Motion Evaluation of Therapy Exercises by Means of Skeleton Normalisation, Incremental Dynamic Time Warping and Machine Learning: A Comparison of a Rule-Based and a Machine-Learning-Based Approach

Julia Richter, Christian Wiede, Ulrich Heinkel and Gangolf Hirtz
Department of Electrical Engineering and Information Technology
Technische Universität Chemnitz, Reichenhainer Str. 70, 09126 Chemnitz, Germany

Keywords: Health Care, Medical Training Therapy, Motion Quality Assessment, Assistance Systems, Machine Learning, Dynamic Time Warping, Skeleton Normalisation.

Abstract: The assessment of motions by means of technical assistance systems is attracting widespread interest in fields such as competitive sports, fitness and rehabilitation. Current research has achieved to generate feedback that concerns quantity or the grade of similarity with regard to correct reference motions. In view of post-operative rehabilitation exercises, such type of feedback is regarded as insufficient. That is why recent research aims at providing a qualitative feedback by communicating motion errors. While existing systems investigated the use of manually defined rules to detect motion errors, we suggest to employ machine learning techniques in combination with dynamic time warping and to train classifiers with sample exercise executions represented by 3-D skeletons joint trajectories. This study describes both a rule-based and a machine-learning-based approach and compares them with regard to their accuracy. In the second place, this study seeks to investigate the effect of using normalised hierarchical coordinates on the classification accuracy if data of different persons is used for the machine-learning-based approach. The results reveal that the performance of the machine-learning-based method compares well with the rule-based concept. Another outcome to emerge from this study is that normalised hierarchical coordinates allow to use data of different persons.

1 INTRODUCTION

Latest results from interviews with medical experts from rehabilitation centres revealed that nowadays a therapist has to supervise up to fifteen patients simultaneously (Lösch et al., 2018). Consequently, patients do not receive continuous feedback from a therapist while they are performing therapy exercises, so that incorrect motion executions are not detected and cannot be instantly corrected.

A variety of studies have already demonstrated that the manner of motion execution influences the muscle activity during an exercise (Kang et al., 2016), (Caterisano et al., 2002), (Gorsuch et al., 2013). For this reason, it is of high importance to control the motion execution by means of feedback. In this way, motion errors can be avoided so that the intended muscles are strained and the desired effects can be achieved.

A possible solution to provide feedback to a patient could be a technical assistance system that imitates the therapist's feedback. Such a system could not only be beneficial in the post-operative rehabilitation phase, but it could also be used by healthy persons at home as a preventive measure for maintaining their mobility.

Therefore, the overall objective that is pursued in our work is to develop an assistance system to supervise persons during their medical training therapy when no therapist is present to supervise their performance, either in rehabilitation centres or at home. For this, the therapist's knowledge and his or her visual perception as well as the therapist's real-time feedback shall be reproduced. This work includes the recognition of exercise-specific motion errors during the exercise execution to ensure the exercise quality.

The paper is structured as follows: Section 2 gives an overview about extant systems and recent research on motion analysis, whereupon the research issue is described. Section 3 introduces the developed assistance system. The system evaluation is presented in Section 4. Finally, Section 5 summarises the paper and draws conclusions about future work.

Richter, J., Wiede, C., Heinkel, U. and Hirtz, G.
Motion Evaluation of Therapy Exercises by Means of Skeleton Normalisation, Incremental Dynamic Time Warping and Machine Learning: A Comparison of a Rule-Based and a Machine-Learning-Based Approach.
DOI: 10.5220/0007260904970504
In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)*, pages 497-504
ISBN: 978-989-758-354-4

2 RELATED WORK

The following sections introduce state-of-the-art feedback systems and recent research that has been carried out so far.

2.1 Extant Systems for Motion Analysis

A variety of systems that provide feedback to the user can be found in the field of fitness training. Examples of prominent providers are Technomex, Balori, Ergofit, HUR and TecnoBody, which equip fitness centres with their devices. Table 1 provides an overview about the above mentioned systems.

When taking a closer look at these existing solutions, however, it becomes obvious that the focus lies on monitoring exercise quantity rather than exercise quality. Additionally, the feedback is mainly generated by means of sensors integrated in the devices, such as force, acceleration and distance sensors. In that way, motion amplitudes, velocities and the number of repetitions can be determined based on indirectly measured parameters in the training machine, e. g. the calculated torsional moment in the lever arm. Another major drawback of these systems is that they do not detect specific motion errors. Consequently, information about the motion quality is not available.

