experiment presented in [28], where the authors evaluated various distance functions in the task of gait identification based on the nearest neighbor classification technique with motion similarity assessment by dynamic time warping, we have chosen Function (1), which gave the best results while using all joints and was the least numerically complicated. Furthermore, visual evaluation of the initial results of application (2)–(4) to DTW done by an expert did not show superiority of any of them over (1).

Because we are dealing with quaternions, the barycenter averaging is replaced by Markley's algorithm, which is a norm-preserving approach [53]. Markley's algorithm is defined as:

$$M = \begin{cases} M_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ M_i = M_{i-1} + w_i \cdot q_i \cdot q_i^T \end{cases} \quad (5)$$

$$\bar{q} = \arg \max_{q \in S^3} q^T \cdot M \cdot q \quad (6)$$

where S^3 denotes the unit 3 sphere, q_i the i -th quaternion to be averaged and w_i the i -th weight from weight vector. The average quaternion \bar{q} is the eigenvector of M corresponding to the maximum eigenvalue.

We use the DTW approach because the activities we want to compare might be performed at different speeds and there might be nonlinear scaling differences between signals. For example, in the case of mae geri kick, the proportion of time between lifting a foot from the ground with simultaneous knee joint flexion in the first part of a kick and knee joint extension in the second part of the kick is not constant and might differ between recordings; see how this phenomenon affects the knee joint rotation in Figure 2. That means that simple resampling of all signals to have the same length (number of samples) and performing, for example, Markley's algorithm directly on them will not result in valuable averaging results.

The proposed averaging algorithm is presented as pseudocode in Algorithm 1.

As we already mentioned, it is possible that persons in each pair of MoCap face different directions. As the body direction, we assume the rotation of the hip joint around the y (vertical) axis. In the hierarchical kinematic model, we can make all persons face the same direction by rotating only the root (hips) joint using the following approach.

Let us assume that we want a person to face direction:

$$X \leftarrow [1, 0, 0] \quad (7)$$

and we want to apply a rotation only by the y (vertical) axis. The problem we have to solve is to find quaternion φ that minimizes the following formula:

$$\min_{\varphi} (R(\varphi, R(\alpha, X)) \circ X) \quad (8)$$

where $R(\beta, Y)$ is a rotation transform of vector Y by quaternion β and α is an initial rotation of X . Equation (8) can be solved relatively fast by the brute-force search approach.

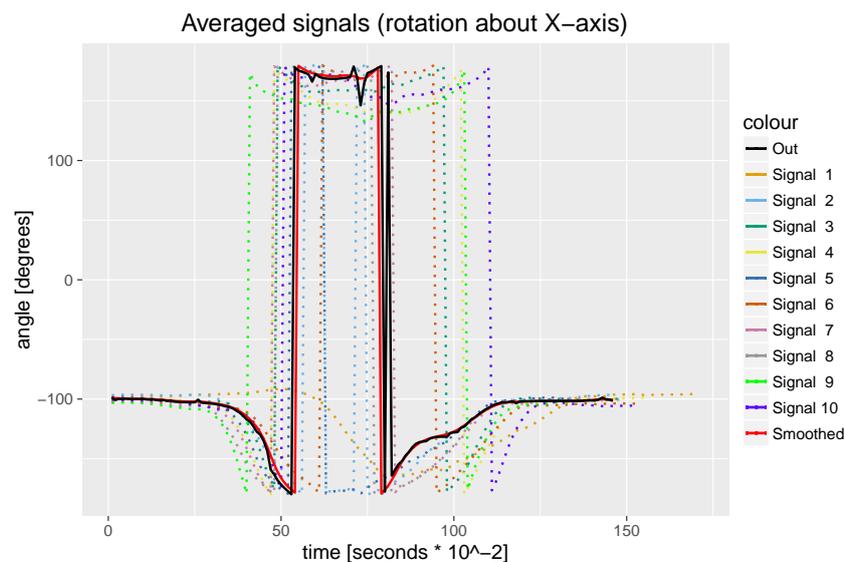


Figure 2. The results of the signal averaging and smoothing algorithm. The plot shows the rotation of hips joint about the X-axis during mae geri kick with left leg. Dotted plots named Signal 1–Signal 10 are the input data. The solid black line (out) is a result of signal averaging, while the solid red line is the smoothed averaged signal.

Algorithm 1: Signal averaging algorithm.

Data: A – initially averaged sequence of size L; S – matrix of sequences, each sequence has length T

Associations := empty list of size L;

for each seq in S do

```

| [Cost, Path] := DTW (A, seq); // DTW algorithm returns Cost matrix and Path
| matrix, cf function is defined in equation (1)
| i := L;
| j := T;
| while i >= 1 and j >= 1 do
| | Associations[i] := Associations[i] sum seq[j]; // add j-th quaternion from seq to
| | Associations
| | [i, j] := Path[i,j];
| end
| end
| for i in 1:T do
| | Avg[i] := Quaternion averaging(Associations[i]); // Markley's algorithm
| end
| return Avg;

```

2.3. Signal Smoothing

The DBA algorithm might introduce in averaged signals high frequency noises that are visible as rapid Euler angle hops in Figure 2 (see the black “out” signal). This is of course the result of the DBA heuristic, which does not prevent these situations, even if input data do not contain that type of noise. Our proposed smoothing algorithm, which is an extension of [37], works similarly to the typical discrete linear convolution algorithm with the Gaussian kernel; however, instead of the linear combination of signal samples in the kernel window and kernel weights, we use once again Markley’s algorithm. The smoothing algorithm of discrete quaternion signal Q is defined as:

$$Qs_j = \arg \max_{q \in S^3} q^T \cdot M_{\{j-ws, \dots, j+ws\}}^G \cdot q \quad (9)$$

where Qs_j is the j -th value of smoothed Q , G is a symmetric vector with a Gaussian distribution of values with length $ws + 1$ and $M_{j-ws, \dots, j+ws}^G$ is Matrix (2) constructed from $ws + 1$ quaternions of Q width indexes $\{j - ws, \dots, j + ws\}$ and weight vector G . The smoothing results are presented in Figure 2 (see the red “smoothed” signal).

The proposed averaging algorithm is presented as pseudocode in Algorithm 2.

Algorithm 2: Smoothing algorithm.

```

Data: Q – quaternion signal of length T which we want to smooth; ws – length of kernel
Qs := Q; // rewrite signal to sequence that hold output data
startInd := floor(ws / 2);
startLoop := startInd;
endInd := ceiling(ws / 2);
endLoop := T - endInd;
G := symmetric vector with Gaussian distribution of values with length ws;
for a in startLoop:endLoop do
  i := 1;
  sampleToSmooth := vector with length ws;
  for b in (-startInd+1):endInd do
    sampleToSmooth[i] := Q[a+b];
    i := i + 1;
  end
  Qs[a] := Quaternion averaging(sampleToSmooth, G);
end
return Qs;

```

2.4. Full-Body Template Generation

To generate the full-body template of an action, we use all MoCap recordings of the given type as input data. The Euler angles are recalculated to quaternions (both global rotation hips joint angle and all local angles). This step removes problems caused by discontinuity in rotation descriptions in the Euler angle domain $[-180, 180)$ for rotation around the X axis, $[-90, 90)$ for Y and $[-180, 180)$ for Z. In the situation when we are using Euler angles to describe joint motion in the domain we mentioned earlier, there might be an event while performing an action when a “bone” rotation around an axis causes the discontinuity of the rotation plot (see the dotted lines in Figure 2 where if the angle is less than -180 degrees, it is flipped to be close to, but not to overcome 180 degrees). These discontinuities might not appear in all signals that represent the same action; for example, Signal 1 does not have it; however, the joint rotation driven by it does not differ much from other signals. Those discontinuities that are present in the Euler angle domain might cause averaging algorithm failure. To overcome this, we decided to use quaternions, where this phenomenon is absent. What is more, quaternions prevent the gimbal lock phenomena and guarantee that the averaged signal retains all degrees of freedom. We apply (8) to correct the rotation of hips joints and then Algorithms 1 and 2 to each quaternion signal separately. The averaged and smoothed quaternions are recalculated to Euler angles once again and integrated into single MoCap. This last quaternion to Euler angle conversion is done to facilitate visual evaluation of the results. The computational complexity of our algorithm (ALG) equals:

$$O(ALG) = O(SN \cdot Iter \cdot O(DTW)) = O(SN \cdot Iter \cdot Len^2)$$

where: SN is the number of sequences; Iter is the number of iterations; Len is the signal length. As can be seen, the computational complexity is polynomial, and the algorithm can operate in real time. By “real time” in this context, we mean that: as the algorithm has a polynomial complexity, it can be

applied to a given time window, and the averaged signal might be computed in time comparable to the acquisition frequency of typical professional MoCap hardware. DTW commonly uses a sample of a certain length to calculate the warping path. Because of that, in the typical scenario, DBA, which uses DTW, will not run on successive samples, but rather on selected parts of previously acquired data.

