JMIR Publications
The Leading eHealth Publisher

Objective Assessment of Rehabilitation Exergames through Motion Sensing and Machine Learning

Reza Haghighi Osgouei, David Soulsby, Fernando Bello

Submitted to: JMIR Rehabilitation and Assistive Technologies
on: December 04, 2019

Preprint
JMIR Publications

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on it's website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressively prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript **4**
Supplementary Files **29**
Figures **30**
Figure 1 ..... **31**
Figure 2 ..... **32**
Figure 3 ..... **33**
Figure 4 ..... **34**
Figure 5 ..... **35**
Figure 6 ..... **36**
Figure 7 ..... **37**
Figure 8 ..... **38**
Figure 9 ..... **39**
Figure 10 ..... **40**
Figure 11 ..... **41**
Figure 12 ..... **42**
Figure 13 ..... **43**
TOC/Feature image for homepages **44**
TOC/Feature image for homepage 0 ..... **45**

Preprint JMIR Publications

Objective Assessment of Rehabilitation Exergames through Motion Sensing and Machine Learning

Reza Haghighi OsgoueiPhD, ; David Soulsby; Fernando BelloPhD,

Corresponding Author:
Reza Haghighi OsgoueiPhD,
Phone: +44940419331
Email: r.haghighi-osgouei@imperial.ac.uk

Abstract

Background: Performing physiotherapy exercises in front of a physiotherapist yields qualitative assessment notes and immediate feedback. However, practicing the exercises at home lacks feedback on how well or not patients are performing the prescribed tasks. The absence of proper feedback might result in patients doing the exercises incorrectly, which could worsen their condition.

Objective: We propose the use of two machine learning algorithms, namely Dynamic Time Warping (DTW) and Hidden Markov Model (HMM), to quantitively assess the patient's performance with respects to a reference.

Methods: Movement data were recorded using a Kinect depth sensor, capable of detecting 25 joints in the human skeleton model, and were compared to those of a reference. 16 participants were recruited to perform four different exercises: shoulder abduction, hip abduction, lunge, and sit-to-stand. Their performance was compared to that of a physiotherapist as a reference.

Results: Both algorithms show a similar trend in assessing participants' performance. However, their sensitivity level was different. While DTW was more sensitive to small changes, HMM captured a general view of the performance, being less sensitive to the details.

Conclusions: The chosen algorithms demonstrated their capacity to objectively assess physical therapy performances. HMM may be more suitable in the early stages of a physiotherapy program to capture and report general performance, whilst DTW could be used later on to focus on the detail.

(JMIR Preprints 04/12/2019:17289)
DOI: https://doi.org/10.2196/preprints.17289

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Preprint JMIR Publications

Original Paper

Reza Haghighi Osgouei[1], PhD, David Soulsby[2], Fernando Bello[1], PhD

[1]Simulation and Modelling in Medicine and Surgery (SiMMS) Research Group, Imperial College Centre for Engagement and Simulation Science (ICCESS), Department of Surgery and Cancer, Faculty of Medicine, Imperial College London
[2]Acting Therapy Lead Children's Services and Dietetics, Chelsea and Westminster NHS Foundation Trust

Objective Assessment of Rehabilitation Exergames through Motion Sensing and Machine Learning

Abstract

Background: Performing physiotherapy exercises in front of a physiotherapist yields qualitative assessment notes and immediate feedback. However, practicing the exercises at home lacks the feedback on how well patients are performing the prescribed tasks. Absence of proper feedback might result in patients doing the exercises incorrectly, which could worsen their condition. We present an approach to generate performance scores to enable tracking the progress by both the patient at home and the physiotherapist in the clinic.
Objective: We propose the use of two machine learning algorithms, namely Dynamic Time Warping (DTW) and Hidden Markov Model (HMM), to quantitively assess the patient's performance with respects to a reference.
Methods: Movement data were recorded using a motion sensor (Kinect V2), capable of detecting 25 joints in human skeleton model, and were compared to those of a reference. 16 participants were recruited to perform four different exercises: shoulder abduction, hip abduction, lunge, and sit-to-stand. Their performance was compared to that of a physiotherapist as a reference.
Results: Both algorithms show a similar trend in assessing participants performance. However, their sensitivity level was different. While DTW was more sensitive to small changes, HMM captured a general view of the performance, being less sensitive to the details.
Conclusions: The chosen algorithms demonstrated their capacity to objectively assess physical therapy performance. HMM may be more suitable in the early stages of a physiotherapy program to capture and report general performance, whilst DTW could be used later on to focus on the detail. The scores enable the patient to monitor their daily performance. They can also be reported back to the physiotherapist to track and assess patient progress, provide feedback and adjust the exercise program if needed.
Trial Registration: Not applicable.

Keywords: Rehabilitation Exergames; Performance Assessment; Similarity Score; Motion Sensing; Machine Learning; Dynamic Time Warping; Hidden Markov Model

Introduction

Rehabilitation is essential to regain lost or weakened functionality after injury or surgery. While it is commonly initiated in a clinic and supervised by a physiotherapist, the prescribed therapeutic exercises normally will need to be practiced at home by patients on their own. Lack of motivation and compliance may hinder the healing process and, in some cases, even worsen the injury. Advances in Virtual Reality (VR) technologies have resulted in various virtual rehabilitation platforms introduced to address this issue [1-3]. They are generally equipped with sensory devices to track and monitor patient's movement whilst performing the exercises. Among them are systems based on the concept of *Exergaming*, exercise gaming. These are interactive video games, with some simple scenarios, that enable a patient to perform a therapeutic exercise by playing the game. Exergames are a subgenre of serious games developed for the purpose of encouraging exercise and activity. While sharing some aspects with medical simulation, here the emphasis is more on the added educational value of fun and entertainment. Accessibility, not needing the constant presence of a physician, and entertainability, turning repetitive tasks to playful activities, are two key advantages of such systems.

The impact of serious games on physical therapy has been studied in term of effectiveness [4-6] and motivational determinants [7, 8]. However, as a virtual guide replacing or complementing a real physician, exergaming systems tend to lack objective, clinically meaningful evaluation of patient's performance. At best, game scores, the extent to which the player achieved the game goals, are reported at the end of the session, along with statistics such as completion time. The work presented here aims to fill this gap by introducing an approach to compare a patient's performance with that of a reference, using the MIRA (Medical Interactive Recovery Assistant) rehabilitation platform. We explore two different machine learning techniques, Dynamic Time Warping (DTW) and Hidden Markov Model (HMM). They have been widely used for gesture recognition to study acquired motion trajectories [9-14]. The former belongs to model-less time domain methods, whereas the latter is a model-based probabilistic technique for time series analysis. Comparing the patient's trajectory with that of a physiotherapist as a reference, each approach generates an objective similarity score indicating how similar were the performances.

Related Work

Rehabilitation Platforms

The concept of exergaming enables exercising while playing games. For players, it is an opportunity to play games in a more active and less passive way. For patients, it offers the chance to practice therapeutic tasks in a more playful and less repetitive manner. Exergames offer various activities, such as aerobic exercises and dancing, balance and stretching workouts, and recreational simulations, such as golf, skiing, and more. However, they exert additional requirements on hardware and software. Hardware-wise, they require proper sensory equipment to track user's motion. In terms of software, the game scenario must accommodate whole body interaction. There are various commercially available game consoles enabling exergames, including Xbox (Microsoft), PlayStation (Sony), and Wii (Nintendo). Each comes with its own dedicated input device for enabling user interaction with the games, i.e., Kinect for Xbox, Move for PlayStation, and Remote Plus for Wii.

Among them, the Kinect has gained higher popularity due to its acceptable performance and versatility [15]. It was introduced by Microsoft in 2010 as a peripheral input device for its gaming

console Xbox and discontinued in 2017. Kinect enables interaction with virtual environments using gestures, rather than conventional controllers. The device includes an RGB camera and a depth sensor, which combined provide full-body 3D motion capture capabilities and gesture recognition. Using the Kinect SDK 2.0, the position of 25 human skeleton joints can be accessed with a sample rate of 30 fps. Since its launch, researchers have used the Kinect for various applications, including rehabilitation [16-21]. Other similar motion sensing devices capable of tracking 3D position of joints include: Azure Kinect, Orbbec Astra, Intel RealSense, Occipital Structure Sensor, and Creative BlasterX Senz3D.

Kinerehab was introduced in [16] to assist therapists in rehabilitating students at a public school. Conducting a user study with two patients, the authors compared two experimental phases, baseline (without assistive technology) and intervention (with Kinerehab), lifting both arms to the front, to the side, and upwards. The data showed a significant increase in motivation (willingness to keep practicing) amongst participants and hence improved exercise performance using Kinerehab, compared to the presence of a therapist. A Kinect-based serious game for physiotherapy (KSGphysio) was proposed in [17] with a mobile interface to facilitate analysis of patients progress by generating relevant statistics. In [18], a web-based serious game called Therasoup was developed to improve the patient's motivation while performing exercises, and to provide technical data to the physiotherapist which can help in the assessment. The game tries to simulate a daily life activity such as cooking, where the player controls an avatar to pick the ingredients from shelves and put them in a pan at the center. In [19], a Kinect-enabled home-based rehabilitation system (KEHR) was developed to assist patients in conducting safe and effective off-hospital rehabilitation, without the immediate supervision of a physician. KEHR supported three different shoulder rehabilitation exercises: shoulder abduction, shoulder anterior elevation and shoulder external and internal rotation. A serious game framework for therapy (Theragame) providing options to imitate the actions performed by an avatar or to play a game that trains specific parts of the body was introduced in [20]. A web-based platform for physical telerehabilitation for patients after hip replacement surgery was described in [21] having two goals in mind: making use of a low-cost motion capture device (Kinect) and real-time automatic assessment of performance. A comprehensive review of technical and clinical impacts of the Kinect in physical therapy and rehabilitation is given in [15]. The survey covers rehabilitation systems before and after the Kinect, and also platforms with and without clinical evaluation. While Kinect-based rehabilitation systems are accepted by both patients and therapists, the lack of objective, clinically meaningful evaluation of performance raises questions regarding their effectiveness.

Commercially available rehabilitation platforms based on Kinect and exergames include MIRA [1], VirtualRehab [2], and REHABILITY [3]. They all have the capacity to be used at home to encourage and monitor performance or at hospital to assist and support physiotherapists. MIRA is a class I medical device that uses games built based on best clinical practice and expertise from specialist physiotherapists, to keep patients engaged and motivated throughout the therapy. VirtualRehab, a CE certified class I medical device, is a product that can be used in clinics and hospitals, as well as in patient's homes, allowing them to continue their rehabilitation treatment. With REHABILITY, patients can carry out their rehabilitation exercises at the clinic/hospital, or remotely, but with constant medical supervision. None of these systems offers automatic, objective assessment comparing patient's performance with that of a reference. They offer game scores such as number of hits and statistics such as execution/completion time.

Performance Evaluation

Automatic performance evaluation of a user carrying out a task has always been a challenge amongst

researchers in both medical and non-medical domains. Such evaluations are usually conducted in the real world subjectively by judges who are experts in the given field. For example, evaluation of the quality of a dance, a gymnastic performance, or a physiotherapy/rehabilitation exercise, may be carried out by expert dancers, sport athletes, and professional therapists, respectively. Such evaluations require the presence of human specialists, who may not be easily accessible or affordable. In addition, the fact that the assessment is subjective, means that a different expert might have a different opinion. The ability to conduct objective automatic evaluations that are repeatable is thus highly desirable. A real-world example is a video game called Just Dance developed by the French company Ubisoft for Microsoft Xbox. Using the Kinect sensor, the players must mimic the on-screen dancer's choreography to a chosen song. The system then continuously evaluates in real-time the quality of a user's dance movements in terms of being ‘Ok’, ‘Good’, ‘Super’, or ‘Perfect’, and reports a total numeric score at the end [22].

Studies in the literature concerned with automated evaluation of therapy motions are scarce [23-26], and not much attention has been paid to the development of metrics for performance evaluation [27]. As a common scheme, a reference model is first captured as the ground truth. Then, a user's performance can be compared to the reference using machine learning approaches. A comprehensive taxonomy of the metrics for evaluation of patient performance in physical therapy is given in [27]. The metrics are classified into quantitative and qualitative categories. Further, quantitative metrics are divided into model-less (based on raw measurements of motions) and model-based (based on a mathematical model of the motions). Of the reviewed metrics, root-mean square distance, Kullback Leibler divergence, log-likelihood, and Fugl-Meyer Assessment have been used to classify a set of five human motions captured with a Kinect sensor.

Utilizing KEHR, the authors apply DTW and fuzzy logic to detect real-time subjective discrepancies between the model exercise and patient's performance [19]. Before applying either algorithm, the user's execution of the prescribed exercise was recorded under the supervision of a professional. Then, two factors were included for the assessment: 1) trajectory disparity, the motional path created by each joint over time and 2) speed variation: the time used to complete a designated exercise. Applying HMM and defining an accept/reject interval, a method to detect deviations from normal repetitions in therapeutic activities was presented in [23]. The authors later compared the performance of their HMM-based technique with that of DTW [28]. A similar approach using HMMs to assess the correctness of telerehabilitation exercises was employed in [24], whereas a cloud-based physical therapy monitoring and guidance system that applies DTW to produce subjective assessments in terms of being too slow/fast or overdone/incomplete was proposed in [25]. An error classification algorithm for therapy exercises based on incremental DTW is presented in [26] to classify the incorrect motions in a hip abduction exercise into four discrete categories: bent knee, foot outside, upper body, and wrong plane. A variance of DTW called multi-template, multi-match DTW was used in [29] to detect and evaluate physical therapy exercises using wearable motion sensors, providing a quantitative measure of similarity between an exercise execution and previously recorded templates. A continuous time-warping algorithm based on automatic motion assessment learning (A-MAL) was introduced in [30]. The resulting models produced numerical scores comparable with those of Fugl-Meyer Assessment. More recently, a DTW-based algorithm was developed for assessing Kinect-enabled home-based rehabilitation exercises in order to support auto-coaching in a virtual gaming environment [31]. Utilizing a simple but innovative method, the DTW distances are converted to meaningful performance scores in terms of percentage. Conducting a user study, the scores are then validated with the expert ratings showing a strong positive linear relationship. In another recent work, a deep learning-based framework for assessment of rehabilitation exercises is proposed [32]. The framework consists of algorithms for dimensionality reduction, performance metrics (based on Gaussian mixture model), scoring functions, and deep

learning models. Evaluating a dataset of 10 rehabilitation exercises, they showed the capacity of the trained models.

Similar techniques have also been employed in other domains, such as dance motion evaluation. In [33] they employed DTW and Laban movement analysis to evaluate the correctness of dance movements in terms of being best, good, bad, and worst. In another work [34], the authors employed HMM for dance gesture recognition and evaluation, comparing the results with those of domain experts in terms of being good, medium, or bad.

What is common between all these efforts, is that they focused on evaluating the incorrectness of the performance on the basis of subjective terms. In most cases, the developed method was used to sort multiple erratic performances with respect to a reference template. This approach motivated us to explore the use of DTW and HMM to generate a similarity score between a participant's performance and a reference.

