

# Design of a Kinematic and Emotional Assessment Module for the Tele-rehabilitation Platform

Yves Rybarczyk, Louis Leconte, Jorge Luis Pérez Medina, Karina Jimenes Vargas, Patricia Acosta-Vargas, and Danilo Esparza

**Abstract**—Tele-rehabilitation is becoming extremely popular in the health scenario. Although this new medical approach has several advantages over traditional therapy, it is important to identify its limitations and risks in the recovery process. This study focuses on the modeling and implementation of a module to assess in real time the quality of movement and the emotional state of patients when they execute the rehabilitation exercises provided by a web-based platform. An algorithm based on dynamic time warping is proposed and tested to assess the movement kinematic. Furthermore, pain is discriminated through a support vector machine classifier. Both methods allow a high identification of gestures and emotions, respectively. These results are discussed in terms of the most suitable solutions to develop an efficient tele-rehabilitation system.

**Index Terms**—Dynamic time warping, machine learning, motion assessment, affective computing, eHealth.

## I. INTRODUCTION

Over the past few years, alternative healthcare deliveries have been developed. For instance, advances in telecommunications have permitted the emergence of tele-health systems. One specialized field of tele-health is tele-rehabilitation, which allows the implementation of a therapeutic program via an interactive multimedia web-based platform [1], [2]. A remarkable increase in the number of patients treated by tele-rehabilitation has been noticed since the beginning of the 21<sup>st</sup> century, especially in physiotherapy [3]. Despite the several medical and economic advantages of this new paradigm, the development of a tele-rehabilitation platform has to follow specific rules to ensure safety and efficiency. Two fundamental rules are (i) making sure the patient is performing the therapeutic movement correctly and (ii) identifying possible pain during execution of the exercises [4].

This study addresses these two key aspects in developing an effective and smart tele-rehabilitation platform to support the recovery of patients after hip replacement surgery. The implementation is based on low-cost technologies (e.g.,

Kinect-based motion capture) and a modular architecture to ensure the viability and scalability of the tool (Fig. 1). The patient's performance is stored in a web-based database that permits a remote monitoring by health professionals. The original contribution of our system is to integrate in a single platform an assessment module of both the quality of the performed movements and the emotional state of the patient when completing the therapeutic exercises.

Because the exercises are carried out in front of a machine instead of a therapist, participants may produce incorrect movements that could be harmful, especially after a surgery. In addition, the lack of human presence diminishes moral support, which may slow down the recovery process. Here the movement assessment is based on a dynamic time warping (DTW) approach, which has several advantages over the other methods (e.g., hidden Markov models) in terms of the size of the referential dataset and the possibility of performing an independent analysis on different portions of the body [5], [6]. The automatic identification of a situation in which patients are faced with difficulties or pain is based on a detection technique of subtle features of facial expressions [7] or body movements [8]. Then, machine learning classifiers are used to distinguish whether the patients are reaching their limits. Once such a situation is identified, a positive reinforcement (e.g., a message of encouragement) can be displayed on the user interface.

The remainder of this paper is organized into four parts. Section II explains the software architecture of the platform. Section III describes the method used to assess the quality of the movements and its experimental validation. Section IV presents the approach to recognize pain and the results of the experiment designed to test our machine learning algorithm. Finally, Section V draws conclusions and proposes future work.

## II. SOFTWARE ARCHITECTURE

### A. Therapeutic Education

The therapeutic education block consists of a set of functionalities that includes the rehabilitation exercises. This block is built in *Django 2.0.5*. The web framework uses the Model-View-Controller design pattern and works under two main principles: (i) maintaining loose coupling between layers of the framework and (ii) Don't Repeat Yourself (DRY) [9]. According to this pattern, the application functionality is divided into three kinds of components. The *model* represents knowledge and has the logic to update the controller. The *view* represents visualization of the data. The *controller* determines data flow into a model object and

Manuscript received August 1, 2018; revised April 12, 2019.

Yves Rybarczyk is with the University of Skövde, Skövde 54128, Sweden (e-mail: y.rybarczyk@fct.unl.pt).