2.5. Signal Matching and Visual Analysis

To compare two actions on the level of a single joint, we will use a procedure that generates the three-dimensional plot of the DTW signal mapping using Measure (1). We use this approach because the activities we want to compare might be performed with different speeds, and there might be nonlinear scaling differences between signals. To solve the problem with various body proportions, we represent the movement only with quaternion angles. The generation of the DTW signal mapping plot is presented in pseudocode in Algorithm 3. This time however, we do not need cost and path matrices, but matching vectors of Signal 1 onto Signal 2 and vice versa, which are also generated by DTW. An example plot of this type is shown in Figure 3.

Algorithm 3: DTW signal mapping plot generation algorithm.

```

Data: singal1, signal2 – quaternion signals to be mapped (compared)
[match1, match2] = DTW(singal1, signal2); // DTW distance with cost function (1),
match1 - matching of signal1 onto signal2, match2 - matching of signal2 onto
signal1
X := [0, -1, 0];
N := length(signal1);
M := length(signal2);
K <- length(match1);
for a in 1:N do
| blue_signal[a] := R(signal1[a], X);
end
for a in 1:M do
| green_signal[a] := R(signal2[a], X);
end
for a in 1:K do
| red_sections[a] := [blue_signal[match1[a]], green_signal[match1[a]]];
end
return blue_signal, green_signal, red_sections;

```

To present distances between two DTW matched signals, one can also plot the following two-dimensional drawing composed of the following components:

- The first signal is represented as a straight, vertical line with n sampled points marked on it (n is the signal length).
- The second signal is plotted over the first one. It consisted of m sampled points (m is the signal length). The vertical coordinates of each point are proportional to the DTW distance between this point and the corresponding point of the first signal.
- It has sections that show DTW mapping (corresponding points according to DTW) between both signals.

An example plot of this type is shown in Figure 4.

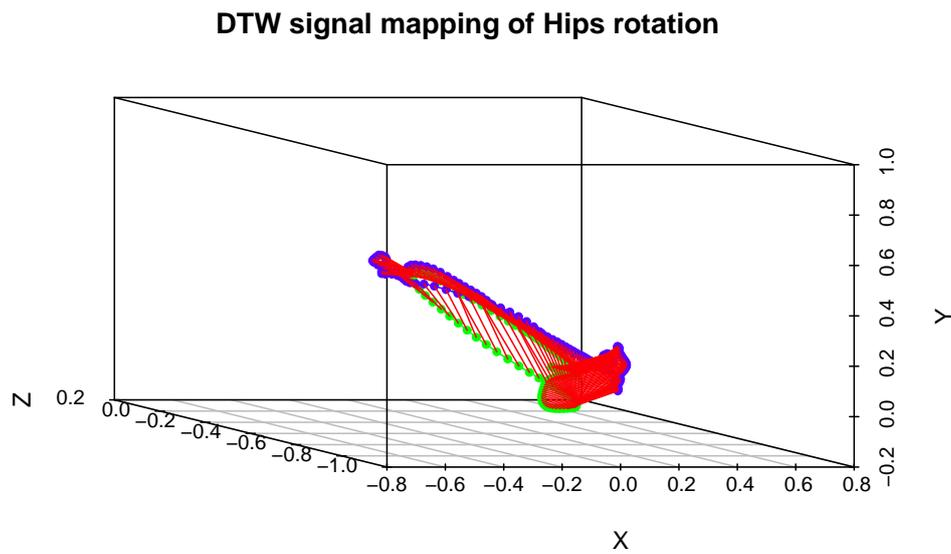


Figure 3. DTW signal mapping of hips rotation between template (green signal) and one selected MoCap recording (blue signal) of mae geri kick with the left leg. Those are the same MoCap as in Figure 5. Distances are presented as red sections.

Distances calculated in DTW

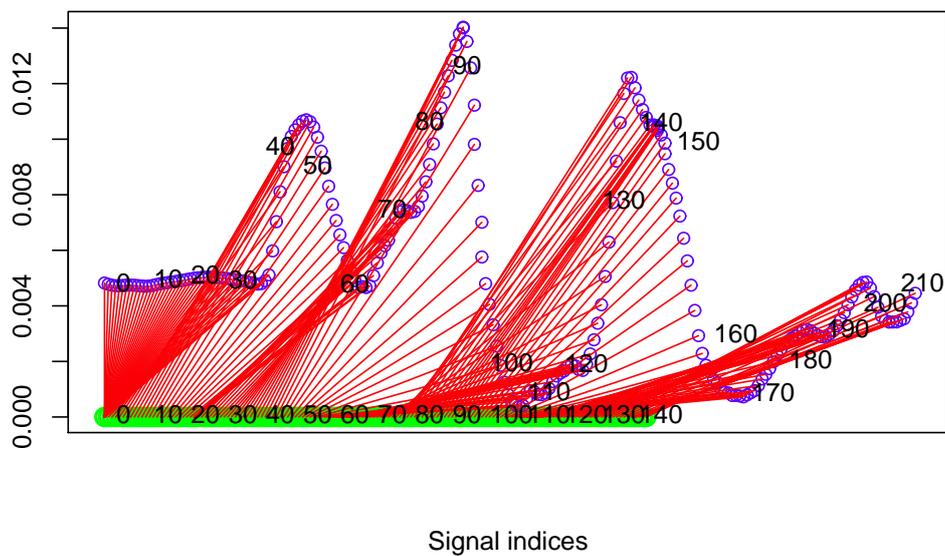


Figure 4. Distance of hips rotation between template (green signal) and one selected MoCap recording (blue signal) of mae geri kick with the left leg. Those are the same MoCap as in Figure 5. Distances are presented as red sections.

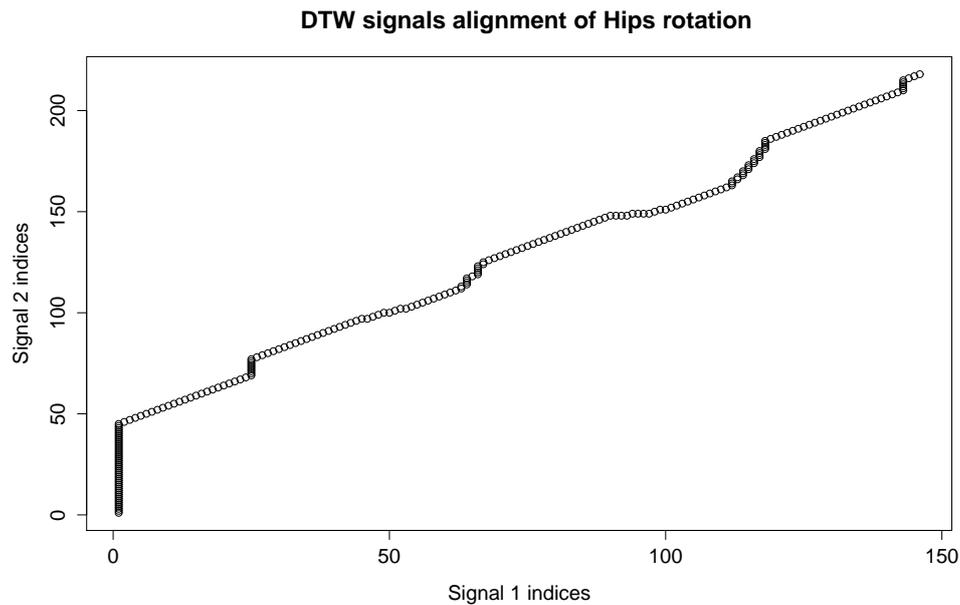


Figure 5. DTW signal alignment of hips rotation between the template (Signal 1) and one selected MoCap recording of mae geri kick with the left leg (Signal 2). The template was generated from the Shorin-ryu karate master data, while the other MoCap came from Oyama karate master.

3. Results

The basis of our averaging method is the DBA algorithm, an averaging method which iteratively refills an initially (potentially arbitrary) selected sequence, in order to minimize its squared distance (DTW) to averaged sequences. As it is a heuristic, it does not guarantee that the obtained solution will be optimal. Our experiments are designed to evaluate if our approach can be utilized for some specialized class of time-varying signals, namely multidimensional MoCap recordings. As our dataset is large (560 recording) and contains a wide range of various techniques (28 classes), we believe that this sample is representative to prove the validity of our methodology. We have prepared and performed several experiments designed to examine the following key aspects of our methodology:

1. Evaluating the quality of generated templates: By quality, we mean the averaging potential of the proposed method and the correspondence of template 3D rendering to the original MoCap recordings. That means that the template has to be not only similar to the input data in the numerical sense, but it also has to present the correct karate technique.
2. Evaluating the potential of the matching algorithm to calculate similarities and differences between various MoCap data.
3. Applying generated templates to human action classification.
4. Evaluating visualizations of important difference and similarities between MoCap data detected by the previous methodology.

The first experiment was performed on the whole dataset introduced in Section 2.1. We have generated templates of all techniques performed by Oyama and Shorin-ryu karate masters separately. Due to the paper space limitation, we cannot present complete evaluation results, so we decided to show them aggregated in the form of various plots.