A step further goes the system developed by Pixformance. Their system processes 3-D skeleton joints in combination with manually defined rules in order to detect motion errors. Similar to Pixformance, Zhao et al. (Zhao et al., 2014) manually defined exercise-specific rules in XML format in cooperation with clinicians. Joint distances, joint angles, body segment orientations and repetition velocities were evaluated to assess motions.

2.2 Recent Research on Motion Analysis

A possibility to identify incorrect motions and to subsequently assess the person's performance is to match the executed motion against pre-defined references or templates. Thereby, a variety of approaches employed sensors that do not provide direct information about joint positions. Yurtman et al. distributed wearable motion sensors, i. e. accelerometers, gyroscopes and magnetometers, over the human body in order to classify a therapy exercise as correct or incorrect (Yurtman and Barshan, 2013). They searched for occurrences of defined template signals by using dynamic time warping (DTW). For post-stroke rehabilitation, Tormene et al. employed DTW on data provided by strain sensors that were worn into a long-sleeve shirt (Tormene et al., 2009). Tak et al. published a method for human abnormality detection in video sequences, for which they used DTW as well (Tak et al., 2011). Their method compared distance curves calculated from the detected foreground regions in each frame to find motion similarities.

Driven by the launch of the Kinect version 1.0 in 2010, researchers have increasingly investigated assessment methods that utilise the Kinect skeleton model (Shotton et al., 2013). Existing work examined motions during sports, such as dancing, ballet training and Tai Chi exercises, using the Kinect skeleton (Gal et al., 2015), (Huang et al., 2013), (Muneesawang et al., 2015), (Lin et al., 2013). These works introduced a variety of distance metrics in order to compare a performed motion against a reference. Such metrics were, for example, angles between joints (Gal et al., 2015), angles between body motion vectors (Huang et al., 2013) or the mean Euclidean distance between joint trajectories (Lin et al., 2013). Although the presented algorithms originally were not designed for therapy exercise assessment, their principle workflows were still inspiring for similarity assessment in view of therapy.

Hence, in the field of physical rehabilitation, similar principles can be found. In order to assess motions, several of the presented works successfully demonstrated the use of DTW (Su et al., 2014), (Antunes et al., 2016) or variants of DTW, such as Incremental Dynamic Time Warping (IDTW) (Khan et al., 2014) and Subsequence DTW (Baptista et al., 2017). The principle of these approaches is to align the two sequences and to calculate similarity measures or scores. Antunes et al., for example, calculated and visualised feedback vectors that show the difference between the performed and the template motion of body parts (Antunes et al., 2016). Baptista et al. extended this approach by investigating the alignment of a template and the performed exercise by means of Subsequence DTW and Temporal Commonality Discovery, whereas one sequence is a subsequence of the other (Baptista et al., 2017). Khan et al. introduced a method for continuous real-time evaluation, which is called IDTW, (Khan et al., 2014). The system they developed provides a similarity score but does not determine specific errors.

2.3 Research Issue

The purpose of this investigation is to explore approaches that allow the recognition of errors that typically occur during the performance of three selected exercises of therapeutic relevance: hip abduction, hip exten-

Table 1: Selection of feedback systems on the market.

System/ Provider	Application Field	Description
Technomex	Fitness games with focus on motivation	Skeleton-based measurement, not meant for rehabilitation, no motion error detection
Pixformance	Motion feedback in the field of fitness training	Skeleton-based measurement, rule-based motion error detection, manual rule specification
Balori	Coordination and balance training	Balance training based on pressure plate
HUR, Ergofit	Motion feedback in the field of fitness	Indirect measurement via sensors integrated in the machine, feedback is visualised by means of intensity bars and amplitude plots
TecnoBody	Motion feedback in the field of fitness	Skeleton-based measurement, no motion error detection by means of joint trajectories

sion and hip flexion, which are performed on a pulley machine and are error-prone due to the unrestricted motion range the machine allows.