Louis Leconte is with the Ecole Normale Supérieure, Paris-Saclay 94235, France (e-mail: louis.leconte@ens-paris-saclay.fr).

Jorge Luis Pérez Medina, Karina Jimenes Vargas, Patricia Acosta-Vargas, and Danilo Esparza are with the Universidad de Las Américas, Quito, Ecuador (e-mail: jorge.perez.medina@udla.edu.ec, karina.jimenes@udla.edu.ec, patricia.acosta@udla.edu.ec, wilmer.esparza@udla.edu.ec).

updates the view whenever data change. It keeps *view* and *model* separated and acts on both.

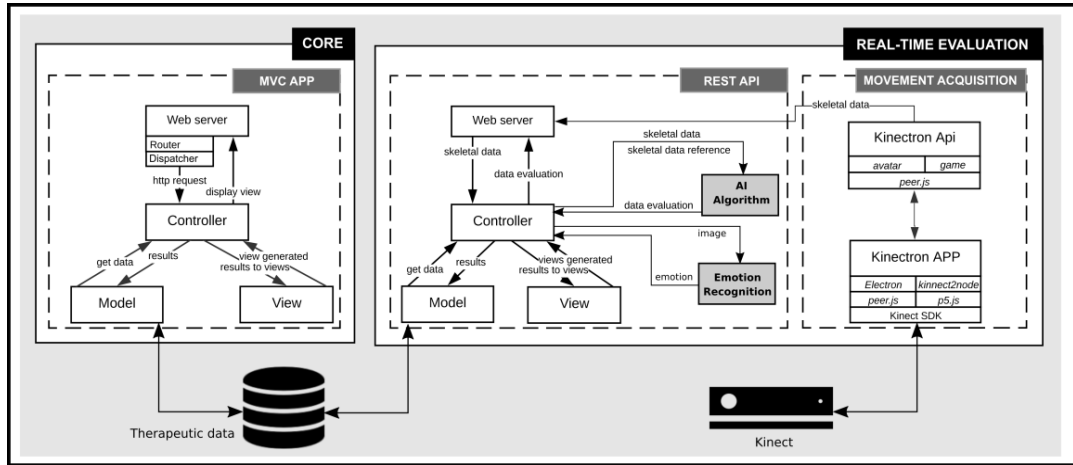


Fig. 1. Architecture of the tele-rehabilitation platform.

### B. Real-Time Evaluation

For real-time evaluation, two additional modules were created: (i) to process skeleton data with a RESTful API and (ii) to acquire movement data as a client application. They communicate with each other to process the movement data and to give the results using JSON data format. The RESTful API was developed using *Django Rest Framework* 3.8.2, which is a powerful and flexible toolkit of Django to build RESTful Web Services [10].

The client application was built with *Kinectron 1.4.2* and *Processing 1.4.8*. Kinectron is an open source tool to capture movement data. It has two components: (i) a server that broadcasts Kinect data and (ii) an API that receives Kinect data [11]. A Kinect v2 camera was used to capture the movements of the patient and extract the coordinates of the main body joints. A cross-browser JavaScript Library was incorporated (Three.js) to display, in a web browser, the game-based exercises, which were made up of 3D animated objects and the user's avatar.

### C. Assessment Module

The API REST package works with the exercise packages and contains intelligent modules to assess both the quality of the movements and the emotions of the patients when performing an exercise. It has a couple of services to process the skeleton and the image frames through different artificial intelligence algorithms.

The first service processes the skeleton to evaluate whether the patient is exercising correctly, and the result is transferred as a JSON structure with a set of angles. Each angle has a type, a DTW metric that compares the performed movement with the reference movement. A graphical-based feedback is displayed on the user's interface to inform patients on the quality of their performance (Fig. 2). A red color represents a bad movement amplitude for a determined limb or body part. This visual feedback is reinforced by a textual description of potential errors in the motion. Other information, such as a progress bar and a score, is also displayed.

The second service processes the image frame every second to determine whether the patient feels pain. A machine learning algorithm based on a support vector machine (SVM) is implemented to operate a Boolean classification of facial expressions in pain (1) vs. no pain (0).