Materials

MIRA

MIRA is a software platform that turns physiotherapy exercises into clinical exergames [35]. It aims to increase engagement levels and improve uptake of exercises by converting the rehabilitation sessions into entertaining activities, making therapy more convenient and easier to follow, and offering greater accessibility. In turn, this has the potential for shorter recovery times, as well as supporting physiotherapists, reducing workload and waiting times at clinics. MIRA has been used in several clinical studies [36-39].



Figure 1. MIRA system including a Kinect motion sensor and software to match an exercise with a game. A child (left) and an elderly gentleman (right) playing the Atlantis game. The child is practicing an arm exercise, the gentleman a hip exercise.

The MIRA system includes a Kinect V2 sensor (Microsoft Corp.) connected to a computer running the MIRA program (Figure 1). Currently, 32 exercises and 25 games are supported. Each rehabilitation session requires selecting an exercise and a suitable game (Figure 2). Shoulder abduction, elbow flexion and side strides are among the exercises, whereas Firefly, Fishing and Football are examples of available games. Once an exercise has been selected, adequate game options are presented. The selected combination of exercise and game is then added to the session and can subsequently be executed. Each execution starts with a process of calibrating the patient's position in front of the Kinect. A short video tutorial explaining the exercise is followed by another

video tutorial describing the game mechanics. As the game starts, the user must play it by moving the intended body part (i.e. left arm, right leg, or neck) in the manner shown in the video.



Figure 2. A snapshot of the MIRA program. Exercises are listed on the left and games are shown on the right. Multiple game options allow practicing the same exercise with different games, thus encouraging patients to cope with the prescribed exercise by discovering the various game scenarios.

At the end of each session, the MIRA system reports various scores. Depending on the game, they reflect to what extent the player followed the game's objectives. For example, the number of fish caught and taken to the boat, or the number of times the spaceship is safely passed through the fire rings. While the scores can be an indication of how well the user played the game, they do not have much value in a clinical context. The aim of this work is to introduce an objective evaluation method that is more meaningful and suitable for clinical evaluation.

Data collection

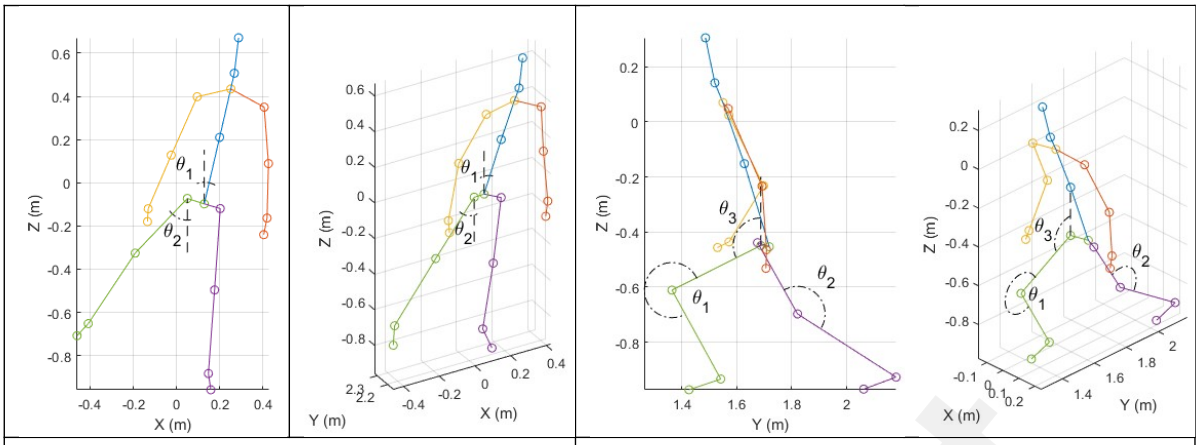
We developed a program in the Unity 3D game engine (Unity Technologies) [40] to capture and store raw 3D position coordinates of the selected joints using the Kinect. This was needed as the MIRA program does not allow accessing joint data while playing an exergame due to regulations imposed on class I medical devices. There were no technical problems executing both programs in parallel as the Kinect SDK permits simultaneous access to the sensor from multiple sources. Using Kinect V2, 3D position coordinates of 25 different human skeleton joints can be tracked with an update rate of 30 fps. However, it is not necessary to track all the joints, but only those that are involved in the chosen exercise. For this study, four different types of exercises were selected in consultation with a physiotherapist (DS): shoulder abduction (both left and right arms), hip abduction (both left and right

legs), lunge (both left and right legs), and sit-to-stand. This resulted in a total of seven exercises to be performed by participants. A brief description of each exercise is given below in Table 1.

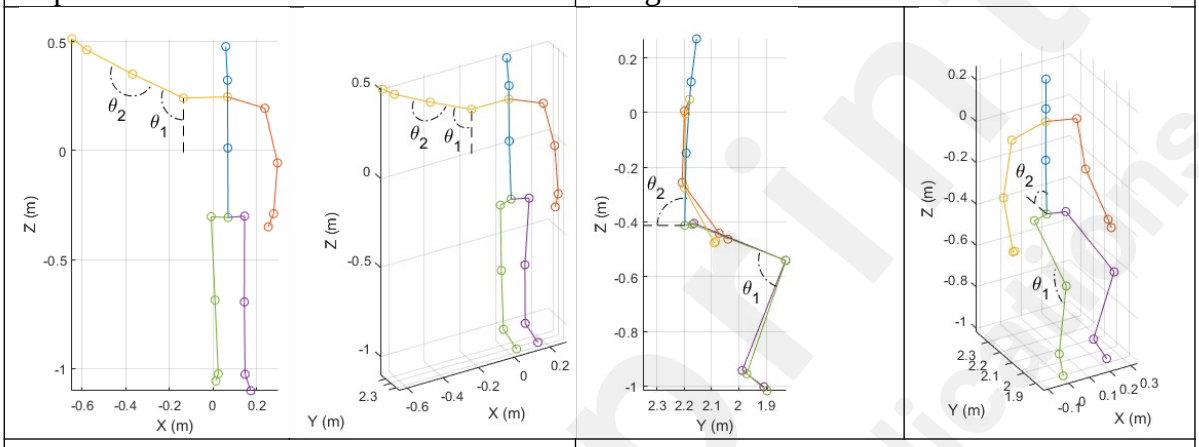
Table 1. Correct execution of each exercise

	Shoulder abduction: The arm should be kept close to the body. The exercise consists of raising the arm away from the side, keeping it in a straight line with the body.
	Hip abduction: The leg should be held straight and on the ground. The exercise involves raising the leg away from the side, keeping it in a straight line with the body.
	Lunge: Stand straight facing forward with the spine and pelvis in a neutral position. Take a step forward with a leg that is long enough so that when the knee bends, it does not go beyond the toes. Bend the back knee until it almost touches the floor, keeping both torso and spine in a neutral position. Return to the starting position.
	Sit-to-stand: Sit on a chair. Without using the hands for support, stand up and sit back down. Make sure each movement is slow and controlled.

As 3D position coordinates are dependent on the user size and location in front of the Kinect camera, we decided to extract invariant features (joint angles) to describe each exercise in an optimal way (Figure 3.). These scalar features were discussed with the physiotherapist and confirmed to be sufficient for capturing the essence of each exercise. For example, in shoulder abduction, the shoulder angle (θ_1) reflects the range of motion and arm angle (θ_2) indicates if the arm is being stretched or not. Except for the lunge exercise, which needed three joint angles, all other exercises could be described with only two angles. A sample plot of extracted features is shown in Figure 4.

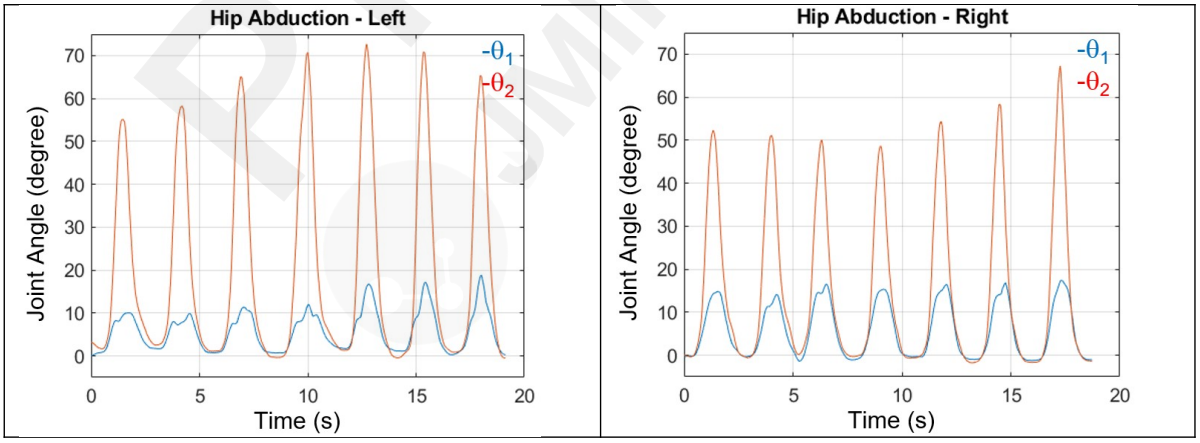


Hip abduction | Lunge



Shoulder abduction | Sit-to-stand

Figure 3. Extracted features (joint angles) from the joint 3D positions for each type of exercise. Both 2D side view and 3D perspective view are provided for clarity.



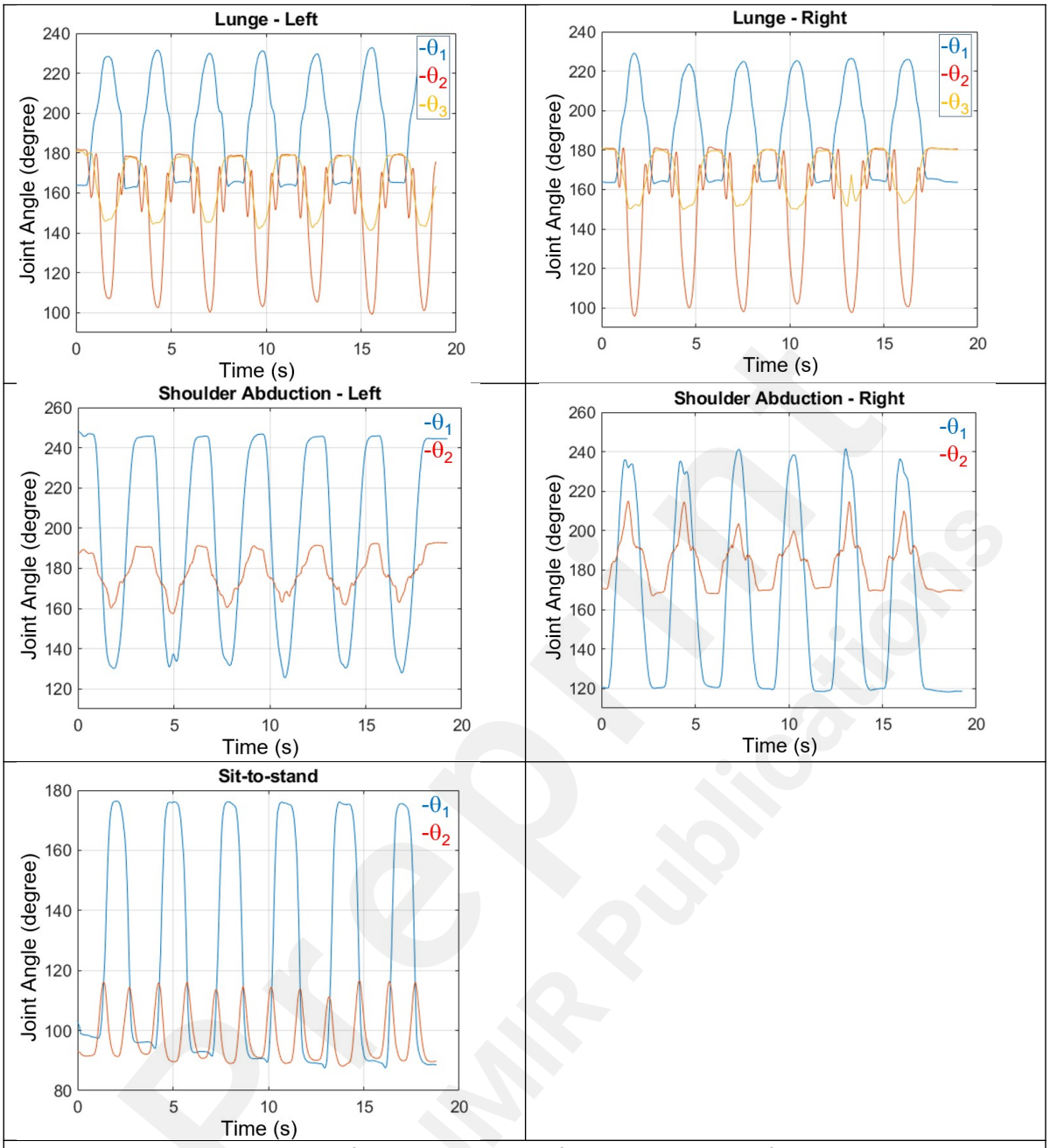


Figure 4. Sample extracted features obtained from 3D position for all seven exercises: hip abduction left (hpl) and right (hpr), lunge left (lul) and right (lur), shoulder abduction left (shl) and right (shr), and sit-to-stand (sit).

A motion trajectory T(l) is formed by the sequence of feature values within the time frame $0 \leq t \leq l$, with l being the execution time. T(l) is a matrix of size l×3 for the lunge exercise and l×2 for the rest.

$$T(l)=\begin{bmatrix} \theta_1(0), \theta_2(0) \\ \theta_1(1), \theta_2(1) \\ \vdots \\ \theta_1(l), \theta_2(l) \end{bmatrix}$$

Methods

Dynamic Time Warping

Dynamic Time Warping (DTW) [41] is a technique to align two time series and find the minimum Euclidean distance between them. It is a frequently used approach in speech recognition to classify sound waves of the same word spoken in different accents and durations. DTW is sensitive to both signal pattern and amplitude. If two signals have the same patterns, for example, the same number of peaks, but different amplitude, then the alignment cannot be done perfectly, thus yielding a large distance between them. If they have the same amplitude, but different patterns, the alignment will also result in a large distance. Therefore, the output distance is a measure of similarity between two time series. The higher the distance, the greater the deviation.

Whilst DTW was initially applied to speech recognition, it has also been widely used in gesture recognition [10, 12, 42]. In these studies, the motion trajectories, each representing a gesture, are classified into the most similar gesture group (i.e. the one with the smallest distance) by converting the distance between two trajectories into a similarity measure.