Renderings of MoCap templates we obtained show the correct performance of all techniques. In the Supplementary Materials, we present two example renderings of MoCap templates. The renderings are visualizations of bvh (Biovision Hierarchy) files we generated with our algorithm. Visualizations were done in Blender 2.69 and played four-times slower than the original data. Ten figures in the rear row are the input (original) Mocap recordings, while the figure in front of them is

the result of averaging. The first visualization contains mae geri kick with the left leg of the Shorin-ryu karate master, while the second mawashi geri kick with the left leg of the Oyama karate master.

In Figure 2, we present the example averaging and smoothing calculated by our method. Figures 6 and 7 summarize our experiment and show bar plots of median values of dynamic time warping normalized distances calculated for each karate technique. The smaller the values of normalized distances, the more similar is a generated template to the recordings from which it was calculated. Due to the large variance of results caused by joints that are not relevant for a given action (for example, hands during kicks), we use the median value, which discriminates outlier values, instead of the average value. The differences in median DTW normalized distances presented in Figures 6 and 7 show the general intuitive rule that a higher median occurs in moves that are more complex and requires larger mobility of the body. For example, it is easier to performer a similar stance than to perform a similar kick or punch.

The more detailed results of this evaluation are presented in Figure 8. It shows in the form of a matrix dynamic time warping normalized distances between template and input signals of mae geri front kick with the left leg of the Shorin-ryu karate master. This matrix demonstrates the results of averaging. The bigger the difference is between the particular signal and template, the higher DTWND (dynamic time warping normalized distances) we obtain. The relatively high DTWND values might be present on the map because of two cases. In Figure 8, we see only relatively high differences in left thigh joints in comparison to other joints. This means that movements of the rest of the body are relatively similar between cases, and differences are present because, as we already mentioned, in practice, it is impossible to perform some action ideally; there are always deviations. The highest deviations happened in the joints that required the highest mobility. The second reason for the differences occurs when the matrix visualizes data from two different actions. If one action is a correct template and the other action we want to examine is not, the matrix indicates in which part of the body actions differ the most. The body joints discovered that way should be the target of further analysis, for example by Algorithm 3; we will discuss this case later.

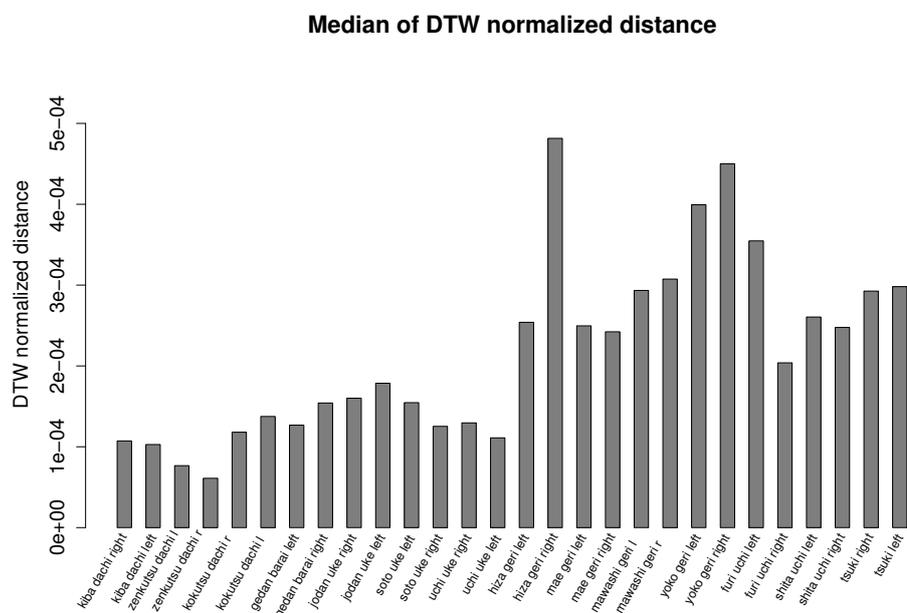


Figure 6. Bar plot of the median value of dynamic time warping normalized distances calculated for each karate technique of the Oyama karate master. At first, the dynamic time warping normalized distance between template and input signals is calculated, then we take the median value of all calculated distances.

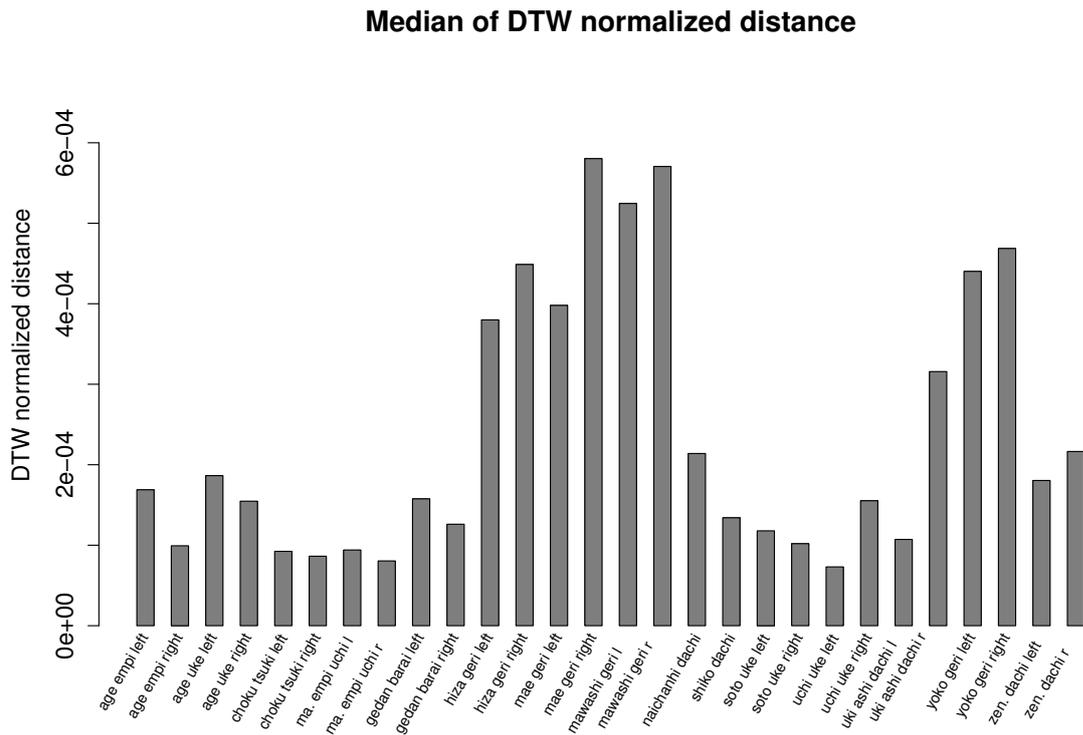


Figure 7. Bar plot of the median value of dynamic time warping normalized distances calculated for each karate technique of the Shorin-ryu karate master. At first, the dynamic time warping normalized distance between template and input signals is calculated, then we take the median value of all calculated distances.

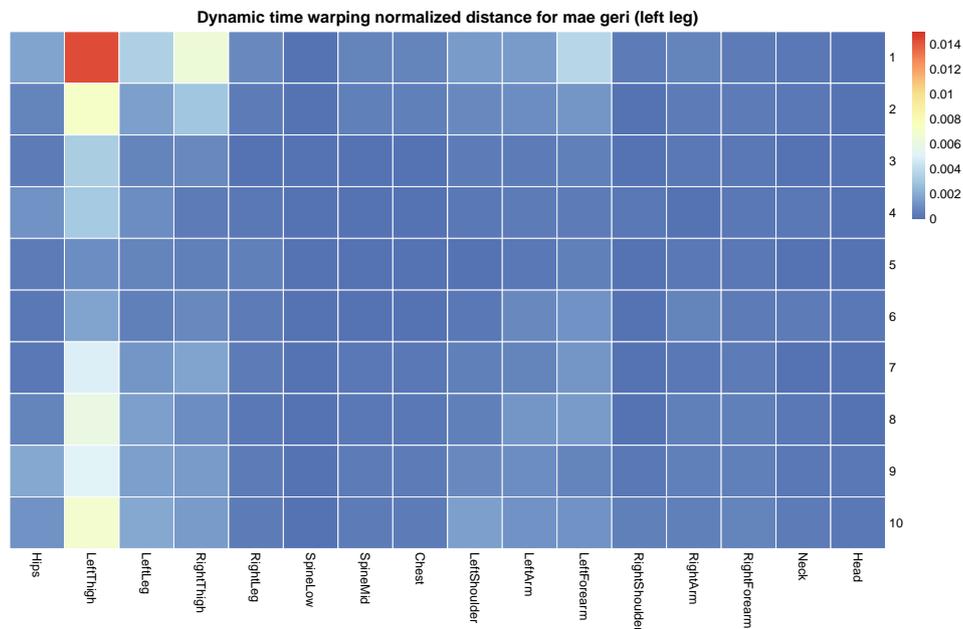


Figure 8. Dynamic time warping normalized distances between the template and input signals of mae geri with the left leg of the Shorin-ryu karate master. Each signal is presented in different rows, while each column is for different joint. Distances are presented on the matrix with colors. Each row represents a single take of the same technique.

