



**HAL**  
open science

# Multi-level Motion Analysis for Physical Exercises Assessment in Kinaesthetic Rehabilitation

Maxime Devanne, Sao Mai Nguyen

► **To cite this version:**

Maxime Devanne, Sao Mai Nguyen. Multi-level Motion Analysis for Physical Exercises Assessment in Kinaesthetic Rehabilitation. *Humanoids 2017: IEEE-RAS 17th International Conference on Humanoid Robots*, Nov 2017, Birmingham, United Kingdom. 10.1109/HUMANOIDS.2017.8246923 . hal-01886757v1

**HAL Id: hal-01886757**

**<https://hal.science/hal-01886757v1>**

Submitted on 9 Oct 2018 (v1), last revised 18 Nov 2021 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Multi-level Motion Analysis for Physical Exercises Assessment in Kinaesthetic Rehabilitation

Maxime Devanne<sup>1</sup> and Sao Mai Nguyen<sup>1</sup>

**Abstract**—Analyzing and understanding human motion is a major research problem widely investigated in the last decades in various application domains. In this work, we address the problem of human motion analysis in the context of kinaesthetic rehabilitation using a robot coach system which should be able to learn how to perform a rehabilitation exercise as well as assess patients’ movements. For that purpose, human motion analysis is crucial. We develop a human motion analysis method for learning a probabilistic representation of ideal movements from expert demonstrations. A Gaussian Mixture Model is employed from position and orientation features captured using a Microsoft Kinect v2. For assessing patients’ movements, we propose a real-time multi-level analysis to both temporally and spatially identify and explain body part errors. This allows the robot to provide coaching advice to make the patient improve his movements. The evaluation on three rehabilitation exercises shows the potential of the proposed approach for learning and assessing kinaesthetic movements.

## I. INTRODUCTION

Low back pain is a leading cause disabling people particularly affecting the elderly, whose proportion in European societies keeps rising, incurring growing concern about healthcare. 50 to 80% of the world population suffers at a given moment from back pain which makes it in the lead in terms of health problems occurrence frequency [2]. To tackle this chronic low back pain, active rehabilitation (physical rehabilitation as opposed to cognitive-behavioral principles) is considered as more effective than usual care [3].

With this perspective, solutions are being developed based on assistive technology and particularly robotics. In KERAAL project, we are developing a robot coach for physical rehabilitation exercises. The goal is to increase the time patients spend exercising, by alleviating the lack of time a physiotherapist can spend monitoring a patient [4]. The system is composed of a low cost stereo vision camera (Microsoft Kinect v2) and an open source humanoid robot called Poppy. Figure 1 shows the setting of our system. The kinect sensor is used to capture human motion of both the therapists and the patients. In addition to its low cost and ease of use, it is a seamless sensor as no markers or specific suit are needed, nor time for setup. The Poppy robot is used to demonstrate exercises to the patient and to provide him feedback. In the literature on robot coach systems, the Nao robot is often used to achieve this task [5], [6]. However,

such a robot has only one DOF for the trunk. This makes the performance of many rehabilitation exercises for low back pain impossible or not natural. Conversely, the Poppy robot is designed to be anthropomorphic [7] with 25 degrees of freedom (DOF) including a 5 DOFs articulated trunk. Given its unique capability of realizing movements of the lumbar spine, this robot fits well with the objectives of rehabilitation programs dedicated to low back pain. We aim to develop a robot coach capable of understanding the requirements of a rehabilitation exercise from the medical expert. Then, it should be capable of demonstrating rehabilitation exercises to a patient. Currently, the learning of rehabilitation exercises is achieved by manually moving the robot in collaboration with physiotherapists and recording motor angles at each timestamps. Finally, the robot should watch him/her carry out the exercise and give him/her feedback so as to improve his/her performance and keep them motivated. To achieve these goals, human motion analysis is crucial. In that context we propose in this paper a multi-level human motion analysis to evaluate and assess rehabilitation movements performed in front of a RGB-D camera.



Fig. 1: Setting of the system including a Microsoft Kinect v2 and an open source humanoid robot called Poppy.

The rest of this paper is organized as follows: Section II reviews existing approaches to address the problem of physical rehabilitation. Section III describes our proposed multi-level human motion analysis approach for both learning ideal rehabilitation movements and thoroughly assessing patients’ movements. The approach is evaluated on several kinaesthetic exercises in Section IV. Finally, Section V concludes the paper and investigates possible future work.