We firstly investigated a rule-based approach to detect motion errors in real-time. Exercise-specific rules had to be defined and suitable thresholds had to be determined. The results of this approach serve as a reference benchmark for the comparison against the proposed machine-learning-based approach. Secondly, we designed, implemented and evaluated a machine-learning base approach, which avoids the manual specification of rules. The principle of the approach has already been presented in (Richter et al., 2017b). In order to avoid taking pre-recordings and examples of incorrect motions for new patients, we investigated a normalisation method. Therefore, the performance of two types of coordinate representations, i. e. local and normalised hierarchical coordinates, was compared against each other by using three different scenarios. Since we intend to give real-time feedback, the approach adapted the IDTW algorithm (Khan et al., 2014).

Two research questions arise at this point and shall be answered in the course of this study:

a) Can we train a machine by means of example sequences, so that the classification performance is similar to that of the rule-based approach?

b) Does the normalisation of skeleton data allow the usage of training data that is different from the person who finally uses the system?

3 DEVELOPED ASSISTANCE SYSTEM

3.1 Concept Overview

The developed assistance system has already been presented in (Richter et al., 2017a) and (Richter et al., 2017b), whereas the focus lied on hip abduction.

Figure 1 presents the concept of the assistance system: It consist of a sensor, i. e. a Kinect version 1.0, that measures the 3-D joint positions of a person's skeleton. Based on these positions, motion errors are determined in real-time. This information is sent to a feedback unit and visualised in form of feedback messages and joints of an animated avatar that are coloured depending on their involvement in the detected error. The feedback system is shown in Figure 2.

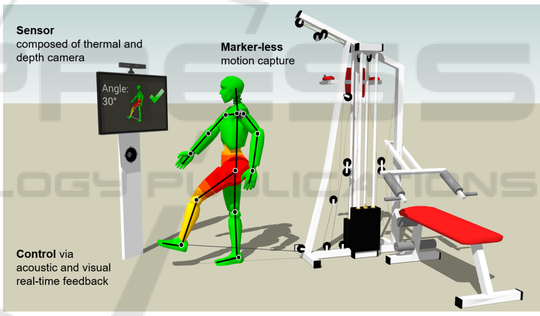


Figure 1: Concept of the assistance system (Richter et al., 2017a).

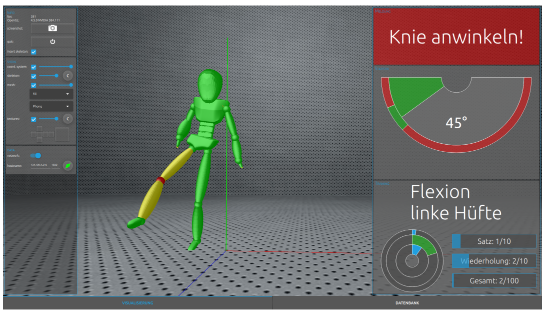


Figure 2: Feedback display: In this example, the person performs the exercise hip flexion and does not bend the knee in the correct way. A message with a suggestion how to correct the motion appears in real time (see top right, translated: Bend your knee!).

3.2 Motion Errors

This section presents the three exercises the study focusses on with their specific motion errors that have to be detected.

In the Figures 3, 4 and 5, the gray part of the skeleton indicates the moving body parts. The regions in red mark the error occurrences. The left skeletons represent the correct exercise executions with the class label C. The remaining class labels UB, FO, BK, WP and NBK correspond to the motion errors.

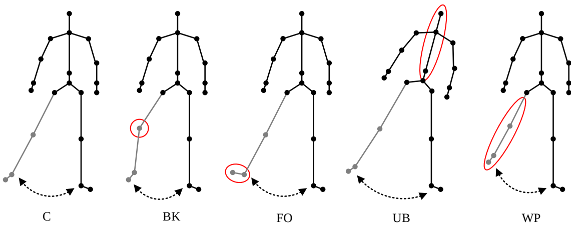


Figure 3: Motion errors for the exercise hip abduction.

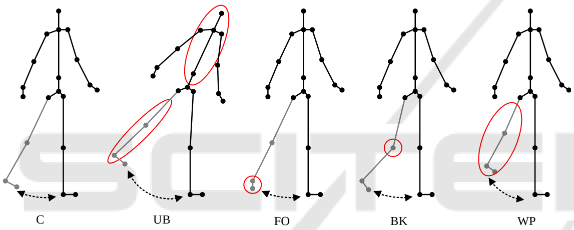


Figure 4: Motion errors for the exercise hip extension.