We define $D_{RP}=DTW(T_R,T_P)$ as a distance measure between the reference (T_R) and participant (T_P) trajectories. The Matlab function *dtw* was used for the implementation. While the lower limit (D_l) of this distance is zero (being perfectly similar), the upper limit is unknown and can take any large value. By estimating an upper limit, it is possible to convert the distance measure into a similarity score. Associating the upper bound with the worst possible performance, an upper limit can be approximated. For all the exercises, not moving would be the simplest worst performance. We refer to this worst trajectory as T_W. Calculating $D_{RW}=DTW(T_R,T_W)$ allows us to establish the upper bound. Knowing both lower and upper bounds, a given distance $D_l=0\leq D\leq D_u$ can be transformed into a similarity measure or percentage score $0\leq S_D\leq 100$:

	$S_D = 100 \times \frac{D - D_l}{D_u - D_l} = 100 \times \frac{D}{D_u}$	(1)

Hidden Markov Modeling

A Hidden Markov Model (HMM) [43] is a stochastic model that considers an observed signal as the result of the transition of a system between several states, each of which has a probability that a particular symbol might be observed. HMMs are useful for recognition of temporal patterns such as speech, handwriting, and gestures. An HMM with discrete observations is mainly specified by the state transition matrix A and the observation matrix B. Assuming that the system goes through N different possible states S_1, S_2, ..., S_N and in each state one of M different symbols v_1, v_2, ..., v_M can be observed (Figure 5).

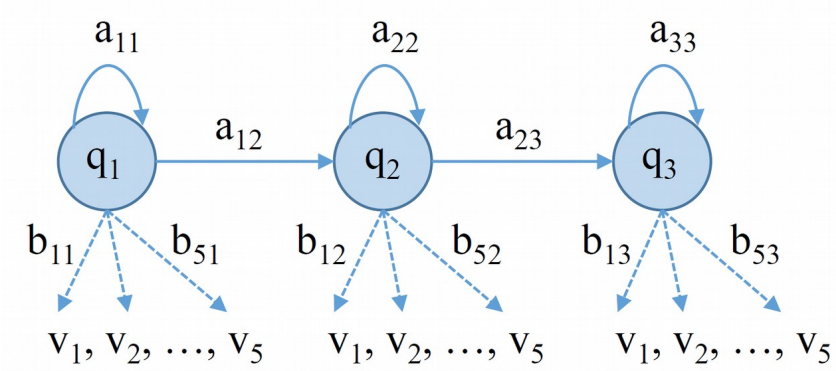


Figure 5. HMM with three states (q_1 to q_3) and five observation symbols (v_1 to v_5). The relationship among states is described by transition matrix $A=[a_{ij}]_{3x3}$, and between states and discreet symbols by observation matrix $B=[b_{ij}]_{3x5}$. It is assumed the system evolves through certain states whose relationship is to be studied.

For performance evaluation, a single HMM, λ_R, is trained based on the reference motion trajectory T_R. We then calculate log-likelihood of T_P given the trained model by $L_{RP}=\log(P(T_P|\lambda_R))/l_P$. Similar to DTW, lower and upper limits of the log-likelihood need to be calculated. The upper limit (L_u) is known and is equal to zero, since the highest probability is one. However, the lower limit is unknown and can be any small value less than zero. Same as before, we assumed that this lower limit reflects the worst possible performance captured by T_W. Letting the lower limit be $L_{RW}=\log(P(T_w|\lambda_R))/l_w$, the similarity score $0\leq S_H\leq 100$ corresponding to log-likelihood $L_l\leq L\leq L_u=0$ is obtained by:

$$S_H = 100 \times \frac{L - L_l}{L_u - L_l} = 100 \times \frac{L_l - L}{L_l} \quad (2)$$

Experimentation

Setup

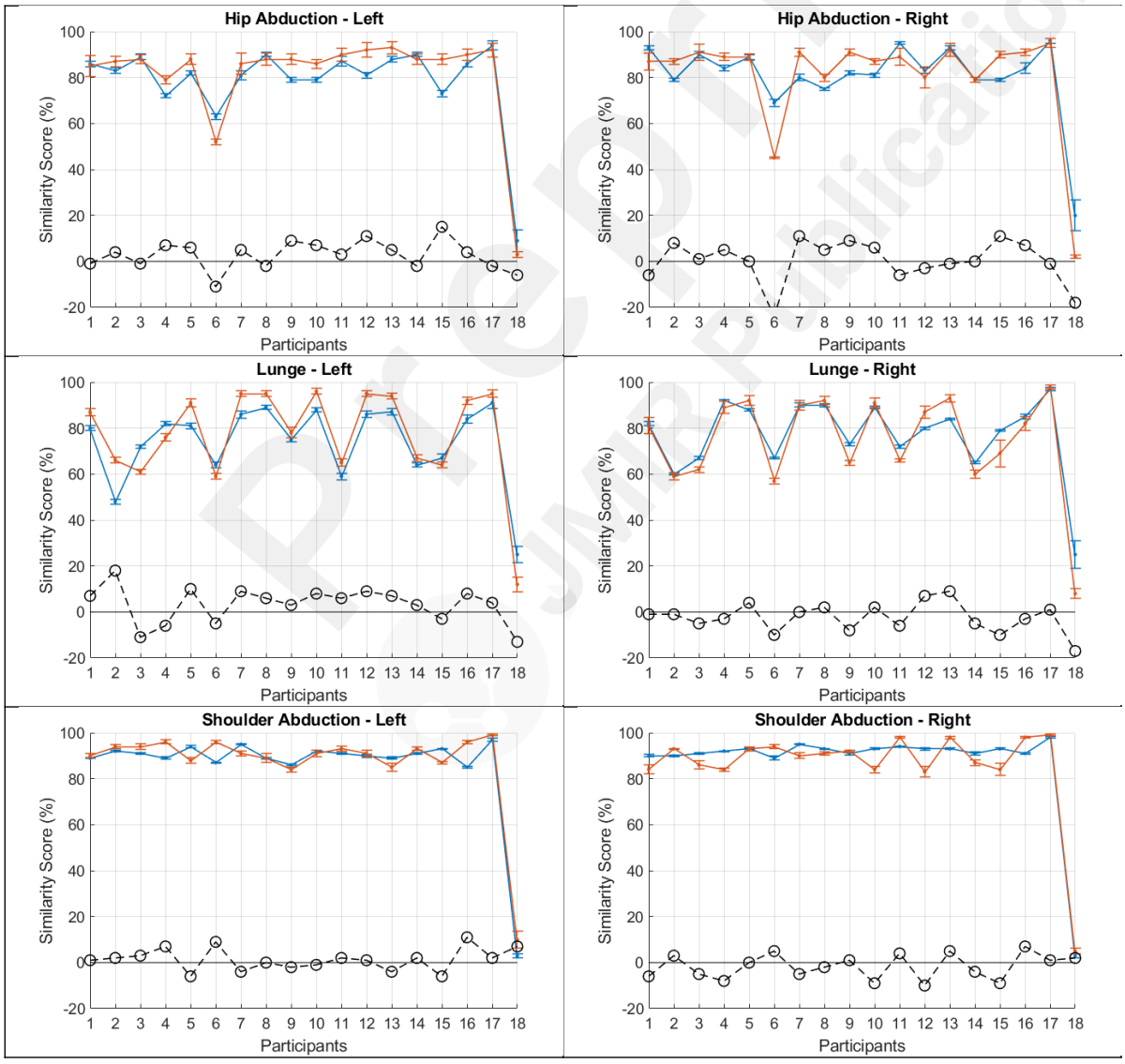
Standard hardware (Computer, TV, Kinect sensor) was used in combination with a Unity program to display the participant's live performance on the screen and store the 3D position data along with a time stamp. As mentioned above, four types of exercises were chosen to be performed by the participants: shoulder abduction, hip abduction, lunge, and sit-to-stand. Except for the sit-to-stand, all other exercises were performed for both left and right side, making the total number of exercises seven.

Participants

16 healthy participants including 8 adult females (22-30 years), 6 adult males (22-40 years), and 2 schoolboys (12 and 17 years), were recruited for the study. Participants were asked to stand in front of the Kinect sensor and perform each of the seven exercises for 20 seconds. They were told to repeat the chosen exercise at-least five times with a short pause between each repetition. In addition to the 16 participants, the physiotherapist involved in the project (DS) was asked to perform the exercises as the reference performance. He repeated each of the seven exercises at-least five times during a period of 20 seconds each.

Results and Discussion

Out of five repetitions, three were extracted (ignoring the first and the last) for each participant. For DTW scores, the distance between each repetition of a participant and each repetition of the physiotherapist is calculated yielding to a total number of nine values. Final DTW score (S_D) is obtained by taking average among them. For HMM scores, the likelihood of each repetition of a participant given the physiotherapist's model is calculated yielding to a total number of three values. Final HMM score (S_H) is obtained by taking average among them. Figure 6. shows the similarity scores obtained applying DTW (S_D, solid blue line) and HMM (S_H, solid red line) for each exercise. The difference between the two scores, S_D-S_H, is also given in the plots by a dashed black line. Participant 17 is the physiotherapist and hence the reference. As expected, he obtained the maximum score in all cases. The worst performance, i.e. making no movement, is included as participant 18 yielding to very low scores in all exercises. Ideally, we would have expected to obtain the maximum score of 100% for the reference and the minimum score of 0% for the worst performance. However, those values were not obtained due to natural variability among the repetitions and cross-comparison between the repetitions.



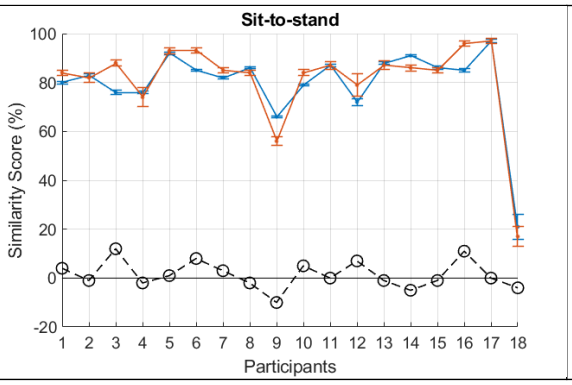


Figure 6. Similarity scores obtained from applying DTW (S_D, solid blue line) and HMM (S_H, solid red line). Their difference, S_H-S_D, is shown by the black dashed line. Participant 17 is the physiotherapist and 18 is the worst performance, i.e. making no movement. Error bars indicate standard error.

Several observations can be made from the plots. Both scores show similar trends. Whenever a participant is not performing well according to DTW score (S_D), his/her HMM score (S_H) is low as well. Examples of this behavior are participant 6 in hip abduction (left and right) and participant 9 in sit-to-stand.

Regarding the difference, S_H-S_D, it is hard to observe any obvious pattern. On occasions, the difference is positive and in others is negative. There are also cases that the difference is negligible. It can be as large as +18% (participant 2, lunge-left) or as small as -24% (participant 6, hip abduction-right). For the reference, participant 17, the difference is generally very low (-1%, -1%, 4%, 1%, 2%, 1%, and 0%). However, this is not the case for the worst performance, participant 18. In most exercises, the difference is negative and is not negligible, meaning that S_D is usually larger than S_H for the worst performance. In addition, in most plots S_H is closer to 0% than S_D for the worst performance.

Among the exercises, shoulder abduction was less challenging and easy to perform, which is reflected by participants performing well and achieved high similarity scores. In contrast, lunge was the most difficult and demanding exercise to perform, which is also reflected in the obtained similarity values.

Time-domain plots of the best and worst performances can be used to visually examine the correlation between the trajectories and the calculated scores. For each of seven exercises, the trajectories of the best and the worst performances are plotted against the reference (Figure 7). Each plot illustrates three repetitions of the chosen participant (solid red lines) and three repetitions of the physiotherapist (dashed blue lines). It should be noted that for hip/shoulder abduction and sit-to-stand, each repetition includes two features and for lunge three features (see Figure 4). Evidently, whenever the participant's trajectories match with those of the reference (in terms of both amplitude and duration), the obtained score is high. Conversely, when the patterns do not match, the score is low. For example, in hip abduction – left, participant 6 failed to achieve full range of motion (compare the peak amplitude 40° vs. 80°) and performed faster (compare the duration, 100 (1s) to 200 (2s)). It is worth mentioning that both DTW and HMM algorithms are sensitive to amplitude than duration. They are both time-series algorithms which take into account sequence of data (amplitudes) regardless of their frequency (duration). DTW does not allow time-scaling of members within the sequence and the HMM algorithm discards time dependency of members by grouping them into clusters during quantization pre-process. If two participants follow the required range of motion, but one performed slower (or faster) than the other, their score would be same (see lunge –

left, both participants 11 and 9).

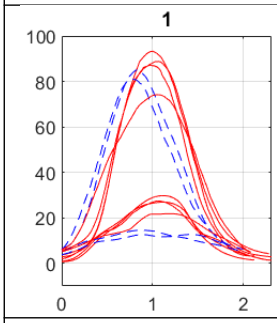
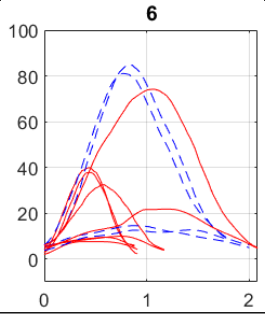
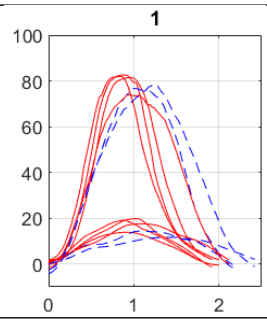
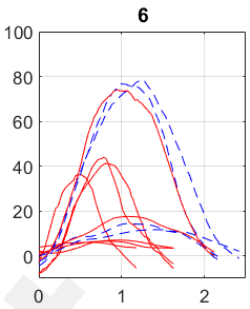
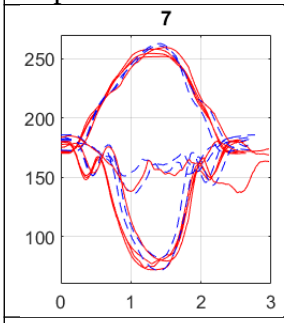
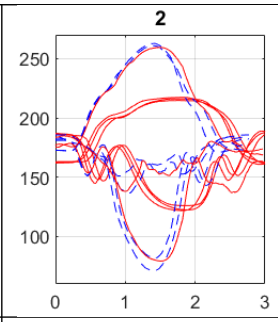
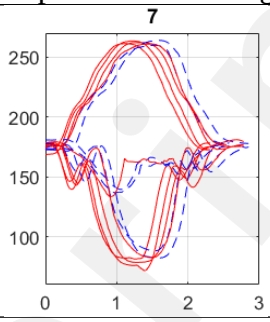
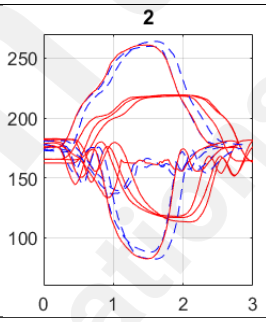
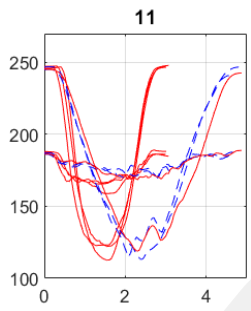
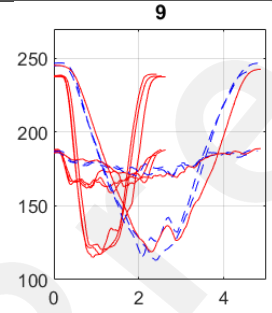
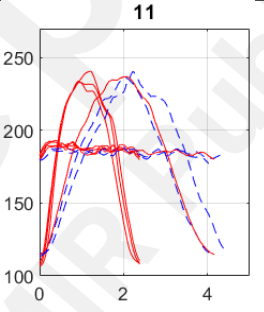
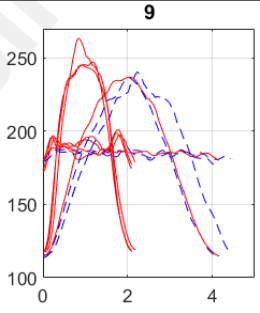
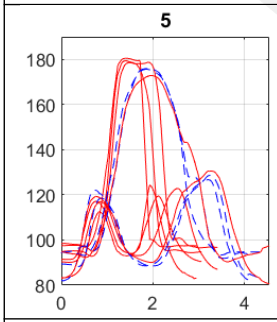
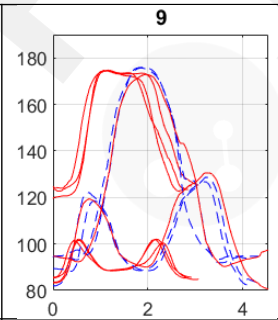
Best	Worst	Best	Worst
			
S_D=86%, S_H=85%	S_D=63%, S_H=52%	S_D=93%, S_H=87%	S_D=69%, S_H=45%
Hip abduction - Left		Hip Abduction - Right	
			
S_D=86%, S_H=95%	S_D=48%, S_H=66%	S_D=90%, S_H=90%	S_D=60%, S_H=59%
Lunge - Left		Lunge - Right	
			
S_D=91%, S_H=93%	S_D=86%, S_H=84%	S_D=94%, S_H=94%	S_D=91%, S_H=93%
Shoulder Abduction - Left		Shoulder Abduction - Right	
			
S_D=92%, S_H=93%	S_D=66%, S_H=56%		
Sit-to-Stand			

Figure 7. Time domain plots of the best and worst trajectories (excluding participants 17 and 18). X-axis indicates time in seconds and Y-axis indicates joint angle in degree. The three repetitions of the reference trajectories are given in dashed blue lines and the chosen participant's trajectories in solid red lines. All plots include two features, except for lunge where three features are presented.