A sequence of 1 and 0 is obtained at the end of each repetition. This information is used by the therapist to identify whether the intensity of the exercises is appropriate or not for a determined patient.

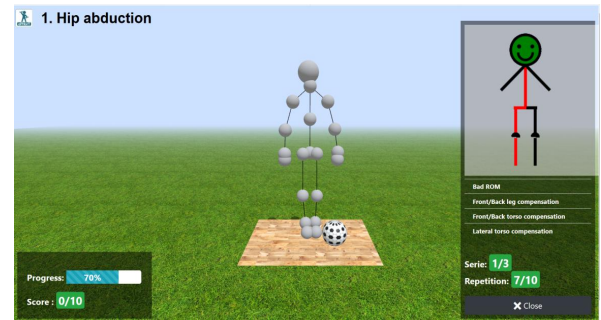


Fig. 2. User's interface of game-based exercises.

## III. MOVEMENT ASSESSMENT

### A. Method

#### 1) Computational modeling

A DTW algorithm was used to measure the similarity between the reference movement and the current movement performed by the patient. The method is based on the nonlinear alignment of two temporal sequences, which allows similar shapes to match even if they are out of phase in the time axis. The implementation of such an algorithm is obtained through a matrix sized with the length of each signal. On each point of the matrix, the distance between the points associated with each of the two signals is calculated. The best alignment of two signals is given by the path through the grid that minimizes the total distance between these two signals. Equation (1) presents the calculation of the quadratic distance used in our implementation:

$$D_{(i,j)} = (x_{(j)} - y_{(i)})^2. \quad (1)$$

Thus, a distance matrix was obtained. Then, a cumulated distance matrix (C) was created, based on the distance matrix (D). The latter was built due to a definition of distance very close to the 8-connectivity in a square tiling (mathematic topology) for each point of the matrix. Equation (2) describes the calculation of C:

$$C_{(i,j)} = \min\{C_{(i-1,j-1)}, C_{(i-1,j)}, C_{(i,j-1)}\} + D_{(i,j)}. \quad (2)$$

To quantify the difference between the two movements, we had to look for the path that minimizes the cumulated distance. To do so, we started from the last point and looked for the point with the lower cumulated distance between all its three inferior neighbors (within the meaning of Moore neighborhood). We reiterated the process until we reached a side or the originate point  $C(0,0)$ . Then, we just had to calculate the length of this path. The result is small when signals are close and becomes larger when signals differ.

As we expected our algorithm to be invariant to 3D translational motion and to different bodies, we had to focus not on coordinates of each joint, but on angles. The trigonometric calculations described in [12] were applied to measure the (i) working angles (range of motion) and (ii) compensation angles for the studied therapeutic movements: hip abduction, slow flexion of the hip and knee, hip extension, and the frontward–sideway–backward sequence (FSBS). During a rehabilitation session, it is important to identify both the movement amplitude that the patient should reach, also called range of motion (ROM) and the compensatory movements that should be avoided for an appropriate re-education of the injured limbs.

Because the raw input data provided by Kinect were too noisy, the last step consisted in applying a low-pass filter to the signal. The Butterworth filter available in the *Python library scipy.signal* was used to smooth the signal. Finally, the assessment module was implemented in such a manner to compare the DTW cost of each described angle (ROM and compensations) to an empirically defined threshold. If the calculated distance was higher than the threshold, the angle of the movement was classified as wrong.

## 2) Experimental protocol

The experiment consisted in comparing the assessment performed by four physiotherapists with the assessment made by the algorithm. Seven subjects participated in the study. They had to perform four different movements of rehabilitation: hip abduction, slow flexion of the hip and knee, hip extension, and FSBS. The movements were repeated 11 times each (one of them was used as a reference). The participant had to introduce a different error in the execution of the movement at each repetition (e.g., wrong ROM or different kinds of compensation). Fig. 3 represents the workflow of the real-time assessment of the movements.