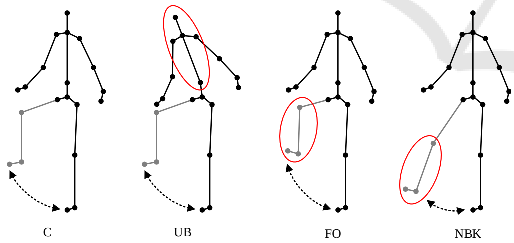


Figure 5: Motion errors for the exercise hip flexion.

3.3 Rule-Based Approach

3.3.1 Rule Modules

The rules are based on three function that are described in the following sections. These rules were applied on the 3-D joint positions represented in the 3-D Kinect camera coordinate system.

M1 – Angle between 3-D Articulated Joints. The angle between two articulated joints is calculated by this module. The input parameters of the module are two arbitrary articulated joints $\underline{a}_t$ and $\underline{b}_t$ in a frame t, which are defined by their corresponding start position $a_{1,t}$ and $b_{1,t}$ as well as their end position $a_{2,t}$ and $b_{2,t}$ respectively. These positions are given by 1×3 vectors. The output angle ρ_t is calculated according to Equation 1.

$$\rho_t = \arccos \frac{(a_{2,t}-a_{1,t})\cdot(b_{2,t}-b_{1,t})}{||a_{2,t}-a_{1,t}||\cdot||b_{2,t}-b_{1,t}||}. \tag{1}$$

M2 – Normalised Euclidean Distance between 3-D Joints. The normalised Euclidean distance dj_t between two joints a and b, whose positions are given by 1×3 vectors, is calculated according to Equation 2. In order to normalise this distance with regard to a person's height, the distance is divided by the sum of the right lower and upper leg $leg_{l,t}$ and $leg_{u,t}$. The right leg was chosen here instead of the left supporting leg, because in contrast to the left leg, the right leg is always visible during the exercise execution.

$$dj_t = \frac{||a_t - b_t||}{leg_{u,t} + leg_{l,t}}. \tag{2}$$

M3 – Normalised Signed Euclidean Distance of a 3-D Joint to a Defined Plane. This module allows to calculate the distance dp_t of a joint of interest $p_{oi,t}$ to the x-y, the x-z or the y-z plane of a Cartesian coordinate system. This coordinate system is defined by three joints a_t, b_t and c_t as well as a joint $p_{0,t}$ that defines the origin of the system. To obtain the distance, the following steps have to be performed: Firstly, a coordinate system with the rotation R and the origin $p_{0,t}$ is calculated using a_t, b_t and c_t. Secondly, $p_{oi,t}$ is transformed into this coordinate system and results in the transformed coordinates $p_{oi,trans,t}$. These steps are represented by the following Equations:

$$x_t = \frac{b_t - a_t}{||b_t - a_t||} \tag{3}$$

$$z_t = (b_t - a_t) \times (a_t - c_t) \tag{4}$$

$$y_t = \frac{z_t \times x_t}{||z_t \times x_t||} \tag{5}$$

$$R_t = \begin{bmatrix} x_t^{\mathsf{T}} & y_t^{\mathsf{T}} & z_t^{\mathsf{T}} \end{bmatrix} \tag{6}$$

$$p_{oi,trans,t} = R_t \cdot \left(p_{oi,t}^{\mathsf{T}} - p_{0,t}^{\mathsf{T}}\right). \tag{7}$$

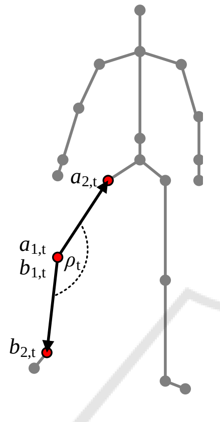
Finally, depending on the plane to which the normalised signed distance shall be calculated, the x, y or z coordinate of the transformed joint is returned according to Equation 8.

$$dp_t = \begin{cases} \frac{p_{oi,trans,t}(z)}{leg_{u,t}+leg_{l,t}}, & \text{if distance to x-y plane,} \\ \frac{p_{oi,trans,t}(y)}{leg_{u,t}+leg_{l,t}}, & \text{if distance to x-z plane,} \\ \frac{p_{oi,trans,t}(x)}{leg_{u,t}+leg_{l,t}}, & \text{if distance to y-z plane.} \end{cases} \tag{8}$$

3.3.2 Rules

Due to limited space, this section presents only the rule of BK for hip abduction in Table 2. If a rule is fulfilled for a frame t, then the corresponding error was detected in this frame. If none of the defined errors were detected, the performed motion was considered to be correct.