<sup>1</sup> IMT Atlantique, Lab-STICC, UBL, France  
maxime.devanne@imt-atlantique.fr,  
nguyensmai@gmail.com

\*The research work presented in this paper is partially supported by the EU FP7 grant ECHORD++ KERAAL and by the European Regional Fund (FEDER) via the VITAAL Contrat Plan Etat Region [1].

## II. RELATED WORK

In the literature, human motion analysis has been widely investigated in different contexts like action recognition, motion segmentation and fall detection. However, only few approaches addressed the challenge of physical rehabilitation through coaching robot systems. In [8], a robotic exercise tutor system for elderly people has been developed. A Nao robot is able to demonstrate several exercises and evaluate the performance of patients. In [9], the authors developed a system of imitation learning for daily physical exercises. The robot could learn new exercises from the therapists and be an exercise demonstrator.

### A. Representing human body

While in these approaches only joint angles derived from joint positions are considered as human motion features, other approaches also consider joint positions [4], [10] or relative transformations (translation and rotation) [11]. Relative positions or orientations are more robust to subjects' sizes and morphology variations. Relative position features are more suitable to describe exercise characteristics like "Place your hands at head's height". Orientation features better correspond to characteristics like "Stretch your arm horizontally". In our project we propose a combination of both relative position and orientation features. This allows to better describe human motion and capture important requirements of each exercise. To represent human body data, traditional Euclidean space is usually employed. However, an increasing number of approaches consider other spaces like Riemannian spaces to represent human postures and motion so as to handle the non linearity of human movements. Its proven effectiveness in human motion analysis for action recognition [12], [13] motivated us to also consider such non linear spaces for physical rehabilitation.

### B. Learning from experts

An important characteristic of a robot coach system for physical rehabilitation is its ability to learn exercises from experts. Whereas [9] has developed a neuro-genetic approach for imitation learning, [8] employs direct transformation of human joint angles into robot motor angles. To improve the imitation, several expert demonstrations are often appreciated. This allows to consider variations among demonstrations to identify which features of exercises are more important. In robotics, imitation learning also called Programming by Demonstration (PbD) has been widely investigated. Particularly, probabilistic methods based on Hidden Markov Models (HMM) and Gaussian Mixture Models (GMM) have proven successful for robots learning by observation of demonstrations such as in [14]. The Gaussian Mixture Model thus learned after a few demonstrations constitute a probabilistic description of the ideal movement, which is robust to noise and small errors in the training data. In this work, we propose a similar approach based on GMM to learn a model representing ideal exercises from human data captured using Kinect and represented on a Riemannian space combining position and orientation features.

### C. Assessing movements

Assessing the patient's performance to provide him adequate and personalized support such as performance feedback to help him correct his errors is also essential for an intelligent tutoring system. [9] only focuses on the algorithms for learning a model from experts and does not tackle the assessment of an imitation attempt. [8] based their automatic evaluation on the distance measure between the user's current arm angles and the specified goal arm angles. It is based on only one template movement. The use of probabilistic models such as GMM are more suitable to analyze deviation according to an ideal movement and has been for instance successfully applied for abnormal gait detection [15]. In addition, only providing overall performance feedback may be insufficient to help the patient correct his exercise. Spatial information about which body part is incorrect facilitates error understanding and improvement by the patient. In human motion analysis, movement segmentation is often adopted to face human motion complexity. In motion assessment, it has mainly been used for segmenting different repetitions of a movement in an online way [16]. In this work we propose to segment exercises online in motion primitives in order to locally analyze patients' movements. In addition, this allows to temporally localize errors which can be beneficial for the patient's understanding. Finally, automatically providing instructions on how to improve the movement can significantly help the patient perform the correct movement and keep him motivated throughout the rehabilitation session. Along these considerations, we propose a multi-level analysis of human motion for assessing physical exercises in a context of a robot coach system for rehabilitation.

## III. MULTI-LEVEL HUMAN MOTION ANALYSIS FOR PHYSICAL REHABILITATION

To guarantee an efficient intelligent tutoring system within our robot coach, two phases have been identified, the learning phase and the assessment phase.