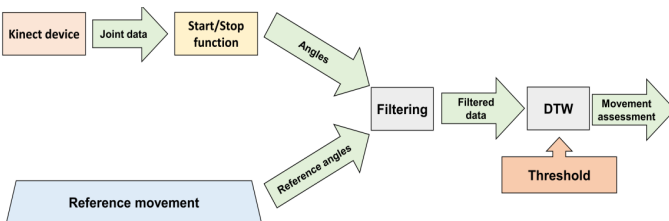


Fig. 3. Workflow of data processing.

## B. Results

Overall, the outcome of the algorithm matches at 88% with the assessment made by the physiotherapists. The accuracy of this matching depends on the threshold of tolerated difference between the reference movement and the analyzed movement (Fig. 4).

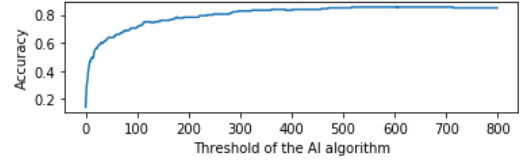


Fig. 4. Overall variation of accuracy depending on the threshold.

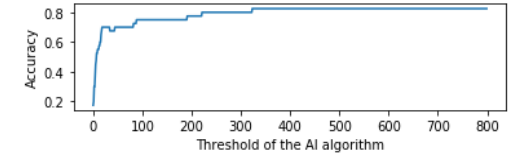


Fig. 5. FSBS variation of accuracy depending on the threshold.

Moreover, there is a certain difference of congruence between the human assessment and the machine assessment according to the type of exercise. Fig. 5 shows that the accuracy rate can reach 92.5% for FSBS.

## IV. PAIN RECOGNITION

### A. Method

An SVM algorithm was used to identify pain. An SVM constructs a hyperplane or a set of hyperplanes in high or infinite dimensional space, which can be used for classification or regression. Intuitively, good separation is achieved by the hyperplane that has the largest distance to the nearest training data points of any class. The model was developed with the *Python library scikit-learn* [13] and was trained on a set of different databases such as JAFFE [14], Tarrlab [15], and FER-2013 [16]. These databases were made up of facial expression images. A preliminary process consisted in selecting relevant pictures for the purpose of the study. Then, the images were converted into vectors of the same size ( $256 \times 256$  pixels) and in a gray scale. These images were also labeled with the correct emotion. The final training dataset was composed of 65,536 instances. Principal component analysis was applied to reduce the dimension of the problem to 150 features. The test set was built with 63 pictures extracted from 5 s videos of patients expressing pain (40 instances) and no pain (23 instances). Fig. 6 represents a sample of emotions for a determined subject.



Fig. 6. Sample of the tested images.

### B. Results

Table I shows that the algorithm based on support vector classification has 100% accuracy to distinguish between a

participant expressing pain and no pain. Equation (3) describes the calculation of the precision ( $p$ ):

$$p = tp / (tp + fp), \quad (3)$$

where  $tp$  is the number of true positives and  $fp$  is the number of false positives. This index defines the capacity of the model to not classify as positive an instance that is negative. Equation (4) presents the calculation of the recall ( $r$ ):

$$r = tp / (tp + fn), \quad (4)$$

where  $tp$  is the number of true positives and  $fn$  is the number of false negatives. This index represents the ability of the classifier to find all the positive instances. The support is the number of occurrences for each class.

TABLE I: RESULT OF THE MAIN CLASSIFICATION METRICS

Emotion	Precision	Recall	Support
Pain	1.00	1.00	40
No pain	1.00	1.00	23
<b>Total</b>	<b>1.00</b>	<b>1.00</b>	<b>63</b>

## V. CONCLUSIONS AND PERSPECTIVES

This study consisted in developing an artificial intelligence module for a tele-rehabilitation platform in order to assess the quality of the movement and the level of pain during the completion of therapeutic exercises. A DTW approach was used to evaluate the motion by comparison to a reference. The experimental results show a high correlation between the assessment performed by the computer and by health professionals. Ten percent of the discrepancies between the two evaluations have both technical and human explanations. The accuracy is affected by partial occlusions of some joints inherent to the motion capture device. In addition, we observed certain inconsistencies in the judgment of the therapists, which limits the reliability of the human reference. The recognition of pain based on a machine learning algorithm (SVM) was even better, because all the emotions were correctly classified. Nevertheless, this result will need to be confirmed at a larger scale (more participants and additional emotions) and under more challenging environmental conditions (e.g., different distances from the sensor and illumination characteristics). Overall, this work shows that the selected methods (DTW and SVM) seem appropriate for a real-time assessment of gestures and pain in the case study of a web-based system for motor rehabilitation.