The aim of the second experiment was to compare a template generated with our approach from actions of one person with actions that were performed by another person. We took the template of the Shorin-ryu karate master that performed mae geri kick with the left leg standing in zenkutsu dachi stance. We compare it to mae geri performed with the left leg from kumite-no-kamae stance of the Oyama master (the initial stances differ from each other). We decided to choose this particular technique because it characterizes not only high dynamics, like all karate attacks, but also it has a large radius of leg movement. What is more, the height of the kick might differ between even the most skilled athletes because it is conditioned by the flexibility of particular persons and their anatomy. Due to this, the proposed evaluation will examine not only the possibility of aligning very similar MoCap data, but also the possibility of dealing with various differences that are present on recordings of people performing the same body action.

Figure 9 presents the comparison of those two MoCap data we mentioned earlier in the form of a matrix. We skipped signals that are irrelevant for comparison, like head and hand positions. In the next Figures 3–5, we evaluate the DTW aligning and mapping (with Algorithm 3), and we plot the DTW distance of the hips joint.

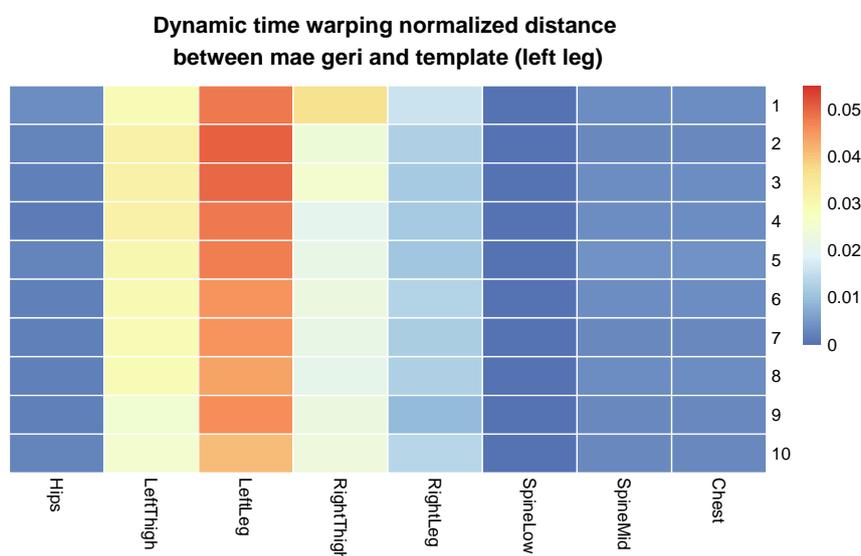


Figure 9. Dynamic time warping normalized distances between the template and input signals of mae geri with the left leg. The template was generated from the Shorin-ryu karate master data, while other MoCap came from the Oyama karate master. Distances are presented on the matrix with colors. Each row represents a single take of the same technique.

We have used data visualized in Figure 4 to find MoCap frames that represent the most distant (local maxima) and least distant (local minima) corresponding points. They allow us to identify the corresponding parts of actions that are least and most similar to each other.

The motion template generated by our method can be used for action recognition purposes. The goal of human action recognition is to assign the recording presenting a person performing some activity to one of the predefined classes. There are many methods to register human movements for example by depth cameras and inertial sensors [34,54,55], and the registration method partially determines the possible classification algorithm. Because we generate a single action template that is averaged from several recordings, the straightforward approach that can be used to classify an action is the DTW classifier [55]. This type of classifier contains templates of actions and compares them with recording to be recognized. This recording will be assigned to the class, to which the corresponding template has the smallest DTW normalized distance value. In order to make the classification effective, we have to select the appropriate feature set. Based on our experience with

karate action recognition [56,57], we will use angle-based features between the axis of the coordinate system that is relative to the body position and vectors designed by coordinates of body joints. As we already mentioned in [57], when we are dealing with whole body action recognition of very similar actions that differ only by some nuance (like various types of kicks or punches), it is better to use different feature sets for various action groups. That is because when we use too many feature, there will be too many constraints for actions to be correctly classified. On the other hand, where the number of features are too small, there will not be a possibility to distinguish between similar actions (like for example various karate kicks). Because of those factors, it is very difficult to create a single feature set that can be successfully used to classify various movement tasks. In the rest of this paper, we will consider a feature set that can be used to classify karate kicks (lower part of the body).

The unnormalized vectors of the coordinate system are defined as:

$$\begin{cases} \bar{X} = LeftThigh - RightThigh \\ \bar{Y} = Chest - Hips \\ \bar{Z} = X \times Y \end{cases} \quad (10)$$

where the names of the vectors correspond to the names in Figure 1 and \times is a cross-product. The three-dimensional coordinates of the features have been recalculated from the hierarchical to the kinematic model. We also use four vectors defined as:

$$\begin{cases} \bar{V1} = LeftLeg - LeftThigh \\ \bar{V2} = LeftLeg - LeftFoot \\ \bar{V3} = RightLeg - RightThigh \\ \bar{V4} = RightLeg - RightFoot \end{cases} \quad (11)$$

We have used twelve angle-based features that are angles on the plane:

$$\begin{cases} F1 = \angle(\bar{X}, \bar{V1}); F2 = \angle(\bar{Y}, \bar{V1}); F3 = \angle(\bar{Z}, \bar{V1}) \\ F4 = \angle(\bar{X}, \bar{V2}); F5 = \angle(\bar{Y}, \bar{V2}); F6 = \angle(\bar{Z}, \bar{V2}) \\ F7 = \angle(\bar{X}, \bar{V3}); F8 = \angle(\bar{Y}, \bar{V3}); F9 = \angle(\bar{Z}, \bar{V3}) \\ F10 = \angle(\bar{X}, \bar{V4}); F11 = \angle(\bar{Y}, \bar{V4}); F12 = \angle(\bar{Z}, \bar{V4}) \end{cases} \quad (12)$$

The distance function for DTW is defined as the Euclidean distance between the vector of features $[F1, F2, F3, F4, F5, F6, F7, F8, F9, F10, F11, F12]$. Using this features set, we have performed leave-one-out action recognition for the Shorin-ryu and Oyama karate master separately. We have taken into account mawashi geri, mae geri, hiza geri and yoko geri (both with the right and left leg), so we have eight classes of actions. Templates were generated from nine motion recordings, and the remaining one was classified using the template. In this experiment, each action was correctly assigned to its class. In another experiment, we used templates generated from Oyama master recordings to classify the Shorin-ryu master recordings and vice-versa. Because the Oyama master performed yoko geri kicking sideways, while the Shorin-ryu master kicked in front, those two motions could not be compared because they create different templates. For both athletes, all mawashi geri (with right and left leg) and all but one mae geri kick were correctly classified. One mae geri kick with the left leg was assigned to the hiza geri left class. Three hiza geri kicks with the right and three with the left leg of the Shorin-ryu master were classified as mae geri kicks. The overall recognition rate was 94.2%, which is a very good result for this type of complex action. The mae geri and hiza geri kicks can be misclassified because hiza geri (knee strike) and mae geri (straight kick) have nearly the same thigh trajectory, and also shin movement at the beginning and end of those actions is very similar. The detailed results are presented in Table 2.

The angle-based feature setup we use in this experiment is often chosen by other researches (see for example the state of the art presented in [13]). Although using quaternions is among the most convenient ways for processing three-dimensional rotations in advanced signal processing (we

explained this in Section 2.4), in the case of actions, recognition features based on Euler angles might be a better choice due to several reasons. At first, in many cases, to successfully perform action recognition, we do not need to use features that represent complete motion of the whole body (like 3D rotations in quaternion form); rather some motion derivatives that are sufficient for the action class we want to classify. For example, for classification purposes, we can replace four-dimensional quaternions by three-dimensional Euler angles on the planes between vectors of the coordinate frame and vectors representing limbs. If we decide to use this simplified approach, the feature set for each body joint has one dimension less compared to quaternion notation (which is four-dimensional). In our case, for example, instead of representing each sample with a 16-dimensional vector using quaternions, we have 12-dimensional vectors of Euler angles. This feature definition makes classification less constrained. The second reason is that results obtained with angles on the plane are easier to interpret than quaternions, because when we calculate angles between vectors using the dot product, they are in domain $[0, 180]$. Furthermore, various classes of actions described by angle-based features can be nicely visualize in three-dimensional space using PCA (see for example [57]).

Table 2. Classification results of the DTW classifier on various karate kicks. Each row represents the correct action class; columns represent classification results. Because there are 10 recordings of each class for the Oyama and Shorin ryu masters, each class has together 20 motions recordings.