Table 2: Hip abduction: Bent Knee (BK).

Module: M1 ρ_t: Angle between upper and lower right leg Rule: $\rho_t < thr_{\mathrm{Abd,BK}}$	

3.3.3 Thresholds

The thresholds for every rule were determined by means of ROC (receiver operating characteristic) curves. The principle is described for the error BK (hip abduction). In this study, the data of all recorded persons was intentionally used, which results in the optimal thresholds for this very test group. This means that, in case of manual threshold selection, the results cannot exceed the accuracy that is determined in this experiment. The determined optimal threshold $thr_{\mathrm{Abd,BK}}$ of the error BK for hip abduction was 159 °.

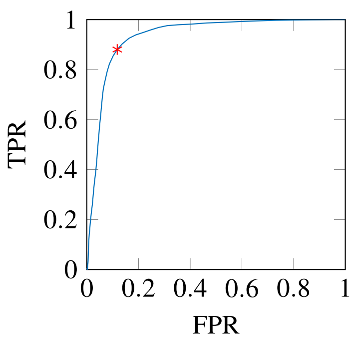


Figure 6: ROC curve for hip abduction, Bent Knee (BK).

Figure 6 presents the according ROC curve with the optimal operating point, i. e. the threshold that separates the samples representing the error classes from the rest of the samples. The red dot on the curve indicates the optimal operating point. Figure 7 shows the separation of the data by means of this determined threshold.

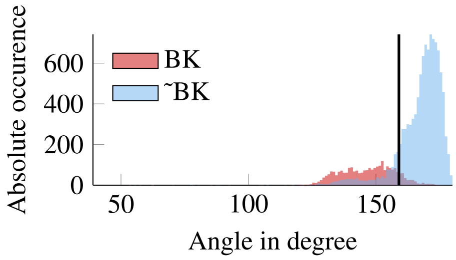


Figure 7: Samples separated by the determined threshold for hip abduction, Bent Knee (BK).

3.4 Machine Learning-Based Approach

The machine learning-based approach has already been published for hip abduction by Richter et al. (Richter et al., 2017b). Their algorithm transforms the obtained 3-D skeleton to normalised hierarchical coordinates. The aim of this transformation is to normalise skeletons of different persons and to make the joint data rotation- and translation-invariant. In the next step, the current motion of a person performing an exercise was aligned with a correctly performed reference by means of IDTW. The new idea in the work of Richter et al. was to calculate a difference vector between the reference and the current joint positions, which was then fed to a set of support vector machines. Thereby, every machine voted for one single error. As a result, the machine could classify whether a current exercise motion is correct or whether a specific motion error occurred.

4 EVALUATION

4.1 Evaluation Measures

To compare the performance of the different approaches presented in this work, the accuracy of the single classifiers η_s and the overall accuracy η_{all}, which is the mean accuracy of all classifiers S for an exercise, are used. Thereby, the overall accuracy is defined by

$$\eta_{\mathrm{all}} = \frac{1}{S} \cdot \sum_{s=1}^{S} \eta_s\,, \tag{9}$$

whereas S is the number of classifiers or error classes used for one exercise.

4.2 Dataset Configurations

To investigate the effect of local and normalised hierarchical coordinates on the classification accuracy, three different scenarios were defined by Richter et al. (Richter et al., 2017b):

Scenario 1. For every patient, an individual machine was trained and tested with the individual patient's data and his or her own reference. In practice, training data as well as the patient's reference have to be recorded for every new patient, which is rather impractical.

Scenario 2. In contrast to scenario 1, one single machine was trained with the motion data of three persons and their individual reference data. Just as in scenario 1, the test was performed with each test person's individual reference. In practice, only the new patient's reference, but not his or her training data, has to be recorded.

Scenario 3. The machine was trained with the motion data of the three persons mentioned in scenario 2. In contrast to scenario 1 and scenario 2, the reference has not been changed according to the person. For testing, the very same reference was used for all test samples of the ten test persons. This means that, in practice, neither training data nor a reference has to be recorded for a new patient.