## REFERENCES

- [1] Y. Rybarczyk, J. Kleine Deters, C. Cointe, and D. Esparza, "Smart Web-based platform to support physical rehabilitation," *Sensors*, vol. 18, pp. 1-21, 2018.
- [2] Y. Rybarczyk, and M. J. Goncalves, "WebLisling: a web-based therapeutic platform for rehabilitation of aphasic patients," *IEEE Latin America Transactions*, vol. 14, pp. 3921-3927, 2016.
- [3] S. Mani, S. Sharma, B. Omar, A. Paungmali, and L. Joseph, "Validity and reliability of Internet-based physiotherapy assessment for musculoskeletal disorders: a systematic review," *Journal of Telemedicine and Telecare*, vol. 23, pp. 379-391, 2017.
- [4] Y. Rybarczyk, J. Kleine Deters, A. Aladro Gonzalvo, M. Gonzalez, S. Villarreal, and D. Esparza, "ePHoRt project: a web-based platform for home motor rehabilitation," in *Proc. 5<sup>th</sup> World Conference on*

*Information Systems and Technologies, Advances in Intelligent Systems and Computing*, pp. 609-618, Portugal, March 2017.

- [5] J. Kleine Deters, and Y. Rybarczyk, "Hidden Markov Model approach for the assessment of tele-rehabilitation exercises," *International Journal of Artificial Intelligence*, vol. 16, pp. 1-19, 2018.
- [6] Y. Rybarczyk, J. Kleine Deters, A. Aladro Gonzalvo, M. Gonzalez, S. Villarreal, D. Esparza, and I. Nunes, "Recognition of physiotherapeutic exercises through DTW and low-cost vision-based motion capture," in *Proc. 8<sup>th</sup> International Conference on Applied Human Factors and Ergonomics, Advances in Intelligent Systems and Computing*, USA, June 2017, pp. 348-360, doi: 10.1007/978-3-319-60366-7\_33.
- [7] L. Y. Mano, "Emotional condition in the health smart homes environment: emotion recognition using ensemble of classifiers," in *Proc. 8<sup>th</sup> IEEE International Conference on INnovations in Intelligent SysTems and Applications*, Greece, July 2018.
- [8] R. Niewiadomski, M. Mancini, G. Varni, G. Volpe, and A. Camurri, "Automated laughter detection from full-body movements," *IEEE Transactions on Human-Machine Systems*, vol. 46, pp. 113-123, 2016.
- [9] S. Jaiswal, and R. Kumar, *Learning Django Web Development*, Birmingham, U.K.: Packt Publishing Ltd, 2015.
- [10] G. C. Hillar, *Django RESTful Web Services: The Easiest Way to Build Python RESTful APIs and Web Services with Django*. Birmingham, U.K.: Packt Publishing Ltd, 2018.
- [11] N. Tisch. (October 2017). Kinectron a realtime peer server for kinect 2 [Online]. Available: <https://kinectron.github.io/docs/server.html>
- [12] Y. Rybarczyk, L. Leconte, J. Pérez Medina, K. Jimenes, P. Acosta-Vargas, and D. Esparza, "Telerehabilitation platform for post-arthroplasty recovery: a Dynamic Time Warping approach," *Proc. 3<sup>rd</sup> IEEE Ecuador Technical Chapters Meeting*, Ecuador, October 2018.
- [13] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, and B. Thirion, "Scikit-learn: machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.
- [14] M. J. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba, "Coding facial expressions with Gabor wavelets," in *Proc. 3<sup>rd</sup> IEEE International Conference on Automatic Face and Gesture Recognition*, Japan, April 1998, pp. 200-205, doi: 10.1109/AFGR.1998.670949.
- [15] Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University. [Online]. Available: <http://www.tarlab.org>
- [16] I. J. Goodfellow *et al.*, "Challenges in representation learning: a report on three machine learning contests," *Neural Networks*, vol. 64, pp. 59-63, 2015.