	Mawashi Geri Right	Mawashi Geri Left	Mae Geri Right	Mae Geri Left	Hiza Geri Right	Hiza Geri Left
Mawashi Geri Right	20					
Mawashi Geri Left		20				
Mae Geri Right			20			
Mae Geri Left				19		1
Hiza Geri Right			3		17	
Hiza Geri Left				3		17

The last experiment evaluates how the proposed matching method behaves when we use it to compare two different techniques that have however some similarities. To complete the tasks, we have chosen the same mae geri (front kick) template as in the previous experiment; however, we compared it with mawashi geri (roundhouse kick) of the Oyama karate master. It is obvious that the front kick has a different hips trajectory than the roundhouse kick; however, the knee joint (relative to the parent joint) rotates in a similar manner. The evaluation was performed in the same manner as in the previous experiment, and results are visualized in Figures 10–13.

The matrix in Figure 10 visualizes the essential differences between two correctly-performed karate techniques: the front kick and the roundhouse kick. While the position of the upper parts of the body remain relatively similar (spine low, spine mid and chest joint), the highest differences are in the movements of the left and right thighs (thigh rotation propels the impact of the kick). What is more, in the case of the roundhouse kick, the athlete also twists the leg that is not used for kicking (in the case the left leg kick, it is a right leg) to increase the range of thigh rotation. The presence of those rules is clearly visible on the obtained matrix. The lack of them would mean that one of those kicks is not performed correctly. For example, insufficient body rotation in the roundhouse kick is a common problem of the young karate students. Without it, the roundhouse kick becomes too similar to the front kick and thus ineffective.

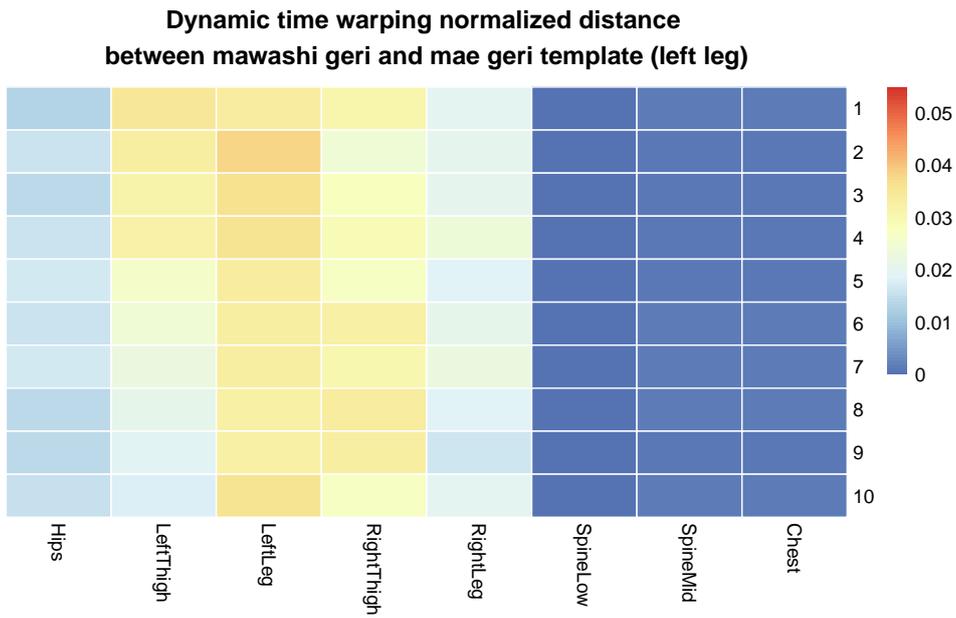


Figure 10. Dynamic time warping normalized distances between the template of mae geri kick with the left leg and the mawashi geri MoCap. The template was generated from the Shorin-ryu karate master data, while other MoCap came from the Oyama karate master. Distances are presented on the matrix with colors.

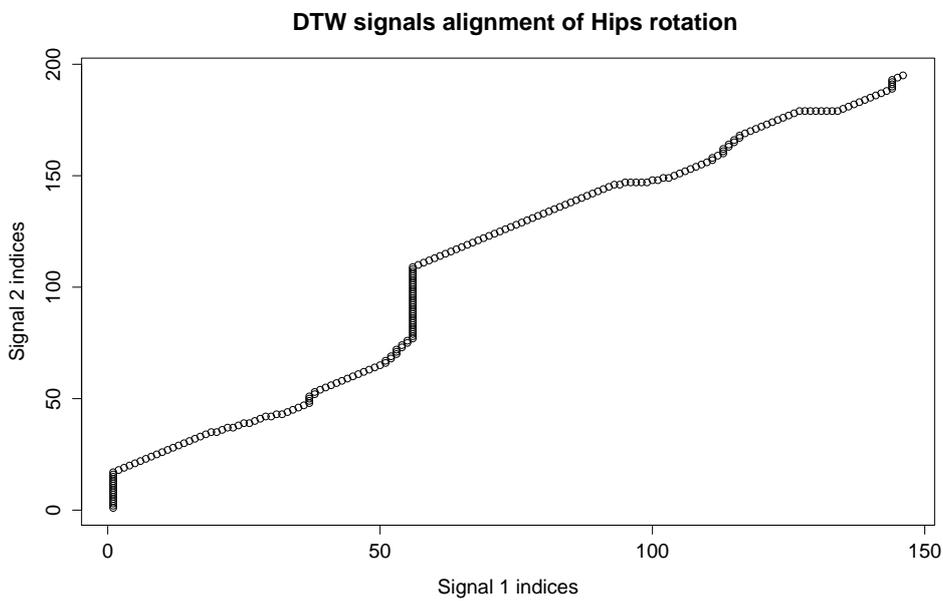


Figure 11. DTW signal alignment of hips rotation between the template of mae geri kick with the left leg (Signal 1) and one selected MoCap recording (Signal 2) of mawashi geri kick. The template was generated from the Shorin-ryu karate master data, while the other MoCap came from the Oyama karate master.

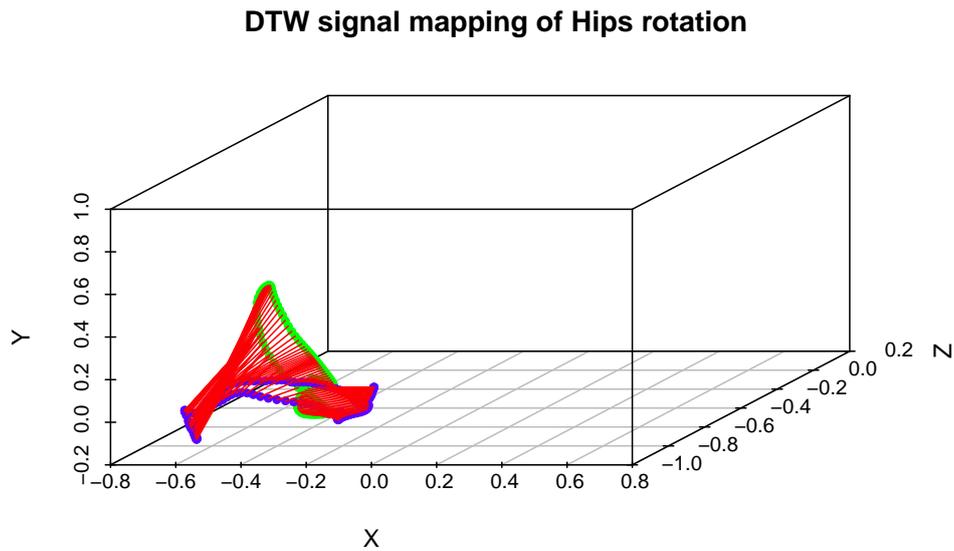


Figure 12. DTW signal mapping of the hips rotation between the template of mae geri kick with the left leg (green signal) and one selected MoCap recording (blue signal) of mawashi geri kick. Those are the same MoCap as in Figure 11. Distances are presented as red sections.

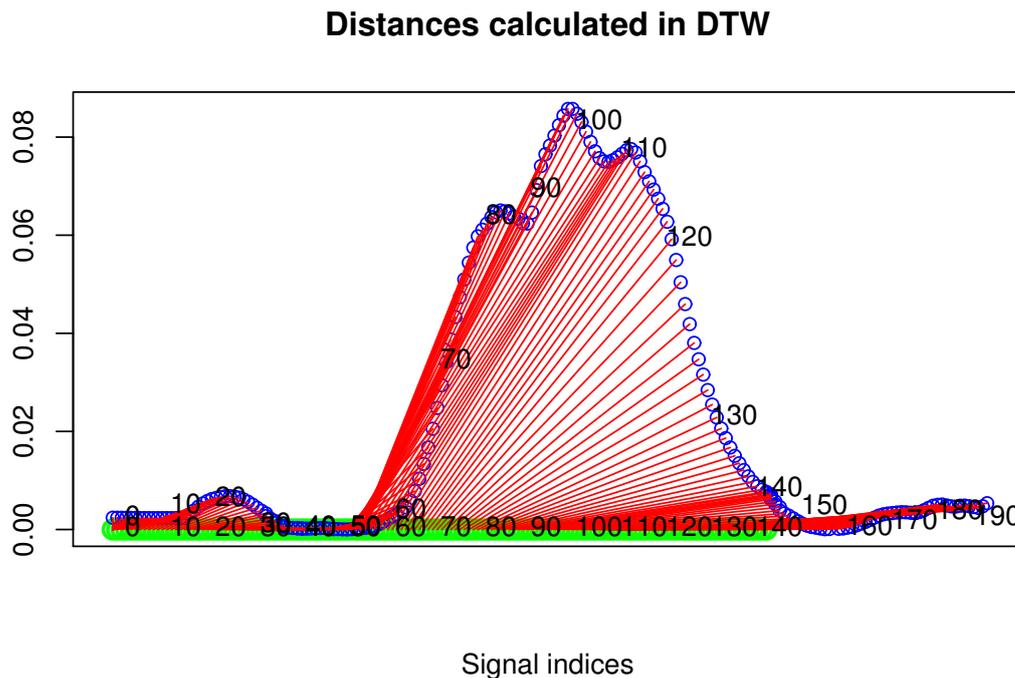


Figure 13. Distance of hips rotation between the template of mae geri kick with the left leg (green signal) and one selected MoCap recording (blue signal) of mawashi geri kick. Those are the same MoCap as in Figure 11. Distances are presented as red sections.