4.3 Results and Discussion

4.3.1 Usage of Normalised Skeleton Data

Table 3 and Figure 8 present the results to answer the research question **b)** raised in Section 2.3.

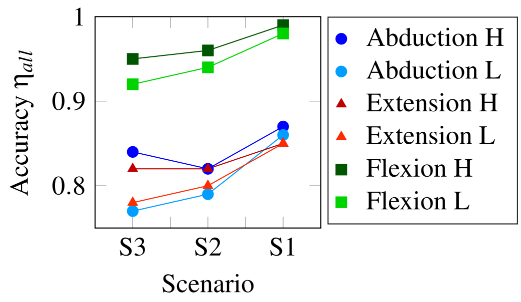


Figure 8: Overall accuracies η_{all} for the three scenarios S1, S2 and S3 with local (L) and normalised hierarchical (H) coordinates.

Table 3: Overall accuracies η_{all} for the three scenarios S1, S2 and S3 with local (L) and normalised hierarchical (H) coordinates and accuracy differences D = S1 - S3 between S1 and S3.

	Abduction		Extension		Flexion	
	L	H	L	H	L	H
S1	0.86	0.87	0.85	0.85	0.98	0.99
S2	0.79	0.82	0.80	0.82	0.94	0.96
S3	0.77	**0.84**	0.78	**0.82**	0.92	**0.95**
D	0.09	0.03	0.07	0.03	0.06	0.04

Table 3 compares the results of local and normalised hierarchical coordinates for the three different scenarios and Figure 8 visualised these numbers in a graph.

When taking a closer look at the influence of hierarchical and local coordinates, we notice the following: In general, the accuracies for hierarchical coordinates are higher than for local coordinates. Moreover, and what is the most important, is that hierarchical coordinates outperform local coordinates especially for S3, which is the scenario of highest practical relevance. From the graph and the table above we can see that the usage of local, i. e. non-normalised coordinates results in a considerable deterioration of the overall accuracy from scenario 1 to scenario 3. This implies that, for non-normalised local coordinates, the less personalised data is used, the higher is the negative influence of factors, such as body size and proportions, on the accuracy. The difference to local coordinates becomes obvious when we have a look at the overall accuracies when using normalised hierarchical coordinates. It is apparent that, in contrast to local coordinates, the accuracy only slightly decreases when using normalised hierarchical coordinates. These results suggest to use normalised hierarchical coordinates instead of local coordinates should the recording of personalised data for each new patient be avoided without loosing classification performance, which correspond to scenario 3. Still, for scenario 2 and even for hip flexion and abduction in scenario 1, normalised hierarchical coordinates outperform the local ones. These results indicate that, in general, normalised hierarchical coordinates are a suitable representation of therapy exercise trajectories.

4.3.2 Rule-Based Versus Machine-Learning-Based

Table 4 and Table 5 present the results to answer the research question **a)**, whereas Table 4 shows the achieved performance of the rule-based approach and Table 5 reveals the performance of the machine-learning-based approach for the scenario 3.

If we compare the results of the rule-based approach in Table 4 with the results of the machine-learning-based approach in Table 5, it can be seen that the performance of the machine-learning-based method is comparable to the rule-based method. As a consequence, these results demonstrate that the detection of motion errors by means of machine lear-

Table 4: Confusion matrices and corresponding accuracies η_s and η_{all} and overall accuracies η_{all} for rule-based approach.

	Abduction				Extension				Flexion		
	L	$\overline{L}$	η_s		L	$\overline{L}$	η_s		L	$\overline{L}$	η_s
C	0.72	0.28	0.81	C	0.73	0.27	0.83	C	0.91	0.09	0.94
$\overline{\text{C}}$	0.09	0.91		$\overline{\text{C}}$	0.08	0.92		$\overline{\text{C}}$	0.04	0.96	
BK	0.88	0.12	0.91	UB	0.97	0.03	0.98	UB	1.00	0.00	1.00
$\overline{\text{BK}}$	0.06	0.94		$\overline{\text{UB}}$	0.02	0.98		$\overline{\text{UB}}$	0.00	1.00	
FO	0.78	0.22	0.73	FO	0.70	0.30	0.69	FO	0.89	0.11	0.90
$\overline{\text{FO}}$	0.32	0.68		$\overline{\text{FO}}$	0.31	0.69		$\overline{\text{FO}}$	0.09	0.91	
UB	0.91	0.09	0.93	BK	0.87	0.13	0.87	NBK	1.00	0.00	0.99
$\overline{\text{UB}}$	0.05	0.95		$\overline{\text{BK}}$	0.12	0.88		$\overline{\text{NBK}}$	0.01	0.99	
WP	0.84	0.16	0.83	WP	0.84	0.16	0.86				
$\overline{\text{WP}}$	0.18	0.82		$\overline{\text{WP}}$	0.13	0.87					
η_{all}			**0.84**				**0.85**				**0.96**