Except for the lunge for which three features (joint angles) are extracted, two features were obtained for all the other exercises. However, not all of the extracted features have the same weight and importance. For instance, in hip abduction, referring to Figure 4 and 7, θ_2 (hip angle) is the more relevant feature and θ_1 (torso angle) having less importance. This is the opposite for shoulder abduction and sit-to-stand. For lunge, θ_2 and θ_1 are both equally important, but θ_3 less so. The main feature has higher influence on the obtained similarity score. This is shown in Figure 8. for hip abduction – left as an example. Generally, the same trend has been maintained for each measure (compare same color solid and dashed lines). However, the effect of removing a minor feature (θ_1) is less influential on S_H than S_D. The reason is that DTW is more sensitive to the detail than HMM.
While single-feature S_D is clearly larger than full-feature S_D, single-feature S_H is almost same as full-feature S_H. Adding more details, i.e. presenting additional minor features, increases the distance values (and hence decreases the similarity scores) obtained by applying DTW.

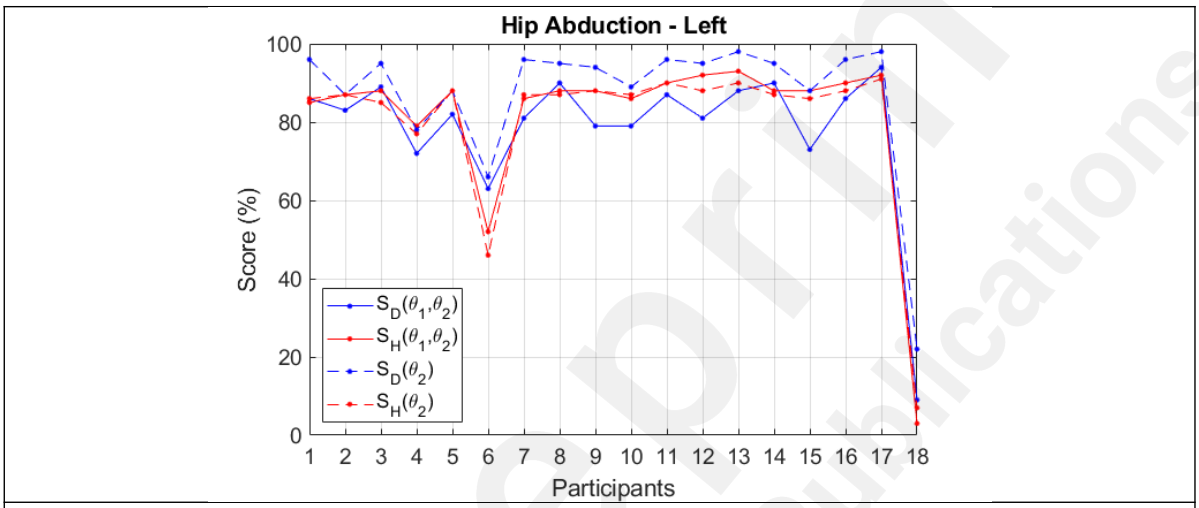


Figure 8. The effect on the scores of removing minor features.

While DTW is more sensitive to the detail, HMM is more sensitive to the way the feature space is quantized. Quantization is a pre-process applied over the extracted features to segment them into several clusters for the purpose of training a discrete HMM. The boundaries and the number of clusters have an obvious effect on the HMM scores. This can be seen in Figure 9 for shoulder abduction – right as an example. Bi-variate histogram plots are used to visualize the clusters. Three cases have been tested changing θ_1 (main feature) boundaries and keeping θ_2 (minor feature) intact. In case A, θ_1 is divided into three clusters (90°, 180°, and 270°), in B four clusters (90°, 150°, 210°, and 270°) and in C five clusters (90°, 135°, 180°, 225°, and 270°). In all cases, θ_2 is divided into four clusters (90°, 180°, and 270°). A closer look reveals that the S_H scores in case B are clearly lower than A and C, which are more similar. The reason for this is that feature points are grouped into different clusters for the participant and the reference. This affects (reduces) the probability that a selected feature point of the participant is generated by the reference model. Most likely, the feature points of the participants and the reference in cases A and C are grouped into similar clusters, hence their similarity score is higher and more similar. In case B, however, they are grouped into different clusters, decreasing the probability values, hence lowering the similarity scores. One can take advantage of this behavior by adjusting the sensitivity of HMM similarity scores to smooth or sharpen the differences.

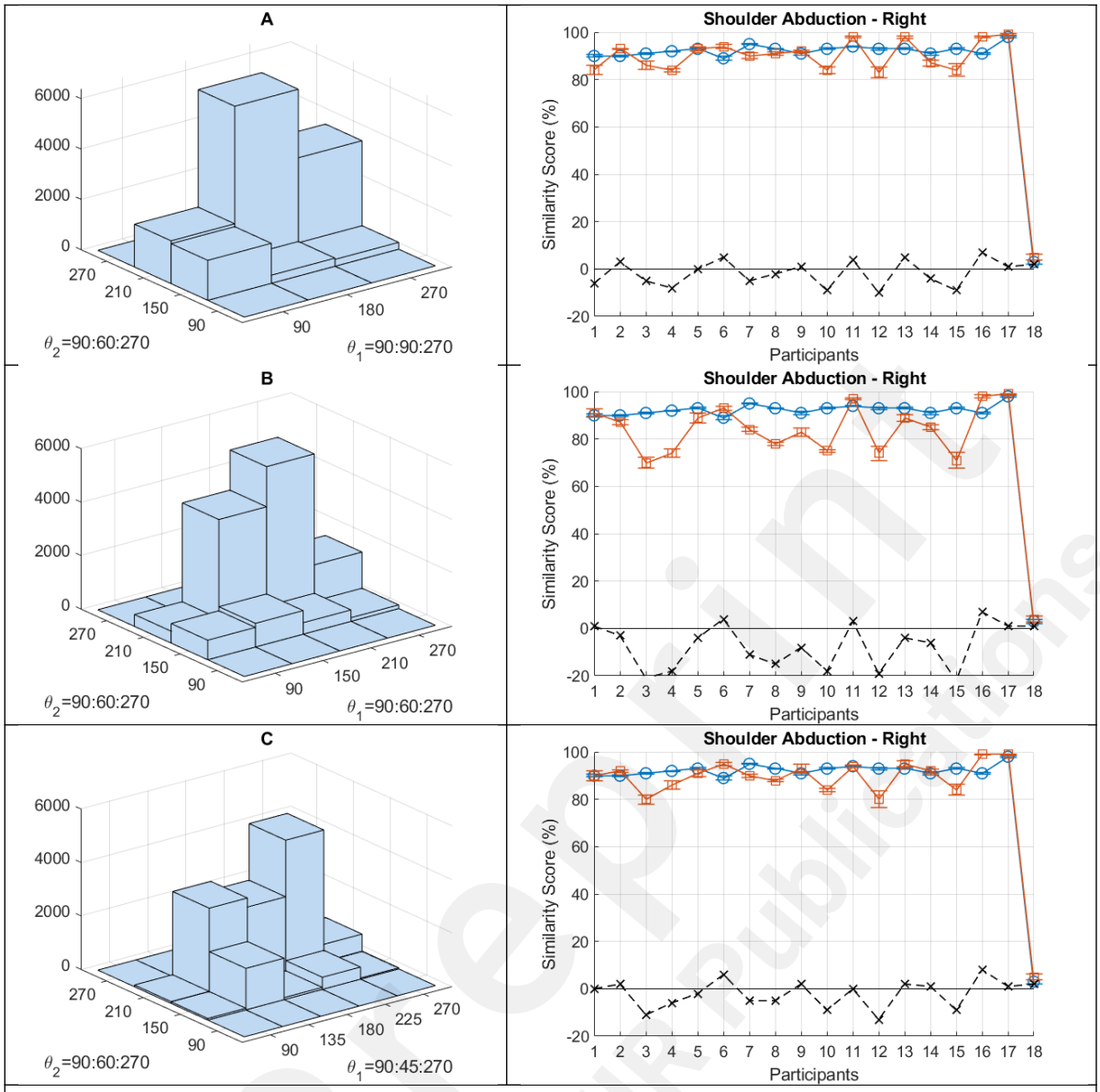


Figure 9. The effect of quantization on the calculated HMM scores.

Worst performance is a key factor affecting the scores in both measures as it corresponds to the upper or lower boundary as previously explained. This is evident referring to equations 1 and 2. Both unknown limits (D_u for DTW and L_l for HMM) are used as denominator to normalize the distance (eqn. 1) and likelihood (eqn. 2) values. The larger the denominator, the smaller the deviations (fluctuations) in the scores. This value can also be intentionally altered to adjust the sensitivity of the scores. With a larger denominator, the scores are smoother and the differences between participants get smaller. With a smaller denominator, the scores get sharper and the differences between participants are highlighted. As explained before, we have chosen no movement as the worst performance for all exercises. Seemingly, a different worst performance can yield different scores if it generates different denominators. For example, one might say closing the elbow in shoulder abduction could be worse than keeping it stretched (the current situation). An example of altering limits (multiplying and dividing D_u and L_l by 2) for sit-to-stand is shown in Figure 10. When the limits are magnified (multiplied by 2), the scores are increased and smoothed (Figure 10, left). When the limits are shrunk (divided by 2), the scores are decreased and sharpened (Figure 10, right). Smoothed scores can be used to encourage patients performing their prescribed exercises in early stages, while sharpened scores could be used in later stages to encouraging further mastering of the

skills involved in performing the exercises.

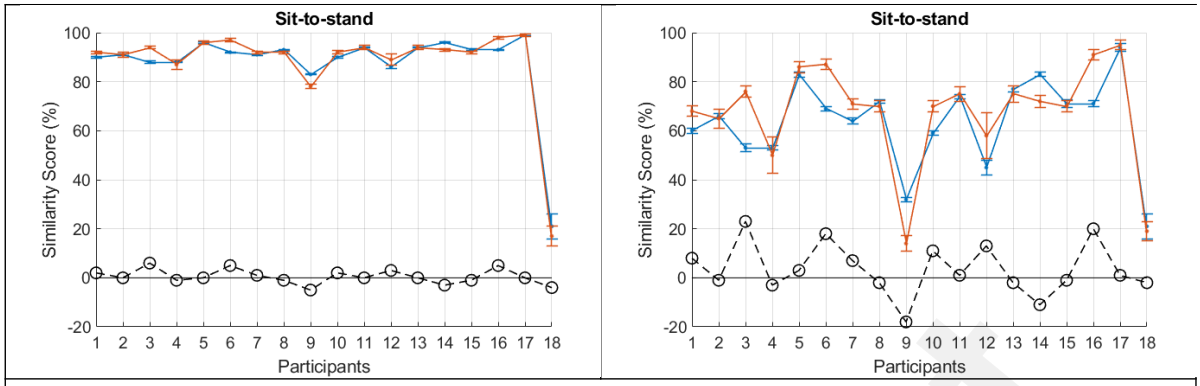


Figure 10. Effect of altering limits (D_u and L_l, obtained from the worst performance) on the scores: multiplied by 2 (left) and divided by 2 (right). Error bars indicate standard error.

Several comparative tests were also conducted. Figure 11 shows the average scores across all seven exercises. For each participant, the average score is obtained from 9×7=72 scores (9 values for each of 7 exercises). As it can be seen, the two plots show very similar trend. This is verified applying t-test over the two measures (p-value=0.49). Small error bars are an indication of high levels of consistency.

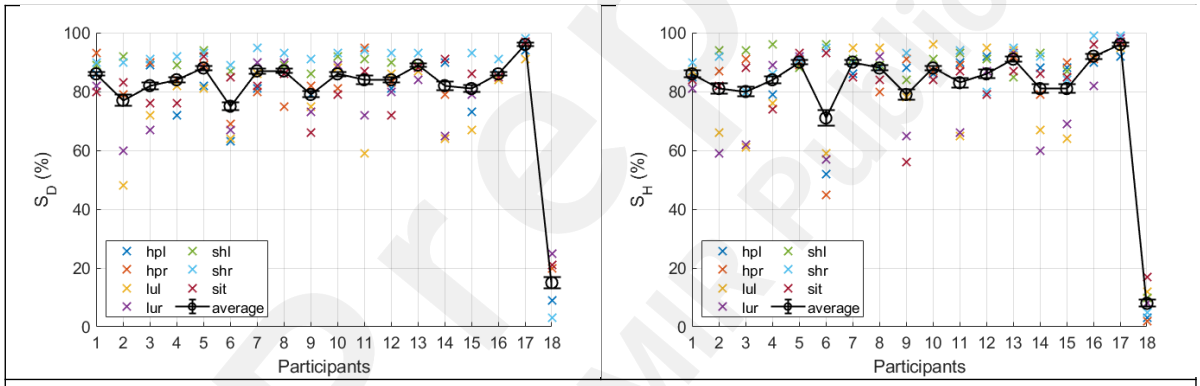


Figure 11. The combined and averaged scores (left: S_D and right: S_H) for all seven exercises. Error bars indicate the standard error.