4. Discussion

The results presented in the previous section prove that all proposed methods work as expected and are easy to use and robust algorithms. The proposed signal averaging method is a combination of the DBA algorithm with Markley's algorithm performed on quaternions. As can be seen in Figure 2, the smoothing algorithm deals very well with high frequency noises. This can be clearly noticed

when comparing a black (out) signal with the red (smoothed) one, which has a nearly identical shape, however without high frequency “leaps”, which are the result of the DBA heuristic. Our smoothing algorithm is basically very similar to the typical low-pass convolution filter that uses the Gaussian kernel; however, instead of the weighted summation of values in the filter window, we used Markley’s approach. Human motion recordings done with a high-frequency MoCap system (like ours) do not contain signal samples with values highly deviating from adjoining samples like those visible in Figure 2 after 75×10^{-2} s (it is virtually impossible for a human being to move that fast). Due to this, we can safely assume that this is a signal disruption introduced either by the MoCap hardware or DBA algorithm, and it should be filtered out. The application of the quaternion representation of angles allows us to overcome the problems of discontinuity in rotation descriptions in Euler angles.

The overall quality of templates measured as the median value of dynamic time warping normalized distances calculated for each karate technique also returns very good (low) results. The median distance for each technique does not overcome the order of 10^{-4} (see Figures 6 and 7). In the case of recordings about 2–4 s in length that we are dealing with, this distance means that the differences between MoCap data nearly cannot be spotted by visual observation. Of course, the lengths of MoCap between template and input data might differ; however, we are nearly sure that we generated the same action. This fact was confirmed by our observation. The more detailed results are presented in Figure 8. As can be observed, most of the DTW distances between template and input data have very small values. Most often, the larger distances are in the joint that moves the most; in the case of mae geri, it might be for example the thigh. However, in the case of highly-skilled class athletes, those differences are very low because they have high repeatability of movements.

In the second experiment, we have compared the template of the same technique (mae geri) generated for two different karate masters. We wanted to find out if the proposed aligning metrology allows us to detect the most similar and least similar parts of the MoCap. We want to do it in order to generate the visual feedback for human action data comparison at a very precise level. The first important step in this type of comparison is to calculate the normalized DTW distance between the template and MoCap we want to compare and to visualize it for example in the form of a matrix (see Figure 9). The output of our algorithm that is presented as a matrix (Figure 9) or more detailed DTW signal mappings (Figures 3–5) speed up the analysis of motion capture renderings. Matrices indicate which body joints should be compared, while DTW mappings allow us to detect frames that correspond to local maximums (top row) and local minimums (bottom row) of distances (Figure 10). All of these methods might be a part of the computer-aided training system giving additional feedback to the coach. Of course, like in all computer-aided systems, the final decision if some differences are relevant from the perspective of the martial art and should be eliminated to improve the techniques is made by the trainer.

We can spot that the biggest differences are between leg (thigh) joints. Differences are caused by different flexibility (or height of the kick) of athletes and different initial stances. However, the most interesting for us is the hips joint that drives the direction of the whole kick. Figure 9 shows that there are not many differences between the template and MoCap of other athlete. In order to make a deeper analysis, we can generate a DTW signal alignment plot (see Figure 5). In this kind of plot, if its direction is more or less diagonal, we know that one signal is warped to another one smoothly. In our case, the first 50 frames of Signal 2 are wrapped to the first frame of Signal 1. That is because the initial stance of the second MoCap differs, and after 0.5 s, the hips position becomes more similar. The rest of the DTW proceeds smoothly. DTW signal mapping in Figure 3 allows us to visualize in 3D how one signal is aligned with another. It may be difficult to follow the signal trajectories when the image is printed on paper; however, this plot is very usable when one can rotate it around all axes and manipulate the number of points on the green and blue plot that are visualized. Thanks to this, we can clearly see the progress of movement and spot the differences and similarities in distance between signals and the timing of an action. When we are limited to two-dimensional visualization, the distance DTW plot presented in Figure 4 is far more understandable. Nevertheless, we have more

complex information in Figure 3 because in Figure 4, both signals are “straightened”, and the biggest emphasis is put on distances between points in the template signal to corresponding points in the second plot. Analyzing local maxima and minima, we can find out which parts of the actions are most and least similar. In Figure 14, the visualization in the first row, which contains the rendering of frames within the local maxima (biggest differences), assures us that the differences are caused by different hips trajectories. That is the result of different initial and final legs positions of both karate masters. However, the most important parts of the mae geri kick action visualized in the bottom row seems to be nearly identical, although the Shorin-ryu master does not lift the leg as high as the Oyama master. In other words, those kicks were performed at different heights. That all proves that the proposed procedure can be easily applied for matching, comparing and visualization of similar actions.

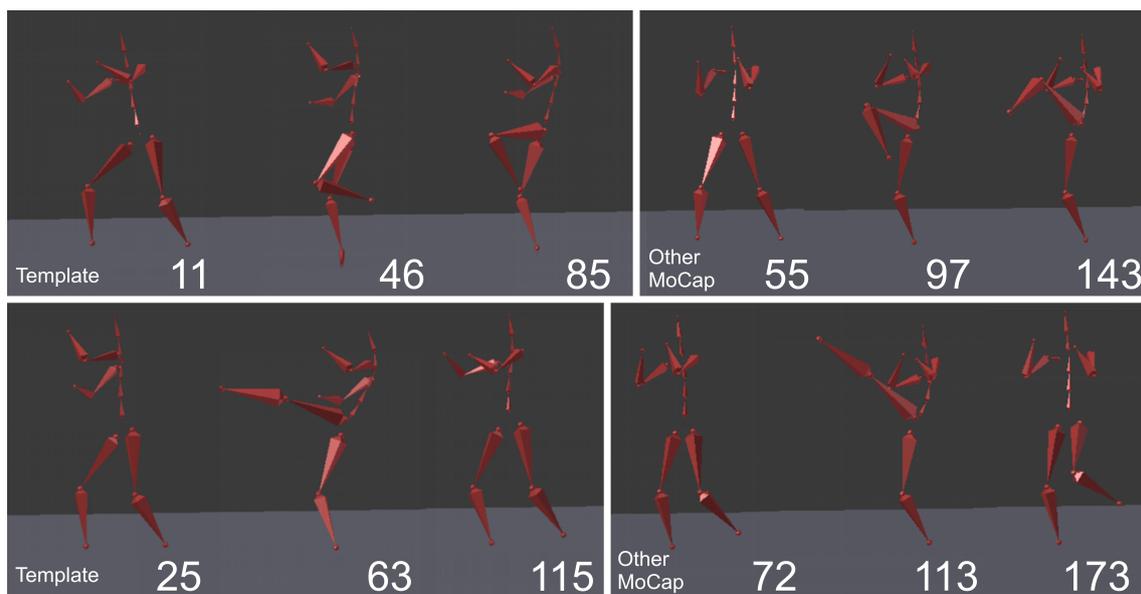


Figure 14. MoCap frames that correspond to local maxima (top row) and local minima (bottom row) of the distances presented in Figure 4. Numbers that are present close to each visualization are indices from Figure 4.

The third experiment is similar to the previous one; however, this time, we have compared the template of the front kick with the MoCap of the roundhouse kick. The matrix indicates this time higher differences than previously in the hips DTW aligning. The differences are very clearly visible in the DTW signal alignment plot (Figure 11), where there is a large section that was not proportionally aligned in the middle of the recording (about 60 frames of Signal 1). That is because the axis of rotation goes in a different plane, which is clearly visible in Figure 12. In the case of mae geri, it goes in the vertical plane (green signal), while in mawashi geri, it goes in the horizontal plane (blue signal). The trajectory difference is also visualized in Figure 13. We can easily see without any MoCap rendering that we are dealing with actions that have totally different hips trajectories that are most similar only in the initial body position. That evaluation proves that the proposed algorithm works very well in comparing actions that are not similar to each other.