Table 5: Confusion matrices and corresponding accuracies η_s and η_{all} for scenario 3 for machine-learning-based approach. Classifiers are regarded independently from each other.

	Abduction				Extension				Flexion		
	L	$\overline{L}$	η_s		L	$\overline{L}$	η_s		L	$\overline{L}$	η_s
C	0.84	0.16	0.80	C	0.92	0.08	0.84	C	0.84	0.16	0.91
$\overline{\text{C}}$	0.24	0.76		$\overline{\text{C}}$	0.24	0.76		$\overline{\text{C}}$	0.01	0.99	
BK	0.95	0.05	0.95	UB	0.90	0.10	0.90	UB	0.98	0.02	0.98
$\overline{\text{BK}}$	0.05	0.95		$\overline{\text{UB}}$	0.09	0.91		$\overline{\text{UB}}$	0.02	0.98	
FO	0.83	0.17	0.70	FO	0.54	0.46	0.67	FO	0.91	0.09	0.92
$\overline{\text{FO}}$	0.44	0.56		$\overline{\text{FO}}$	0.20	0.80		$\overline{\text{FO}}$	0.08	0.92	
UB	0.89	0.11	0.88	BK	0.87	0.13	0.86	NBK	1.00	0.00	0.99
$\overline{\text{UB}}$	0.13	0.87		$\overline{\text{BK}}$	0.16	0.84		$\overline{\text{NBK}}$	0.02	0.98	
WP	0.80	0.20	0.87	WP	0.73	0.27	0.84				
$\overline{\text{WP}}$	0.07	0.93		$\overline{\text{WP}}$	0.04	0.96					
η_{all}			**0.84**				**0.82**				**0.95**

ning techniques is equally reliable as the detection by means of manually defined rules. In conclusion, we provided evidence that a system can be trained with a set of example sequences, so that motion errors can be detected. This finally avoids the manual definition of rules for every exercise, because the machine can automatically derive the rules based on the provided example sequences.

5 CONCLUSIONS AND FUTURE WORK

To sum up, this study has found that the introduced machine-learning-based approach works equally accurate as the rule-based version. As a consequence, the manual definition of rules can be avoided because the machine is able to learn motion errors by means of annotated sample sequences. In addition, this study has identified that normalised hierarchical coordinates allow to employ training samples of persons that are different from those who finally use the system. Hence, we lay a foundation towards a more qualitative feedback during therapy exercises. This work contributes to an automatic inclusion of new exercises by simply recording sample sequences of an exercise along with sequences containing specific errors. In future work, we therefore aim at an automated annotation process for the training data, which is labelled manually at the moment. Finally, further exercises shall be included, for example by a therapist who specifies both the correct execution and the errors that shall be detected. In this way, it could be possible in future to remotely create exercises for different patients or to simply pre-train systems with new exercises without taking pre-recordings of every person that uses the system.

REFERENCES

Antunes, M., Baptista, R., Demisse, G., Aouada, D., and Ottersten, B. (2016). Visual and human-interpretable feedback for assisting physical activity. In *Euro-*

pean Conference on Computer Vision, pages 115–129. Springer.

Baptista, R., Antunes, M., Aouada, D., and Ottersten, B. (2017). Video-based feedback for assisting physical activity. In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 5: VISAPP, (VISIGRAPP 2017)*, pages 274–280.

Caterisano, A., Moss, R., Pellinger, T., Woodruff, K., Lewis, V., Booth, W., and Khadra, T. (2002). The effect of back squat depth on the EMG activity of 4 superficial hip and thigh muscles. *The Journal of Strength & Conditioning Research*, 16(3):428–432.