A comparison between left and right performances, excluding sit-to-stand, is given in Figure 12. No significant differences can be observed between the left and the right scores. This was verified by applying a t-test over the scores for each measure (p-values are given in Table 2).

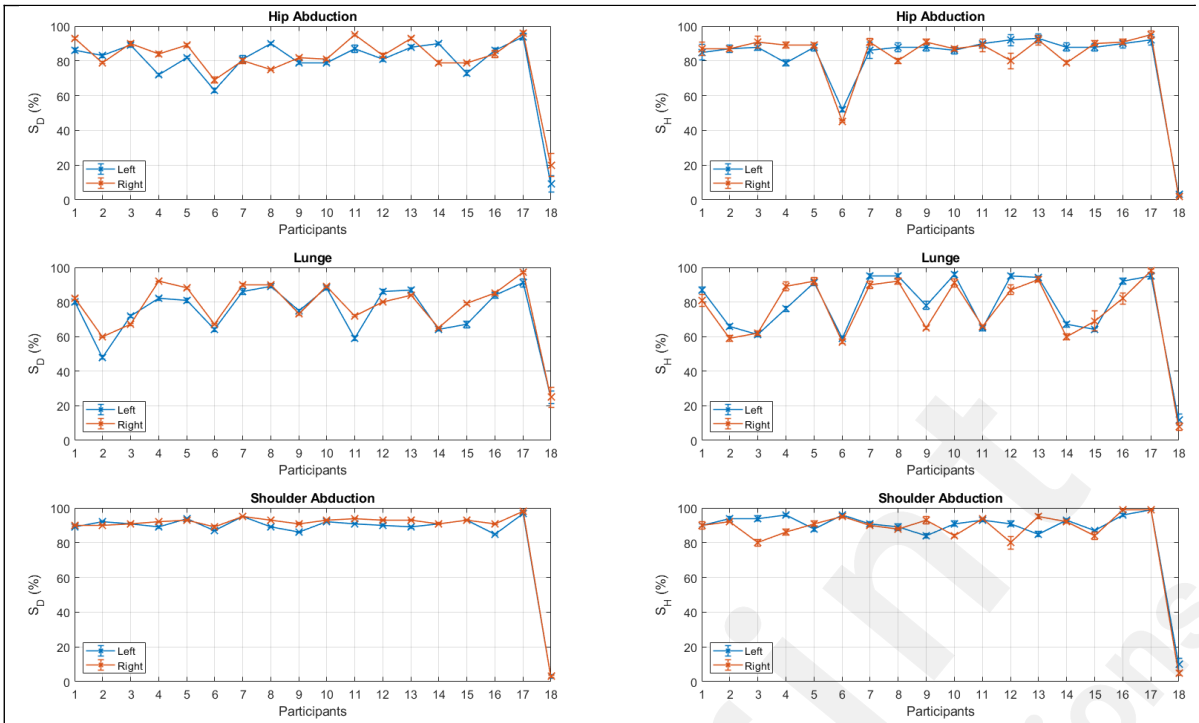


Figure 12. Comparison between left and right scores. Error bars indicate standard error.

Table 2. P-values obtained applying t-test over left and right scores.

	Hip abduction	Lunge	Shoulder abduction
S_D	0.27	0.09	0.44
S_H	0.85	0.25	0.45

Furthermore, a comparison between female and male participants is given in Figure 13. The scores are averaged across 8 female and 8 male participants (excluding reference). As it can be seen, there is no significant difference between the two genders. This is verified by applying a t-test over S_D (p-value=0.22) and S_H (p-value=0.86).

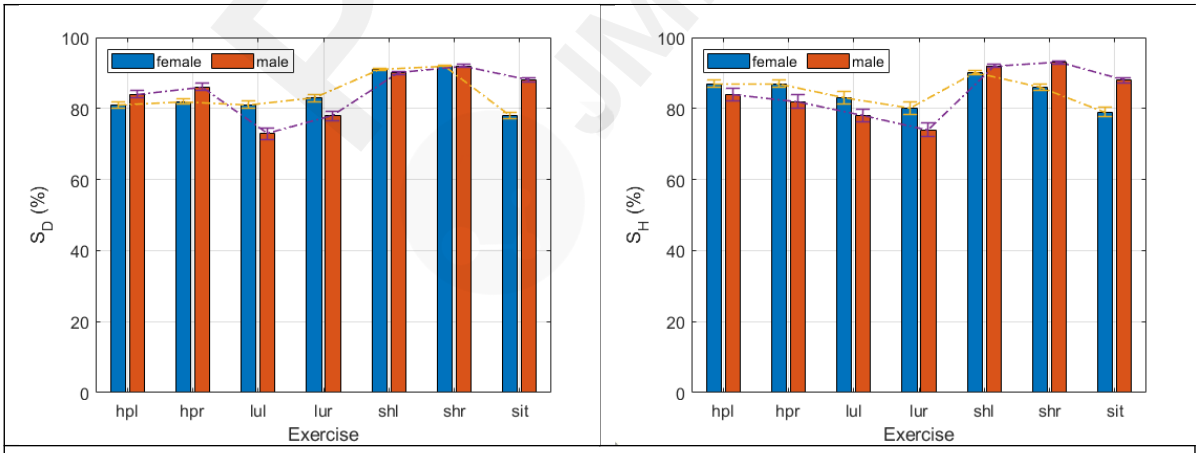


Figure 13. Comparison between female and male participants, left: DTW scores (S_D) and right: HMM scores (S_H). Error bars indicate standard error.

Preprint
JMIR Publications

Conclusions and Future Remarks

We implemented and compared two commonly used machine learning algorithms, DTW and HMM, to objectively evaluate the performance of patients using a rehabilitation exergaming platform. 3D movement data was obtained using the Kinect depth camera and invariant features (joint angles) describing each exercise were extracted. The extracted features are independent of body fit, size, and the position and distance of the user to the Kinect. They are also independent of the hardware being used and can be adapted for any motion sensing device able to track human skeleton joints such as those mentioned in the Related Work section.

Setting a physiotherapist performance as the “reference” and making no movement as the “worst performance”, we applied both DTW and HMM algorithms to compare participants performance and report a similarity score. The idea of worst performance was the key to convert the distance measures, obtained from DTW, and likelihood values, obtained from HMM, to similarity scores between 0% and 100%. Overall, both algorithms showed similar trends, but with different sensitivity. DTW was observed to be more sensitive to small changes, whilst HMM was more sensitive to the boundaries and clusters resulting from the quantization process. Both DTW and HMM are inherently more sensitive to range of motion than duration. In addition, both measures are sensitive to the worst performance. This suggests ways to utilize both algorithms to monitor a patient progress at different stages: monitoring could start with HMM similarity scores in early stages for more general comparison, switching to DTW similarity scores in later stages for finer comparison.

The application of these similarity scores is twofold. The scores can be used by the patients at home to encourage them to keep practicing the Exergames to achieve higher similarity scores. In addition, the scores can be reported back to the physiotherapist to monitor patient progress and provide feedback. The exercise program can also be adjusted by the physiotherapist given the level of progress to better fit the patient’s needs and progression.

Our proposed method has the potential for significant impact in the context of rehabilitation exergames by enabling remote therapy home-based sessions where performance can still be adequately monitored. This can help better assess the quality of physical exercises performed by patients, fine tune rehabilitation programmes and enhance the efficiency of home-based rehabilitation. In turn, cost reductions and freeing up of physiotherapy units time may also be achieved.

Future work will include testing our proposed system on a public dataset such as UTD-MHAD [44]. In addition, we intend to recruit patients with reduced movement range or other constraints to participate in a study to validate our proposed performance metrics in a clinical setting. The study will also investigate the correlation between the objective scores used by our system and the subjective notes taken by physiotherapists whilst observing patients performing the exergames.

Acknowledgements

The project was funded through MedCity’s Collaborate to Innovate programme and received funding from the European Regional Development Fund and the Higher Education Funding Council for England. We also acknowledge the support of MIRA Rehab Ltd as well as members of the Simulation and Modelling in Medicine and Surgery (SiMMS) research group at Imperial College London.

JMIR Preprints

Conflicts of Interest

None declared.

Abbreviations

JMIR: Journal of Medical Internet Research
RCT: randomized controlled trial
MIRA: Medical Interactive Recovery Assistant
DTW: Dynamic Time Warping
HMM: Hidden Markov Model

References

1. [MIRA] “MIRA Rehab.” [Online]. Available: www.mirarehab.com (last accessed 31/10/2019)

2. “VirtualRehab.” [Online]. Available: evolvrehab.com/virtualrehab/ (last accessed 31/10/2019)

3. “REHABILITY.” [Online]. Available: rehability.me (last accessed 31/10/2019)

4. Whitehead, Anthony, et al. "Exergame effectiveness: what the numbers can tell us." Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games. 2010.

5. Almeida, João, and Francisco Nunes. "The Practical Work of Ensuring the Effective Use of Serious Games in a Rehabilitation Clinic: Qualitative Study." JMIR Rehabilitation and Assistive Technologies 7.1 (2020): e15428.

6. Bossen, Daniël, et al. "Effectiveness of Serious Games to Increase Physical Activity in Children With a Chronic Disease: Systematic Review With Meta-Analysis." Journal of Medical Internet Research 22.4 (2020): e14549.

7. Meekes, Wytske, and Emma Kate Stanmore. "Motivational determinants of exergame participation for older people in assisted living facilities: mixed-methods study." Journal of medical Internet research 19.7 (2017): e238.

8. Tobaigy, Abdullah, et al. "The feasibility of using exergames as a rehabilitation tool: the attitudes, awareness, opinions and experiences of physiotherapists, and older people towards exergames." Journal of physical therapy science 30.4 (2018): 555-562.

9. Moni, M. A., and ABM Shawkat Ali. "HMM based hand gesture recognition: A review on techniques and approaches." In 2009 2nd IEEE International Conference on Computer Science and Information Technology, pp. 433-437. IEEE, 2009.

10. Ibanez, Rodrigo, Alvaro Soria, Alfredo Teyseyre, and Marcelo Campo. "Easy gesture recognition for Kinect." Advances in Engineering Software 76 (2014): 171-180.

11. Antón, David, Alfredo Goni, and Arantza Illarramendi. "Exercise recognition for Kinect-based telerehabilitation." Methods of information in medicine 54, no. 02 (2015): 145-155.

12. Raheja, J. L., M. Minhas, D. Prashanth, T. Shah, and A. Chaudhary. "Robust gesture recognition using Kinect: A comparison between DTW and HMM." Optik 126, no. 11-12 (2015):

1098-1104.

13. Calin, Alina Delia. "Gesture recognition on Kinect time series data using dynamic time warping and hidden Markov models." In 2016 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), pp. 264-271. IEEE, 2016.

14. A. D. Calin, H. F. Pop, and R. F. Boian, “Improving movement analysis ˇ in physical therapy systems based on kinect interaction,” in Proceedings of the 31st British Computer Society Human Computer Interaction Conference. BCS Learning & Development Ltd., 2017, p. 87.

15. H. Mousavi Hondori and M. Khademi, “A review on technical and clinical impact of microsoft kinect on physical therapy and rehabilitation,” Journal of medical engineering, vol. 2014, 2014.

16. Chang, Yao-Jen, Wen-Ying Han, and Yu-Chi Tsai. "A Kinect-based upper limb rehabilitation system to assist people with cerebral palsy." Research in developmental disabilities 34, no. 11 (2013): 3654-3659.

17. N. Duarte, O. Postolache, and J. Scharcanski, “Ksgphysio-kinect serious game for physiotherapy,” in Proceedings of the International Conference and Exposition on Electrical and Power Engineering (EPE). IEEE, 2014, pp. 606–611.

18. F. Cary, O. Postolache, and P. Girao, “Kinect based system and serious game motivating approach for physiotherapy assessment and remote session monitoring,” Proceedings of the International Conference on Sensing Technology (ICST), pp. 2–4, 2014.

19. C.-J. Su, C.-Y. Chiang, and J.-Y. Huang, “Kinect-enabled home-based rehabilitation system using dynamic time warping and fuzzy logic,” Applied Soft Computing, vol. 22, pp. 652–666, 2014.

20. D. Ferreira, R. Oliveira, and O. Postolache, “Physical rehabilitation based on kinect serious games,” in Proceedings of the 11th International Conference on Sensing Technology (ICST). IEEE, 2017, pp. 1–6.

21. Y. Rybarczyk, J. K. Deters, C. Cointe, and D. Esparza, “Smart web-based platform to support physical rehabilitation,” Sensors, vol. 18, no. 5, 2018.

22. “Just Dance 2018.” [Online]. Available: https://en.wikipedia.org/wiki/Just_Dance_2018 (last accessed 03/12/2019)

23. Palma, Carlos, Augusto Salazar, and Francisco Vargas. "Hmm based evaluation of physical therapy movements using kinect tracking." In International Symposium on Visual Computing, pp. 174-183. Springer, Cham, 2015.

24. Deters, J. Kleine, and Yves Rybarczyk. "Hidden Markov Model approach for the assessment of telerehabilitation exercises." International Journal of Artificial Intelligence 16, no. 1 (2018): 1-19.

25. W. Wei, Y. Lu, E. Rhoden, and S. Dey, “User performance evaluation and real-time guidance in cloud-based physical therapy monitoring and guidance system,” Multimedia Tools and

Applications, pp. 1–31, 2017.

26. J. Richter, C. Wiede, B. Shinde, and G. Hirtz, “Motion error classification for assisted physical therapy-a novel approach using incremental dynamic time warping and normalised hierarchical skeleton joint data.” in Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM), vol. 1, 2017, pp. 281–288.

27. A. Vakanski, J. M. Ferguson, and S. Lee, “Metrics for performance evaluation of patient exercises during physical therapy,” International journal of physical medicine & rehabilitation, vol. 5, no. 3, 2017.

28. Palma, Carlos, Augusto Salazar, and Francisco Vargas. "HMM and DTW for evaluation of therapeutical gestures using kinect." arXiv preprint arXiv:1602.03742 (2016).

29. Yurtman, Aras, and Billur Barshan. "Automated evaluation of physical therapy exercises using multi-template dynamic time warping on wearable sensor signals." Computer methods and programs in biomedicine 117.2 (2014): 189-207.
30. Hakim, Tal, and Ilan Shimshoni. "A-MAL: Automatic Motion Assessment Learning from Properly Performed Motions in 3D Skeleton Videos." Proceedings of the IEEE International Conference on Computer Vision Workshops. 2019.

31. Yu, Xiaoqun, and Shuping Xiong. "A Dynamic Time Warping Based Algorithm to Evaluate Kinect-Enabled Home-Based Physical Rehabilitation Exercises for Older People." Sensors 19.13 (2019): 2882.

32. Liao, Yalin, Aleksandar Vakanski, and Min Xian. "A Deep Learning Framework for Assessing Physical Rehabilitation Exercises." IEEE Transactions on Neural Systems and Rehabilitation Engineering 28.2 (2020): 468-477.

33. Jang, Minsu, et al. "Automated dance motion evaluation using dynamic time warping and Laban movement analysis." IEEE International Conference on Consumer Electronics (ICCE). 2017.