**Yves Rybarczyk** was born in Beaumont, France, in 1975. He received a Ph.D. degree in robotics from the University of Evry, France, in 2004. His teaching and research activities focus on human-machine interaction, artificial intelligence, models, and simulation.

He was an assistant professor in the Department of Electrical Engineering at the Nova University of Lisbon, Portugal, between 2007 and 2015. Then, he moved to South America where he held the position of associate professor and head of the Intelligent & Interactive Systems Lab (SI<sup>2</sup> Lab) at the Universidad de Las Américas, Ecuador, until 2019. Currently, he is associate professor in the School of Informatics at the University of Skövde, Sweden. He has participated in several projects in the area of the modeling of complex systems (e.g., prediction of environmental phenomena) and the development of interactive systems (e.g., assistive and rehabilitation technologies). He was involved in several national and international funded projects (five of them as coordinator), and he is the author of over 70 scientific papers and book chapters.

Prof. Rybarczyk was the chairman of the ninth International Summer Workshop on Multimodal Interfaces (eNTERFACE), and he is currently a member of the steering committee of this scientific event. Recently, he was also elected as a member of the editorial board of the journal IEEE Latin America Transactions.



**Louis Leconte** is a normalien civil servant of the Ecole Normale Supérieure (ENS) Paris-Saclay, France. He has done two years of Classe préparatoire (2014–2016) in Collège Stanislas, Paris, France. He received a bachelor's degree in electrical engineering in 2016 at the ENS Paris-Saclay and is now studying for a master's degree. His research activities focus on signal processing and artificial intelligence. He is currently doing an internship in Montreal, Canada, at the CoCoLab of the Université de Montréal (UdeM).





**Jorge Luis Pérez Medina** received his M.S. degree in information system from the Joseph Fourier University of Grenoble, France, in 2006 and a Ph.D. in computer science from the same university in 2010. He was responsible for a Human–Computer Interaction course at the Université Catholique de Louvain, Belgium, during the 2015–2016 academic year. He is currently professor and researcher at Intelligent & Interactive Systems Lab (SI<sup>2</sup> Lab), Universidad de Las Américas, Quito, Ecuador. Throughout his career, he has developed a deep knowledge of computer science and human–computer interaction. His interest in scientific research has been promoted by the desire to facilitate the work of IT managers and model designers by helping them in choosing methods, models, and modeling environments adapted to their specific needs.



**Karina Jimenes Vargas** collaborates in a tele-rehabilitation project for the Intelligent & Interactive Systems Lab (SI<sup>2</sup> Lab) of the Universidad de Las Américas, Quito, Ecuador, as an assistant researcher. She has a master's degree in computer science, and her primary areas of interest are software development, artificial intelligence, and data mining.



**Patricia Acosta-Vargas** is currently studying for her Ph.D. in computer science at the University of Alicante, Spain. She is a member of the Intelligent and Interactive Systems Lab (SI<sup>2</sup> Lab), Universidad de Las Américas, Ecuador. In recent years, she has been involved in web accessibility research. She also participates in the tele-rehabilitation project for the CEDIA network. Her work has been published in IEEE, IEEE Access, ACM, and Springer. Her research interests involve engineering applied to education, health, government, web accessibility metrics, and heuristics. She collaborates as a reviewer of scientific articles for several journals and conferences. She is the author of several books, journals, and more than 20 scientific articles, indexed in Scopus and WoS.



**Danilo Esparza** is a physical therapist and a full-time professor at the Universidad de Las Américas, Ecuador. He received a PhD in physical activity and sport science from the University of Orleans, France. His areas of interest are the analysis of movement, kinematic, musculoskeletal pain, manual lymphatic drainage, and motor control and learning. He is currently teaching and researching in neurology, anatomy, and practice-based evidence in physical therapy. He is the author/coauthor of several articles in peer-reviewed journals in the physical therapy field.