Gal, N., Andrei, D., Neme, D. I., Ndan, E., and Stoicu-Tivadar, V. (2015). A Kinect based intelligent e-rehabilitation system in physical therapy. *Digital Healthcare Empowering Europeans*, pages 489–493.

Gorsuch, J., Long, J., Miller, K., Primeau, K., Rutledge, S., Sossong, A., and Durocher, J. J. (2013). The effect of squat depth on multiarticular muscle activation in collegiate cross-country runners. *The Journal of Strength & Conditioning Research*, 27(9):2619–2625.

Huang, T.-C., Cheng, Y.-C., and Chiang, C.-C. (2013). Automatic Dancing Assessment Using Kinect. In *Advances in Intelligent Systems and Applications-Volume 2*, pages 511–520. Springer.

Kang, S.-Y., Choung, S.-D., and Jeon, H.-S. (2016). Modifying the hip abduction angle during bridging exercise can facilitate gluteus maximus activity. *Manual therapy*.

Khan, N. M., Lin, S., Guan, L., and Guo, B. (2014). A visual evaluation framework for in-home physical rehabilitation. In *Multimedia (ISM), 2014 IEEE International Symposium on Multimedia*, pages 237–240. IEEE.

Lin, T.-Y., Hsieh, C.-H., and Lee, J.-D. (2013). A kinect-based system for physical rehabilitation: Utilizing tai chi exercises to improve movement disorders in patients with balance ability. In *2013 7th Asia Modelling Symposium*, pages 149–153. IEEE.

Lösch, C., Weigert, M., Nitzsche, N., Richter, J., Wiede, C., and Schulz, H. (2018). Einsatz und Bedeutung von Seilzügen in der Medizinischen Trainingstherapie am Beispiel Hüft-Totalendoprothese Eine Expertenperspektive. In *Bewegungstherapie und Gesundheitssport*.

Muneesawang, P., Khan, N. M., Kyan, M., Elder, R. B., Dong, N., Sun, G., Li, H., Zhong, L., and Guan, L. (2015). A machine intelligence approach to virtual ballet training. *IEEE MultiMedia*, 22(4):80–92.

Richter, J., Wiede, C., Apitzsch, A., Nitzsche, N., Lösch, C., Weigert, M., Kronfeld, T., Weisleder, S., and Hirtz, G. (2017a). Assisted Motion Control in Therapy Environments Using Smart Sensor Technology: Challenges and Opportunities. In *Ambient Assisted Living, 9. AAL-Kongress, Frankfurt/M, Germany, April 20 - 21, 2016*, pages 119–132. Springer Verlag.

Richter, J., Wiede, C., Shinde, B., and Hirtz, G. (2017b). 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 - Volume 1: ICPRAM*, pages 281–288.

Shotton, J., Girshick, R., Fitzgibbon, A., Sharp, T., Cook, M., Finocchio, M., Moore, R., Kohli, P., Criminisi, A., Kipman, A., et al. (2013). Efficient human pose estimation from single depth images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12):2821–2840.

Su, C.-J., Chiang, C.-Y., and Huang, J.-Y. (2014). Kinect-enabled home-based rehabilitation system using Dynamic Time Warping and fuzzy logic. *Applied Soft Computing*, 22:652–666.

Tak, Y.-S., Rho, S., and Hwang, E. (2011). Motion Sequence-Based Human Abnormality Detection Scheme for Smart Spaces. *Wireless Personal Communications*, 60(3):507–519.

Tormene, P., Giorgino, T., Quaglini, S., and Stefanelli, M. (2009). Matching incomplete time series with dynamic time warping: an algorithm and an application to post-stroke rehabilitation. *Artificial intelligence in medicine*, 45(1):11–34.

Yurtman, A. and Barshan, B. (2013). Detection and evaluation of physical therapy exercises by dynamic time warping using wearable motion sensor units. In *Information Sciences and Systems 2013*, pages 305–314. Springer.

Zhao, W., Lun, R., Espy, D. D., and Reinthal, M. A. (2014). Rule based realtime motion assessment for rehabilitation exercises. In *IEEE Symposium on Computational Intelligence in Healthcare and e-health (CICARE), 2014*, pages 133–140. IEEE.