The results of our experiment might be influenced by the errors generated by the inertial MoCap system we used. Those might be both static and dynamic measuring errors. Various researchers report that while internal systems have good absolute static accuracy, clinically acceptable absolute accuracy under the condition of slow motion, the absolute and relative accuracies are significantly affected by velocity during sustained motions [44,58]. Those potential errors might be a source of additional differences between acquired body actions that were later used to generate the single template. If the introduced errors are random, the DBA algorithm might decrease their impact on the resulting template.

The systematic error cannot be overcome by DBA. Furthermore, the environmental conditions might perturb the measurements to some extent, although we did not detect any electromagnetic disturbances with the electromagnetic field meter in the gymnasium where data acquisition had been performed.

5. Conclusions

Based on the results discussed in the previous section, we can conclude that our proposed methodology can be easily applied to human actions analysis, that is to template generation, matching, comparing and visualization. It works the best when we are dealing with relatively short actions (from 2–4 s) that might be averaged and aligned with the DTW framework. In practice, the methodology is designed to optimize the performance of some full-body techniques performed in various sports, for example combat sports and martial arts [59]. We can also use it to generate templates or to compare the correct performance of techniques between various top sportsmen to generate a knowledge base of reference MoCap videos. This type of knowledge base might be a very important source of information for coaches and athletes. The averaged sequences generated by our algorithm might be directly applied in the classification algorithms that use DTW.

In further research, we want to use our methodology to deeply examine the movement repeatability of athletes at different levels of advancement. With a large enough database, we will be capable of drawing conclusions about numerical values returned by our method; for example, what is an accepted variance of DTW distance between the template and MoCap data to consider actions similar and performed correctly. We also want to apply the comparison and visualization potential of our algorithm to aid in the process of training.

Supplementary Materials: The following are available online at www.mdpi.com/1424-8220/17/11/2590/s1, 1.mpg: mae geri kick with the left leg of the Shorin-ryu karate master; 2.mpg: mawashi geri kick with the left leg of the Oyama karate master.

Acknowledgments: This work has been supported by the National Science Centre, Poland, under Project Number 2015/17/D/ST6/04051.

Author Contributions: The ideas and concept was carried out in collaboration with all authors. T.H. designed the framework, carried out the experiments, and wrote this paper. M.P. and T.H. gathered the data, M.O. and M.P. contributed to the data analysis, the verification of the approach, and reviewed the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Neto, O.P.; Silva, J.H.; Marzullo, A.C.; Bolander, R.P.; Bir, C.A. The effect of hand dominance on martial arts strikes. *Hum. Mov. Sci.* **2012**, *31*, 824–833, doi:10.1016/j.humov.2011.07.016.
2. Matsumoto, T.; Konno, A.; Gou, L.; Uchiyama, M. A Humanoid Robot that Breaks Wooden Boards Applying Impulsive Force. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 9–15 October 2006; pp. 5919–5924, doi:10.1109/IROS.2006.282473.
3. Voigt, M.; Klausen, K. Changes in muscle strength and speed of an unloaded movement after various training programmes. *Eur. J. Appl. Physiol. Occup. Physiol.* **1990**, *60*, 370–376.
4. Neto, O.P.; Magini, M. Electromyographic and kinematic characteristics of Kung Fu Yau-Man palm strike. *J. Electromyogr. Kinesiol.* **2008**, *18*, 1047–1052, doi:10.1016/j.jelekin.2007.03.009.
5. Vences Brito, A.M.; Rodrigues Ferreira, M.A.; Cortes, N.; Fernandes, O.; Pezarat-Correia, P. Kinematic and electromyographic analyses of a karate punch. *J. Electromyogr. Kinesiol.* **2011**, *21*, 1023–1029, doi:10.1016/j.jelekin.2011.09.007.
6. Moreira, P.V.; Goethel, M.F.; Gonçalves, M. Neuromuscular performance of Bandal Chagui: Comparison of subelite and elite taekwondo athletes. *J. Electromyogr. Kinesiol.* **2016**, *30*, 55–65, doi:10.1016/j.jelekin.2016.06.001.
7. Kolahi, A.; Hoviattalab, M.; Rezaeian, T.; Alizadeh, M.; Bostan, M.; Mokhtarzadeh, H. Design of a marker-based human motion tracking system. *Biomed. Signal Process. Control* **2007**, *2*, 59–67.
8. Ilg, W.; Mezger, J.; Giese, M. Estimation of Skill Levels in Sports Based on Hierarchical Spatio-Temporal Correspondences. *Pattern Recognit.* **2003**, *2781*, 523–531, doi:10.1007/978-3-540-45243-0_67.

9. Quinzi, F.; Camomilla, V.; Felici, F.A.; Mario, D.; Sbriccoli, P. Differences in neuromuscular control between impact and no impact roundhouse kick in athletes of different skill levels. *J. Electromyogr. Kinesiol.* **2013**, *23*, 140–150, doi:10.1016/j.jelekin.2012.09.006.
10. Sbriccoli, P.; Camomilla, V.; Di Mario, A.; Quinzi, F.; Figura, F.; Felici, F. Neuromuscular control adaptations in elite athletes: The case of top level karateka. *Eur. J. Appl. Physiol.* **2010**, *108*, 1269–1280. doi:10.1007/s00421-009-1338-5.
11. Irina, B.; Dan, D. Kinematic Analysis of the Cross Punch Applied in the Full-contact System Using Inertial Navigation Technology and Surface Electromyography. *Procedia* **2014**, *117*, 335–340.
12. Pliske, G.; Emmermacher, P.; Weinbeer, V.; Witte, K. Changes in dual-task performance after 5 months of karate and fitness training for older adults to enhance fall prevention. *Aging Clin. Exp. Res.* **2016**, *28*, 1179–1186, doi:10.1007/s40520-015-0508-z.
13. Hachaj, T.; Ogiela, M.R.; Koptyra, K. Application of Assistive Computer Vision Methods to Oyama Karate Techniques Recognition. *Symmetry* **2015**, *7*, 1670–1698, doi:10.3390/sym7041670.
14. Burke, M.; Lasenby, J. Estimating missing marker positions using low dimensional Kalman smoothing. *J. Biomech.* **2016**, *49*, 1854–1858, doi:10.1016/j.jbiomech.2016.04.016.
15. Jin, M.; Zhao, J.; Jin, J.; Yu, G.; Li, W. The adaptive Kalman filter based on fuzzy logic for inertial motion capture system. *Measurement* **2014**, *49*, 196–204.
16. El-Gohary, M.; McNames, J. Human Joint Angle Estimation with Inertial Sensors and Validation with A Robot Arm. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1759–1767. doi:10.1109/TBME.2015.2403368.
17. Joukov, V.; Karg, M.; Kulic, D. Online tracking of the lower body joint angles using IMUs for gait rehabilitation. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2014**, *2014*, 2310–2313, doi:10.1109/EMBC.2014.6944082.
18. Masiero, A.; Cenedese, A. A Kalman filter approach for the synchronization of motion capture systems. In Proceedings of the 51st IEEE Conference on Decision and Control, Maui, HI, USA, 10–13 December 2012; pp. 2028–2033. doi:10.1109/CDC.2012.6425864.
19. Qi, Y.; Soh, C.B.; Gunawan, E.; Low, K.S. A wearable wireless ultrasonic sensor network for human arm motion tracking. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2014**, 5960–5963, doi:10.1109/EMBC.2014.6944986.
20. Gheller, R.G.; Dal Pupo, J.; Ache-Dias, J.; Detanico, D.; Padulo, J.; dos Santos, S.G. Effect of different knee starting angles on intersegmental coordination and performance in vertical jumps. *Hum. Mov. Sci.* **2015**, *42*, 71–80, doi:10.1016/j.humov.2015.04.010.
21. Vuk, S.; Markovic, G.; Jaric, S. External loading and maximum dynamic output in vertical jumping: The role of training history. *Hum. Mov. Sci.* **2012**, *31*, 139–151, doi:10.1016/j.humov.2011.04.007.
22. Hadizadeh, M.; Amri, S. Mohafez, H.; Roohi, S.A.; Mokhtar, A.H. Gait analysis of national athletes after anterior cruciate ligament reconstruction following three stages of rehabilitation program: Symmetrical perspective. *Gait Posture* **2016**, *48*, 152–158, doi:10.1016/j.gaitpost.2016.05.002.
23. Vishnoi, N.; Mitra, A.; Duric, Z.; Gerber, N.L. Motion based markerless gait analysis using standard events of gait and ensemble Kalman filtering. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2014**, 2512–2516, doi:10.1109/EMBC.2014.6944133.
24. Soltani, P.; Figueiredo, P.; Fernandes, R.J.; Vilas-Boas, J.P. Do player performance, real sport experience, and gender affect movement patterns during equivalent exergame? *Comput. Hum. Behav.* **2016**, *63*, 1–8, doi:10.1016/j.chb.2016.05.009.
25. Müller, M.; Röder, T. Motion templates for automatic classification and retrieval of motion capture data. In Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation, Vienna, Austria, 2–4 September 2006; pp. 137–146.
26. Jabłoński, B. Quaternion dynamic time warping. *IEEE Trans. Signal Process.* **2012**, *60*, 1174–1183, doi:10.1109/TSP.2011.2177832.
27. Sempena, S.; Maulidevi, N.U.; Aryan, P.R. Human action recognition using dynamic time warping. In Proceedings of the International Conference on Electrical Engineering and Informatics (ICEEI), Bandung, Indonesia, 17–19 July 2011.
28. Wojciechowski, K.; Polański, A.; Josiński, H.; Świtoński, A.; Michalczuk, A. Dynamic Time warping in gait classification of motion capture data. *Proc. World Acad. Sci. Eng. Technol.* **2012**, *6*, 1289–1294.
29. Świtoński, A.; Josinski, H.; Zghidi, H.; Wojciechowski, K. Selection of pose configuration parameters of motion capture data based on dynamic time warping. *AIP Conf. Proc.* **2015**, *1648*, 660003, doi:10.1063/1.4912879.