34. Laraba, Sohaib, and Joëlle Tilmanne. "Dance performance evaluation using hidden Markov models." *Computer Animation and Virtual Worlds* 27.3-4 (2016): 321-329.

35. Cantea, Andrei, Cosmin Mihaiu, Andrei Dascălu, and Alina Călin. "Mira." In Proceedings of the 31st International BCS Human Computer Interaction Conference (HCI 2017) 31, pp. 1-3. 2017.

36. Moldovan, I. M., A. D. Călin, A. C. Cantea, L. A. Dascălu, C. A. Mihaiu, O. Ghircău, S. Onac, O. Rîză, R. A. Moldovan, and L. V. Pop. "Development of a new scoring system for bilateral upper limb function and performance in children with cerebral palsy using the MIRA interactive video games and the Kinect sensor." In Proceedings of the International Conference on Disability, Virtual Reality and Associated Technologies, pp. 2-4. 2014.

37. M. McGlinchey, A. Curtis, R. Ritchie, and G. D. Jones, “Feasibility of using mira with adult patients presenting with upper limb motor dysfunction post neurological damage,” Studies in Health Technology and Informatics, vol. 217, pp. 696–702, 2015.

38. N. Czako, C. Silaghi, and R. S. Vizitiu, “Rehabilitation of a 7-year-old boy with brachial plexus palsy using mira exergames and vojta therapy,” in Proceedings of the E-Health and Bioengineering Conference (EHB). IEEE, 2017, pp. 349–352.

39. Moldovan, I. M., L. Tric, R. Ursu, A. Podar, A. D. Călin, A. C. Cantea, L. A. Dascălu, and C. A. Mihaiu. "Virtual rehabilitation programme using the MIRA platform, Kinect and Leap Motion sensors in an 81 years old patient with ischemic stroke." In 2017 E-Health and Bioengineering Conference (EHB), pp. 325-328. IEEE, 2017.

40. “Unity Real-Time Development Platform.” [Online]. Available: unity.com (last accessed 03/12/2019)

41. H. Sakoe and S. Chiba, “Dynamic programming algorithm optimization for spoken word recognition,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 26, no. 1, pp. 43–49, 1978.

42. M.-C. Hu, C.-W. Chen, W.-H. Cheng, C.-H. Chang, J.-H. Lai, and J.-L. Wu, “Real-time human movement retrieval and assessment with kinect sensor,” IEEE Transactions on Cybernetics, vol. 45, no. 4, pp. 742–753, 2015.

43. Rabiner, Lawrence R. "A tutorial on hidden Markov models and selected applications in speech recognition." Proceedings of the IEEE 77, no. 2 (1989): 257-286.

44. Chen, Chen, Roozbeh Jafari, and Nasser Kehtarnavaz. "UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor." IEEE International conference on image processing (ICIP). 2015.

Preprint JMIR Publications

Supplementary Files

Preprint JMIR Publications

Figures

Preprint JMIR Publications

MIRA system including a Kinect motion sensor and software to match an exercise with a game. A child (left) and an elderly gentleman (right) playing the Atlantis game. The child is practicing an arm exercise, the gentleman a hip exercise.

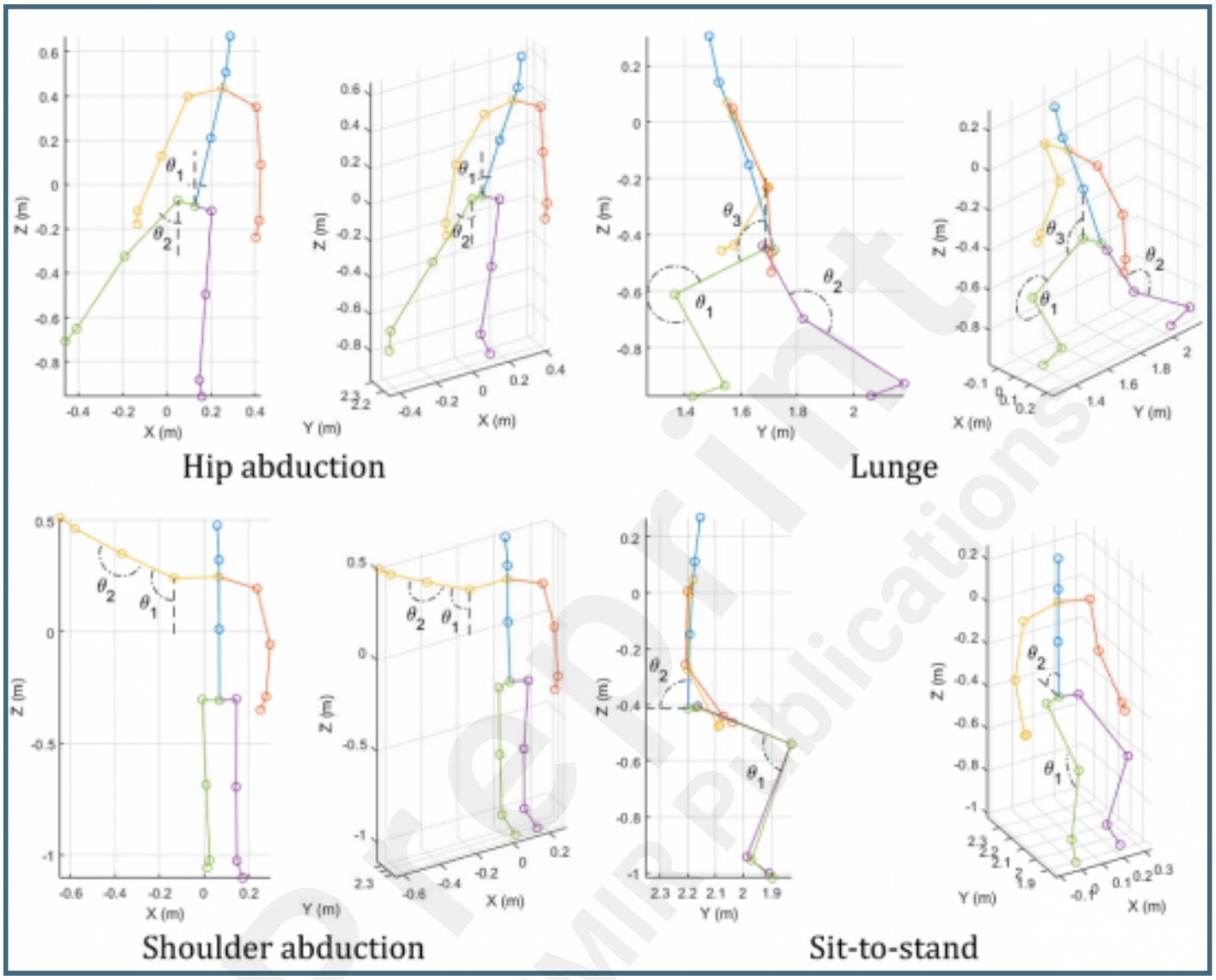


Preprint JMIR Publications

A snapshot of the MIRA program. Exercises are listed on the left and games are shown on the right. Multiple game options allow practicing the same exercise with different games, thus encouraging patients to cope with the prescribed exercise by discovering the various game scenarios.

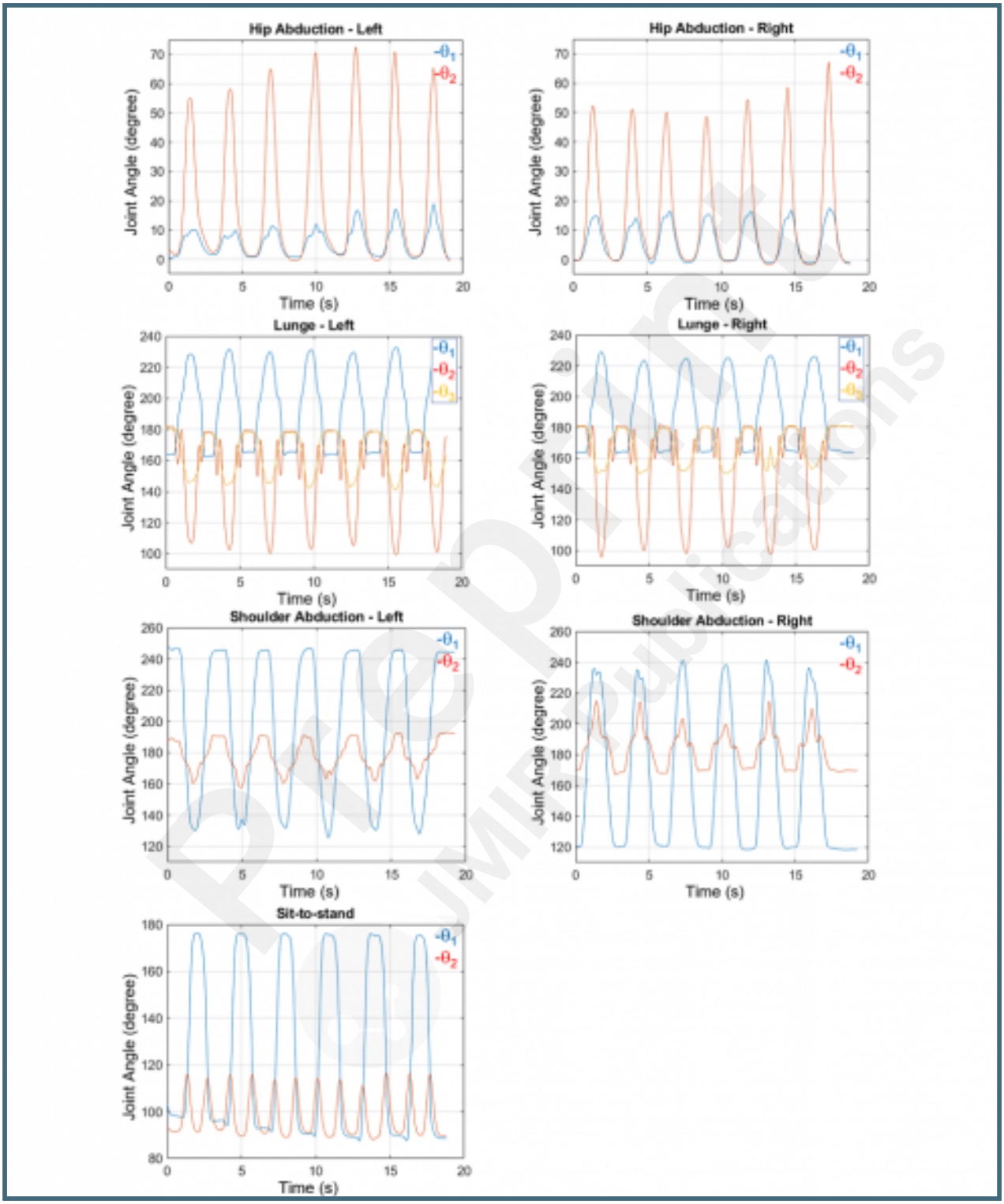


Extracted features (joint angles) from the joint 3D positions for each type of exercise. Both 2D side view and 3D perspective view are provided for clarity.



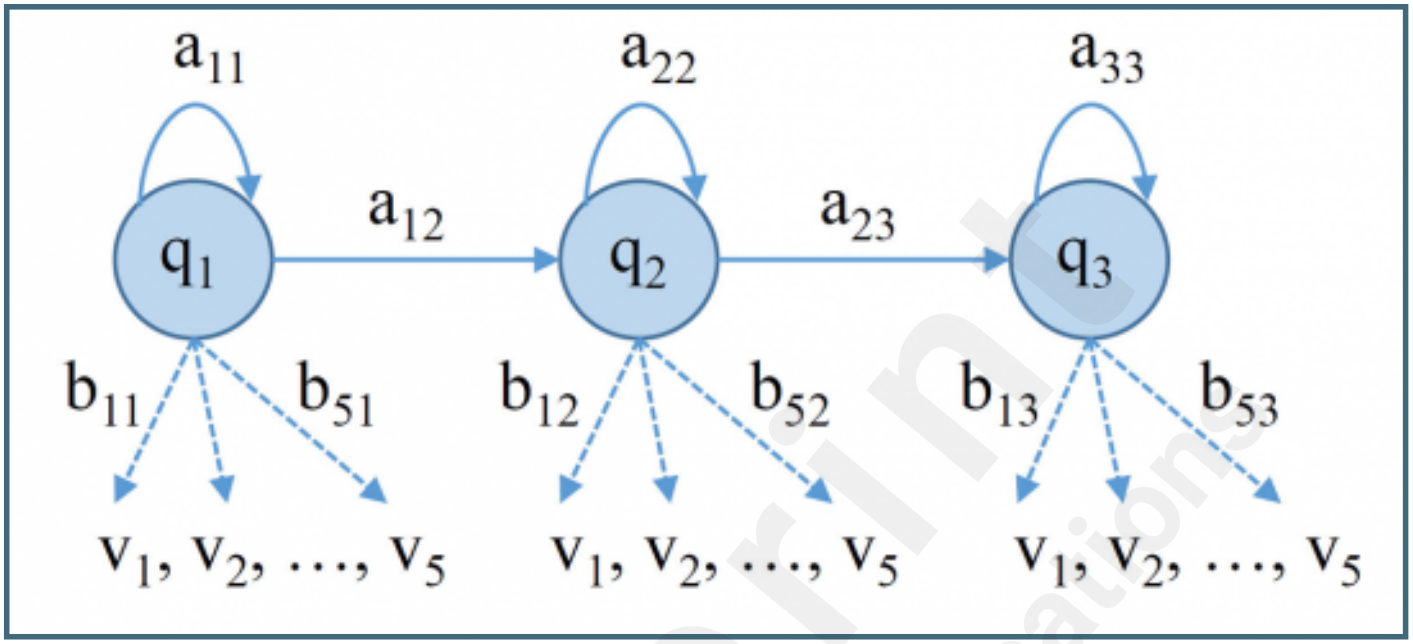
Preprint JMIR Publications

Sample extracted features obtained from 3D position for all seven exercises: hip abduction left (hpl) and right (hpr), lunge left (lul) and right (lur), shoulder abduction left (shl) and right (shr), and sit-to-stand (sit).