30. Liu, X.P.; Xia, X.Y.; Lu, J.T. Normalization of motion sequences based on DTW and hermite interpolation. *J. Syst. Simul.* **2013**, *25*, 1598–1603.
31. Skurowski, P.; Prusowski, P.; Peşzor, D. Synchronization of motion sequences from different sources. *AIP Conf. Proc.* **2016**, *1738*, 180013, doi:10.1063/1.4951960.
32. Tunçel, O.; Altun, K.; Barshan, B. Classifying human leg motions with uniaxial piezoelectric gyroscopes. *Sensors* **2009**, *9*, 8508–8546, doi:10.3390/s91108508.
33. Stančin, S.; Tomažič, S. Early Improper Motion Detection in Golf Swings Using Wearable Motion Sensors: The First Approach. *Sensors* **2013**, *13*, 7505–7521, doi:10.3390/s130607505.
34. Chen, C.; Jafari, R.; Kehtarnavaz, N. Improving human action recognition using fusion of depth camera and inertial sensors. *IEEE Trans. Hum. Mach. Syst.* **2015**, *45*, 51–61, doi:10.1109/THMS.2014.2362520.
35. Liu, H.; Ju, Z.; Ji, X.; Chan, C.S.; Khoury, M. *Human Motion Sensing and Recognition*; Springer: Berlin, Germany, 2017, doi:10.1007/978-3-662-53692-6.
36. Ju, Z.; Gao, D.; Cao, J.; Liu, H. A novel approach to extract hand gesture feature in depth images. *Multimed. Tools Appl.* **2016**, *75*, 11929–11943, doi:10.1007/s11042-015-2609-2.
37. Hachaj, T.; Ogiela, M.R.; Piekarczyk, M.; Koptyra, K. *Averaging Three-Dimensional Time-Varying Sequences of Rotations: Application to Preprocessing of Motion Capture Data*; Springer: Berlin, Germany, 2017; pp. 17–28, doi:10.1007/978-3-319-59126-1_2.
38. The Website Containing Motion Database. Available online: <http://gdl.org.pl/> (accessed on 1 November 2017).
39. Yamazaki, J.; Muneta, T.; Ju, Y.-J.; Koga, H.; Morito, T.; Sekiya, I. The kinematic analysis of female subjects after double-bundle anterior cruciate ligament reconstruction during single-leg squatting. *J. Orthopaedic Sci.* **2013**, *18*, 284–289, doi:10.1007/s00776-012-0350-5.
40. Bu, B.; Han, H.; Liu, Y.; Zhang, C.; Yang, X.; Singh, M.F. Effects of martial arts on health status: A systematic review. *J. Evid. Based Med.* **2010**, *3*, 205–219, doi:10.1111/j.1756-5391.2010.01107.x.
41. Vignais, N.; Kulpa, R.; Brault, S.; Presse, D.; Bideau, B. Which technology to investigate visual perception in sport: Video vs. virtual reality. *Hum. Mov. Sci.* **2015**, *39*, 12–26, doi:10.1016/j.humov.2014.10.006.
42. Marin, J.; Blanco, T.; Marin, J.J. Octopus: A Design Methodology for Motion Capture Wearables. *Sensors* **2017**, *17*, 1875; doi:10.3390/s17081875.
43. Vamsikrishna, K.M.; Dogra, D.P.; Bhaskar, H. Classification of head movement patterns to aid patients undergoing home-based cervical spine rehabilitation, Acoustics. In Proceedings of the Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 849–853, doi:10.1109/ICASSP.2016.7471795.
44. Lebel, K.; Boissy, P.; Hamel, M.; Duval, C. Inertial Measures of Motion for Clinical Biomechanics: Comparative Assessment of Accuracy under Controlled Conditions - Effect of Velocity. *PLoS ONE* **2013**, *8*, e79945. doi:10.1371/journal.pone.0079945.
45. Zhang, B.; Jiang, S.; Wei, D.; Marschollek, M.; Zhang, W. State of the Art in Gait Analysis Using Wearable Sensors for Healthcare Applications. In Proceedings of the IEEE/ACIS 11th International Conference on Computer and Information Science, Shanghai, China, 30 May–1 June 2012; pp. 213–218, doi:10.1109/ICIS.2012.100.
46. Schulze, M.; Liu, T.-H.; Xie, J.; Zhang, W.; Wolf, K.-H.; Calliess, T.; Windhagen, H.; Marschollek, M. Unobtrusive ambulatory estimation of knee joint angles during walking using gyroscope and accelerometer data—A preliminary evaluation study. In Proceedings of the IEEE-EMBS International Conference on Biomedical and Health Informatics, Hong Kong, China, 5–7 January 2012; pp. 559–562, doi:10.1109/BHI.2012.6211643.
47. An, Q.; Ishikawa, Y.; Nakagawa, J.; Kuroda, A.; Oka, H.; Yamakawa, H.; Yamashita, A.; Asama, H. Evaluation of wearable gyroscope and accelerometer sensor (PocketIMU2) during walking and sit-to-stand motions. In Proceedings of the IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, Paris, France, 9–13 September 2012; pp. 731–736, doi:10.1109/ROMAN.2012.6343838.
48. Cloete, T.; Scheffer, C. Benchmarking of a full-body inertial motion capture system for clinical gait analysis. In Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, USA, 20–24 August 2008; pp. 4579–4582, doi:10.1109/IEMBS.2008.4650232.
49. Valtazanov, A.; Arvind, D.K.; Ramamoorthy, S. Using wearable inertial sensors for posture and position tracking in unconstrained environments through learned translation manifolds. In Proceedings of the

- ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), Philadelphia, PA, USA, 8–11 April 2013; pp. 241–252, doi:10.1145/2461381.2461411.
50. Solberg, R.T.; Jensenius, A.R. Optical or Inertial? Evaluation of two motion capture systems for studies of dancing to electronic dance music. In Proceedings of the SMC Conference on Creative Commons, Hamburg, DE, USA, 23 August–3 September 2016; pp. 469–474.
 51. Funakoshi, G. *Karate-Do Kyohan: The Master Text, Kodansha International*, 1st ed.; Kodansha International: Tokio, Japan, 2013.
 52. Petitjean, F.; Ketterlin, A.; Gançarski, P. A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognit.* **2011**, *44*, 678–693.
 53. Markley, F.L.; Cheng, Y.; Crassidis, J.L.; Oshman, Y. Averaging Quaternions. *J. Guid. Control Dyn.* **2007**, *30*, 1193–1197, doi:10.2514/1.28949.
 54. Xu, W.; Zhang, M.; Sawchuk, A.A.; Sarrafzadeh, M. Robust Human Activity and Sensor Location Corecognition via Sparse Signal Representation. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 3169–3176, doi:10.1109/TBME.2012.2211355.
 55. Chen, C.; Kehtarnavaz, N.; Jafari, R. A medication adherence monitoring system for pill bottles based on a wearable inertial sensor. In Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Chicago, IL, USA, 26–30 August 2014, doi:10.1109/EMBC.2014.6944743.
 56. Hachaj, T.; Koptyra, K.; Ogiela, M.R. Initial Proposition of Kinematics Model for Selected Karate Actions Analysis. In Proceedings of the Ninth International Conference on Machine Vision (ICMV 2016), Nice, France, 18–20 November 2016, doi:10.1117/12.2268402.
 57. Hachaj, T.; Ogiela, M.R.; Koptyra, K. Human actions modeling and recognition in low-dimensional feature space. In Proceedings of the BWCCA 2015, 10th International Conference on Broadband and Wireless Computing, Communication and Applications, Krakow, Poland, 4–6 November 2015; pp. 247–254, doi:10.1109/BWCCA.2015.15.
 58. Brodie, M.A.; Walmsley, A.; Page, W. Dynamic accuracy of inertial measurement units during simple pendulum motion. *Comput. Methods Biomech. Biomed. Eng.* **2008**, *11*, 235–242. doi:10.1080/10255840802125526.
 59. Polak, E.; Kulasa, J.; Vences Brito, A.; Castro, M.A.; Fernandes, O. Motion analysis systems as optimization training tools in combat sports and martial arts. *Rev. Artes Marciales Asiaticas*. **2015**, *10*, 105–123, doi:10.18002/rama.v10i2.1687.



© 2017 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/>).