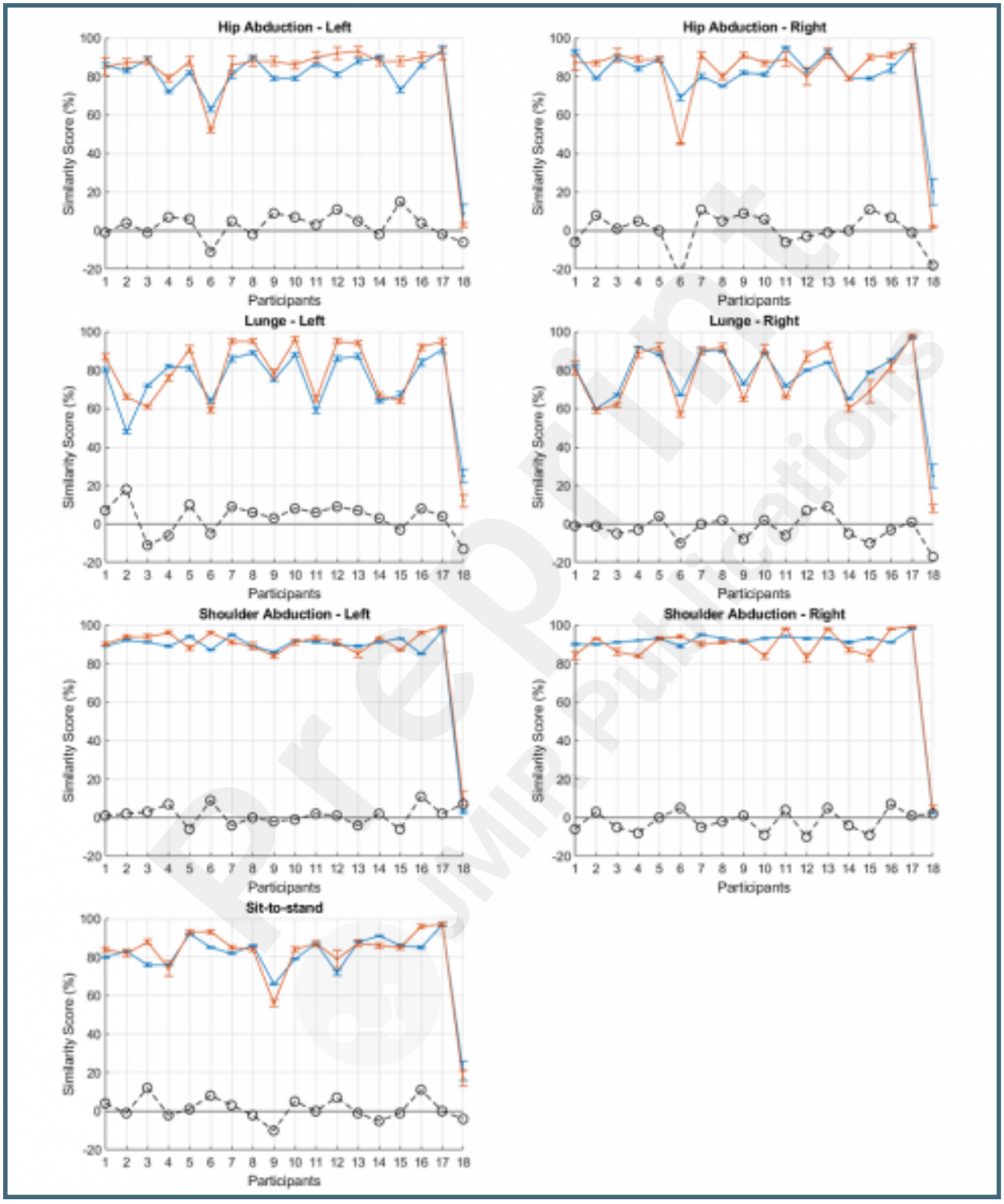
JMIR Publications

HMM with three states (q1 to q3) and five observation symbols (v1 to v5). The relationship among states is described by transition matrix A=[aij]3x3, and between states and discreet symbols by observation matrix B=[bij]3x5. It is assumed the system evolves through certain states whose relationship is to be studied.



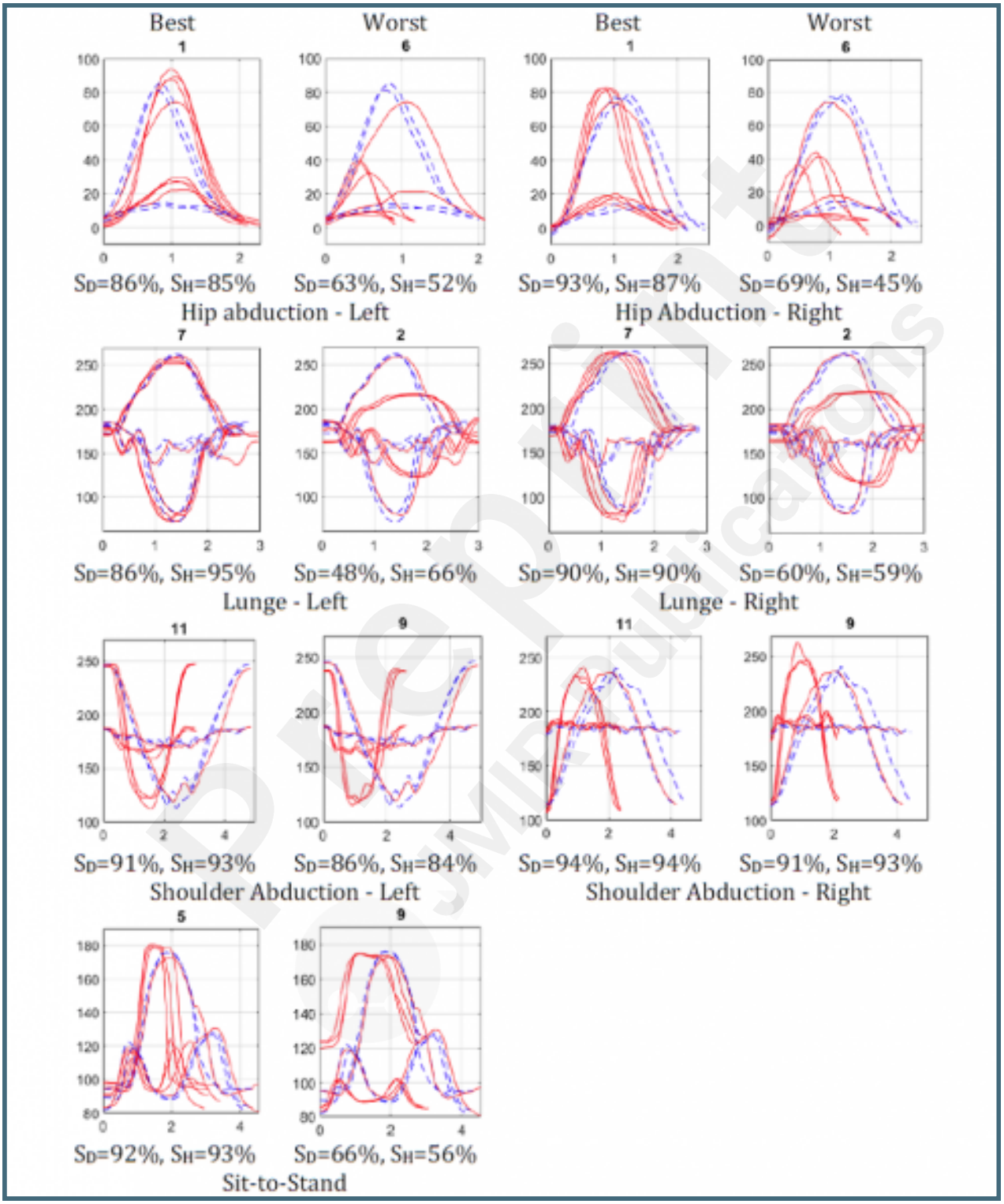
Preprint JMIR Publications

Similarity scores obtained from applying DTW (SD, solid blue line) and HMM (SH, solid red line). Their difference, SH-SD, is shown by the black dashed line. Participant 17 is the physiotherapist and 18 is the worst performance, i.e. making no movement. Error bars indicate standard error.

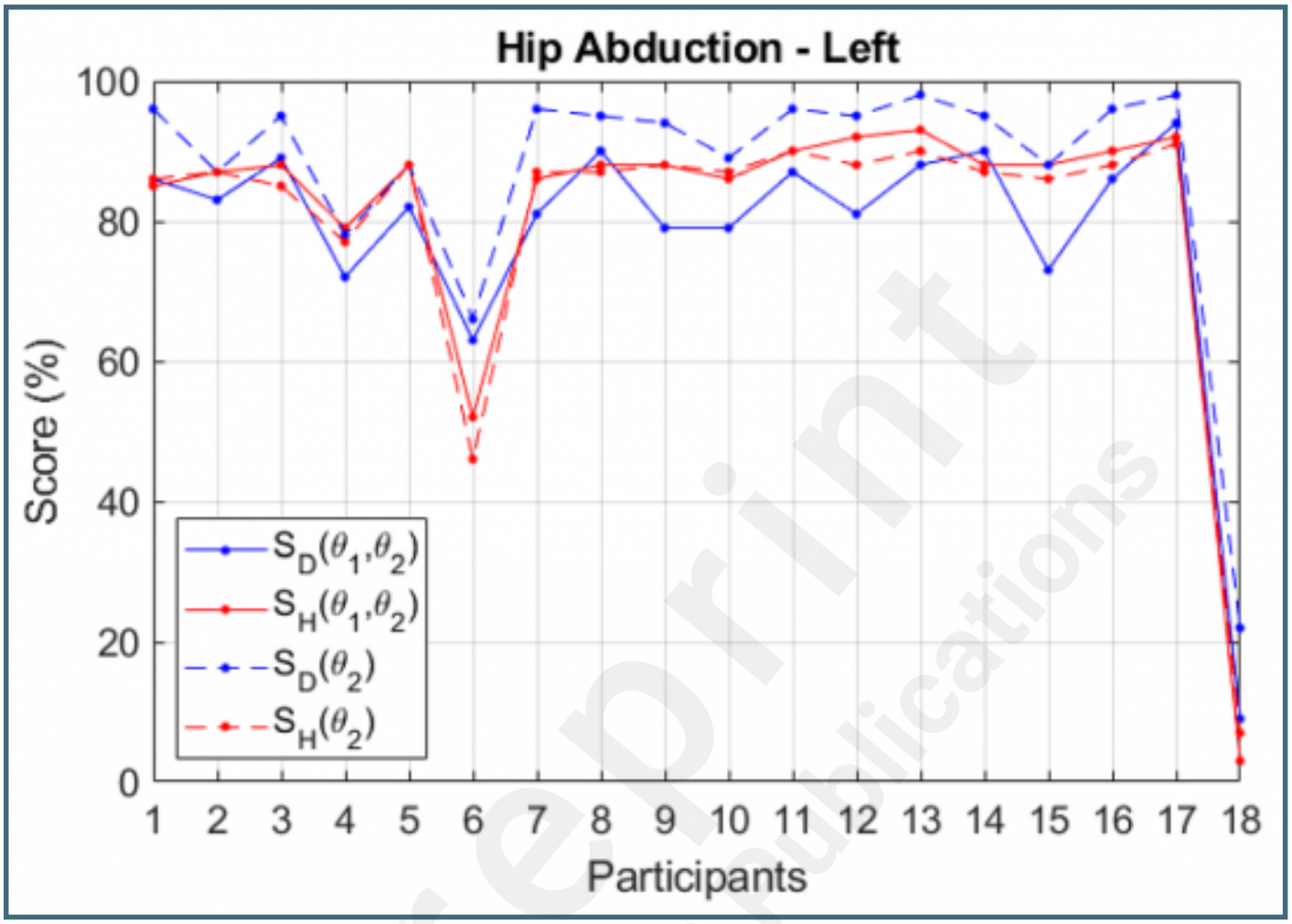


Pre-Print JMIR Publications

Time-domain plots of the best and worst trajectories (excluding participants 17 and 18). X-axis indicates the time in seconds and Y-axis indicates the joint angle in degree. The three repetitions of the reference trajectories are given in dashed blue lines and the chosen participant's trajectories in solid red lines. All plots include two features, except for lunge where three features are presented.

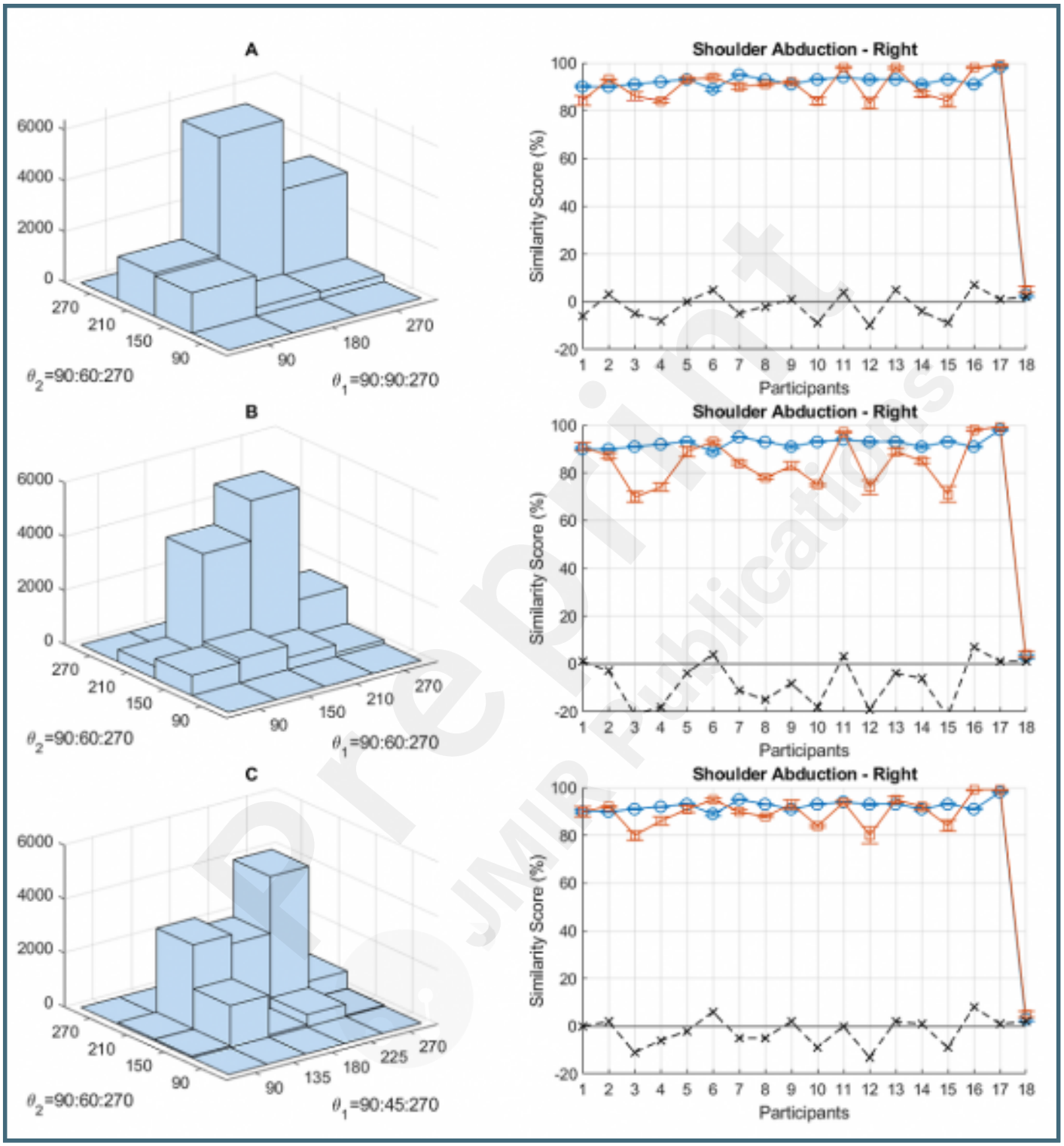


The effect on the scores of removing minor features.



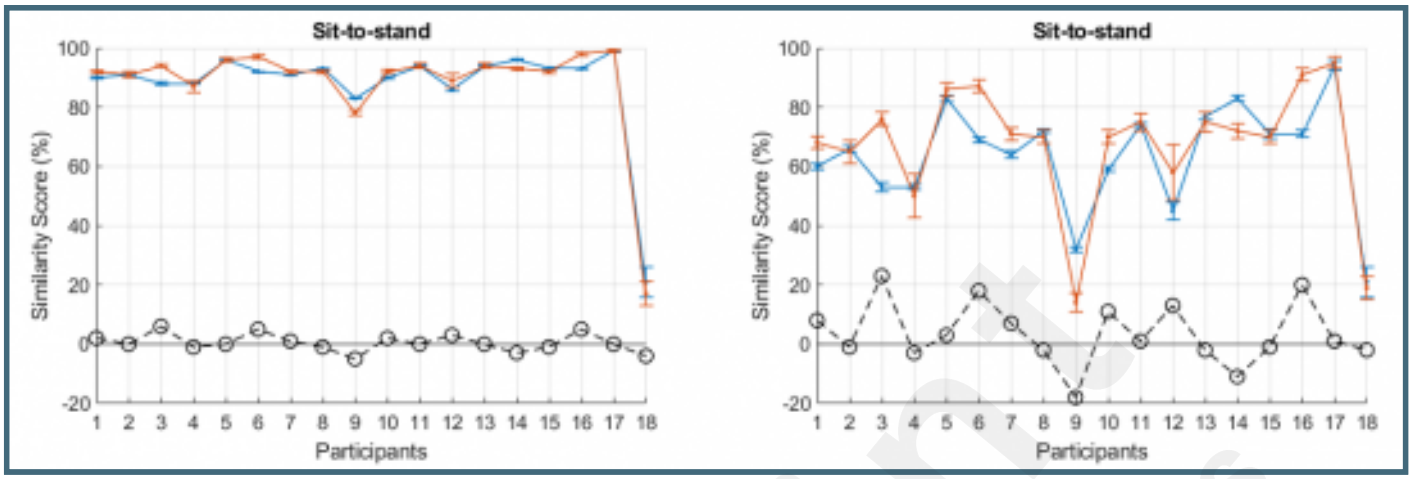
Preprint JMIR Publications

The effect of quantization on the calculated HMM scores.



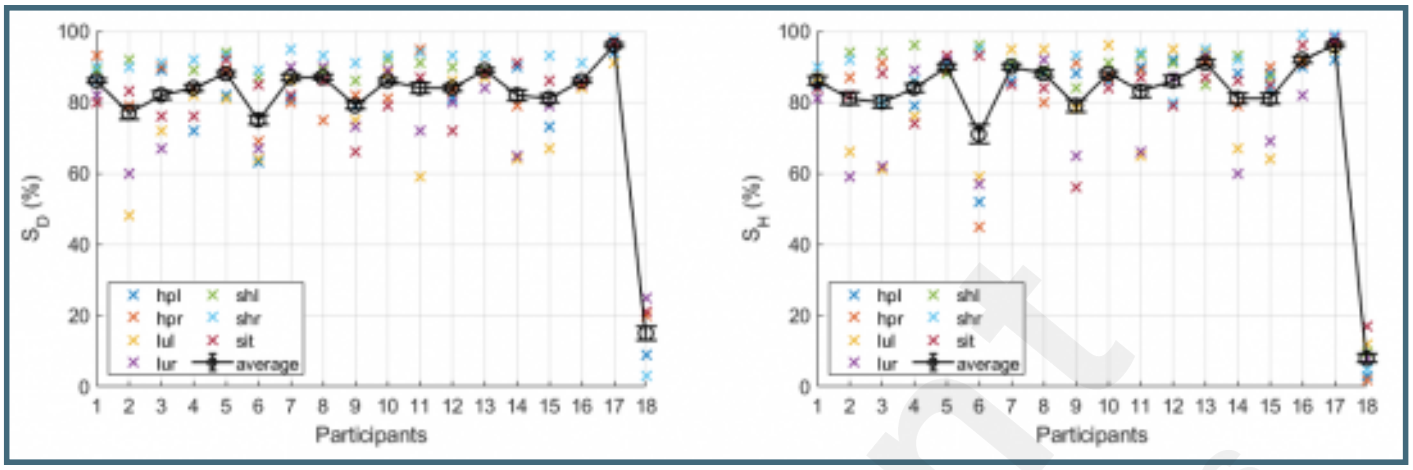
Preprint JMIR Publications

Effect of altering limits (Du and Ll, obtained from the worst performance) on the scores: multiplied by 2 (left) and divided by 2 (right). Error bars indicate standard error.



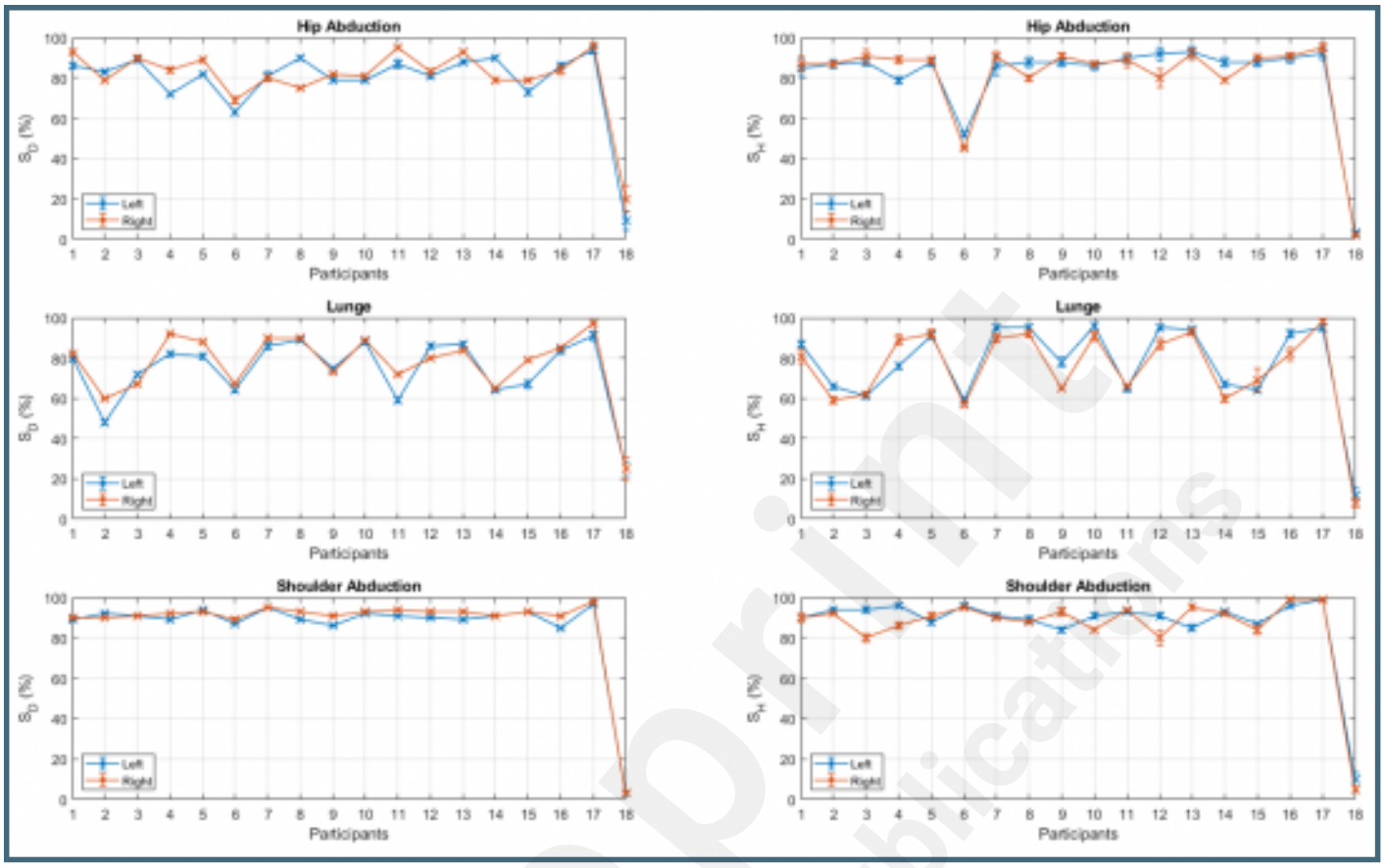
Preprint JMIR Publications

The combined and averaged scores (left: SD and right: SH) for all seven exercises. Error bars indicate the standard error.



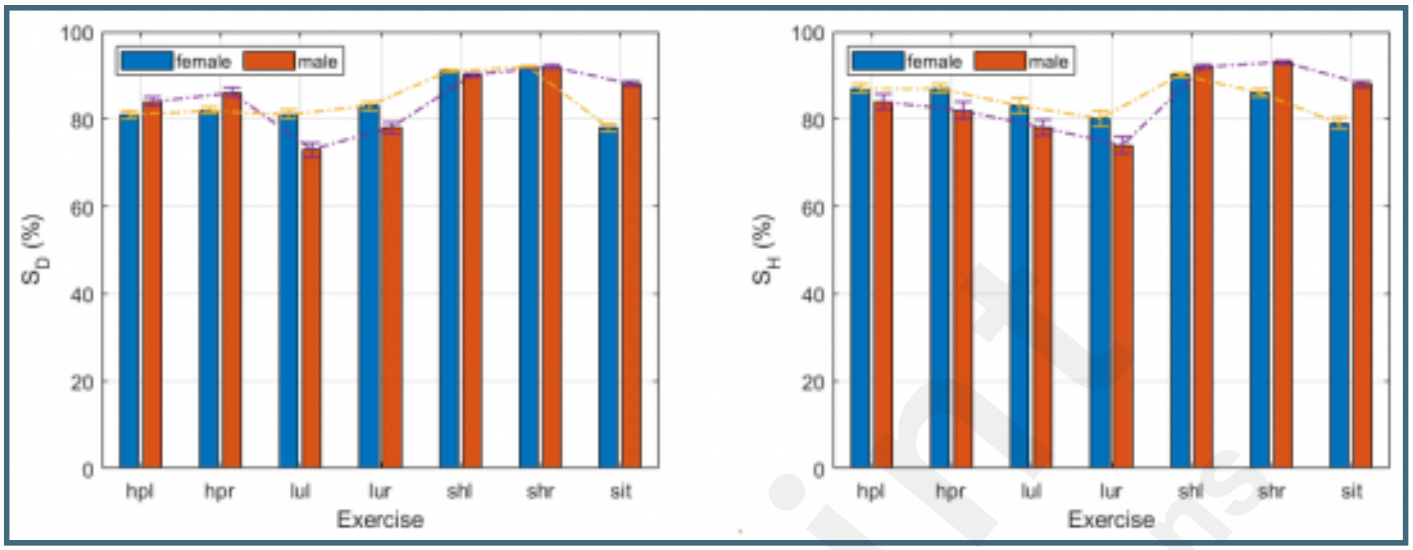
Preprint
JMIR Publications

Comparison between left and right scores. Error bars indicate standard error.



Preprint JMIR Publications

Comparison between female and male participants, left: DTW scores (SD) and right: HMM scores (SH). Error bars indicate standard error.

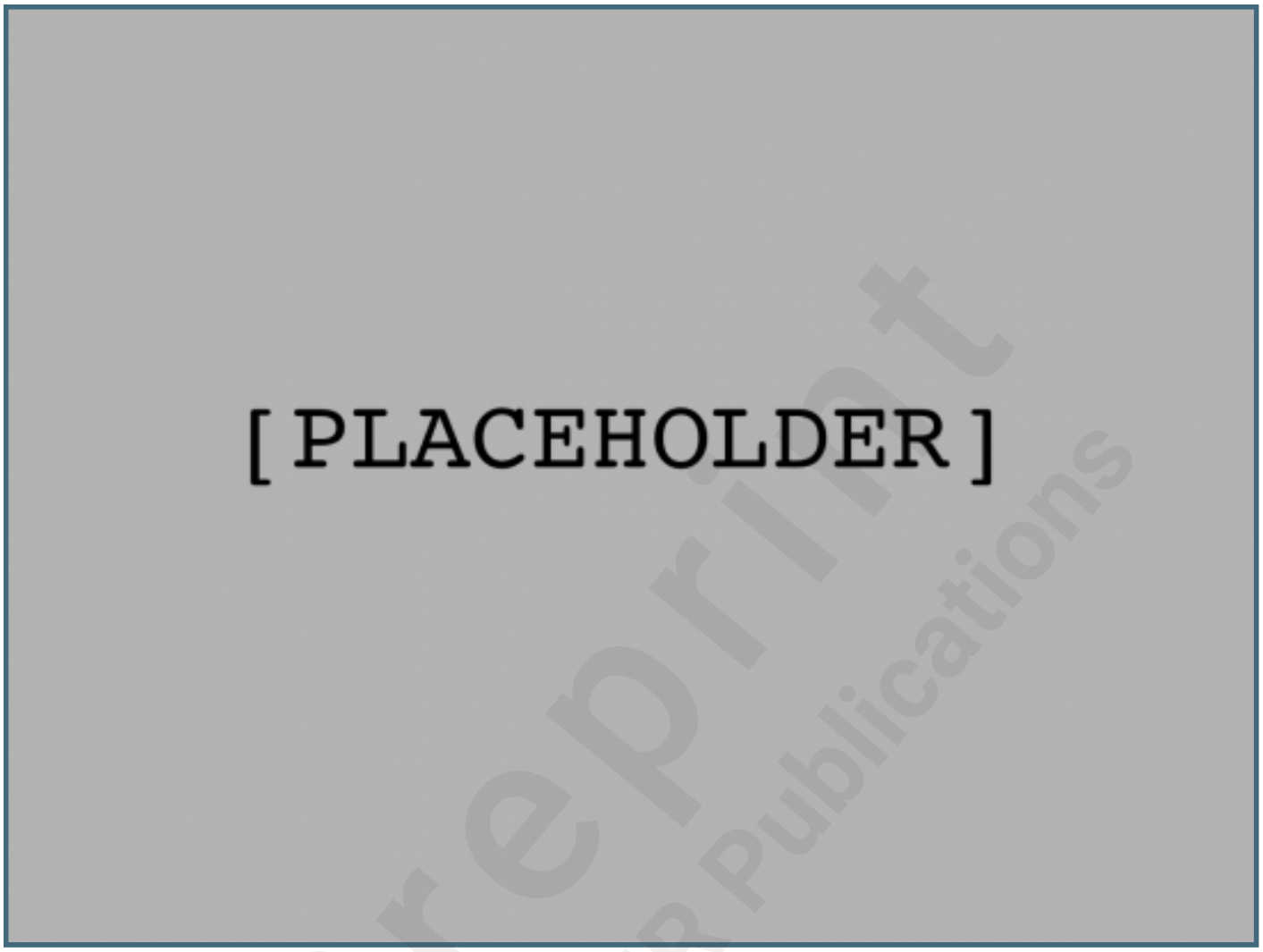


Preprint JMIR Publications

TOC/Feature image for homepages

Preprint JMIR Publications

Placeholder.



Preprint JMIR Publications