### A. Human Motion Learning

In robotics, Gaussian Mixture Models (GMM) have proven successful for robots learning by observation of demonstrations [14]. In our context of physical rehabilitation, expert demonstrations correspond to human motion sequences. Hence an efficient representation of human motion is needed.

1) *Human Pose Space*: The Microsoft Kinect provides in real-time the 3D position and the orientation of 25 joints forming a humanoid skeleton. Figure 2a shows the structure of the skeleton detected using Kinect. In this work, we focus our analysis on upper body joints. In addition, we do not consider hand and finger joints as they may be very noisy and do not add relevant information for physical rehabilitation exercises. As a result the number of joints is  $J = 11$ .

As quaternions provided by Kinect sensors represent orientations relative to the parent's bones, it naturally allows invariance to subjects' sizes and positions. However, this is not the case for joint positions. Therefore, we employ normalized relative positions. For a given joint  $j$ , its 3D

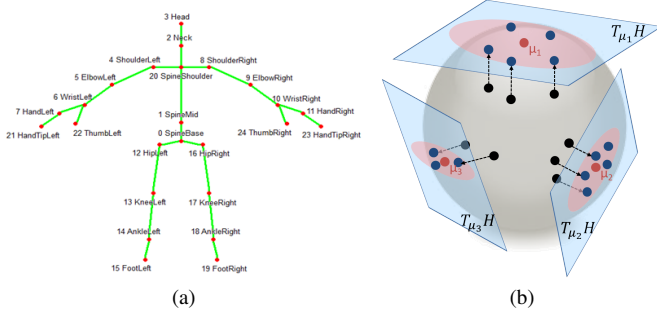


Fig. 2: Representation of the human pose. (a) The structure of the skeleton captured by Kinect. (b) Illustration of the human pose space  $\mathcal{H}$  with three Gaussians computed on tangent space at means  $\mu_k$  (red dots). Black dots are elements on the manifold and blue dots are their projection on tangent spaces.

position  $P_j$  is computed relatively to the Spine Shoulder absolute position  $p_{ss}$  and normalized using the length  $L_{spine}$  of the spine bone (between Spine Shoulder joint and Spine Mid joint). As a result, a skeleton pose  $y_t$  at frame  $t$  can be represented as:

$$y_t = [O_1, P_1, O_2, P_2, \dots, O_J, P_J], \quad (1)$$

where  $O_j$  is the orientation of joint  $j$  and  $P_j$  its relative position defined as  $P_j = (p_j - p_{ss})/L_{spine}$ .

Note that data extracted from Kinect sensor are processed using a Butterworth filter to remove noise.

While joint positions are naturally viewed in 3D Euclidean space, quaternions can be represented as elements of the 3-sphere  $\mathcal{S}^3$  which is a 3 dimensional Riemannian manifold. A Riemannian manifold is a smooth space that locally resembles Euclidean space and is equipped with the Riemannian metric defined on the tangent space at each point of the manifold [17]. Furthermore, the Cartesian product of several Riemannian manifolds is again a Riemannian manifold. This property allows us to consider combinations of joint quaternions and positions corresponding to the whole body. We define the human pose space  $\mathcal{H}$  as the Cartesian product of quaternion and position of all skeleton joints:

$$\mathcal{H} = \mathbb{R}^3 \times \mathcal{S}^3 \times \mathbb{R}^3 \times \mathcal{S}^3 \times \dots \times \mathbb{R}^3 \times \mathcal{S}^3. \quad (2)$$

2) *Imitation Learning Algorithm*: For learning a model representing an ideal movement from several expert demonstrations, we have employed the recent framework proposed in [18] extending common imitation learning techniques, such as GMM, to Riemannian manifolds. Such framework is particularly convenient for our work as our skeleton features are represented in the Riemannian human pose space as explained above. Indeed, as Riemannian manifolds are non linear spaces, we cannot compute statistics as mean and covariance similarly to Euclidean spaces.

A common way to handle the non-linearity of Riemannian manifolds  $\mathcal{M}$  is to consider tangent spaces  $\mathcal{T}_p\mathcal{M}$  at a reference point  $p \in \mathcal{M}$  as a linear approximation of the

neighborhood of  $p$ . To map a point  $g$  from the manifold to the tangent space at  $p$  resulting in  $v$ , the distance preserving logarithmic map is defined as  $\text{Log}_p(\cdot) : \mathcal{M} \rightarrow \mathcal{T}_p\mathcal{M}$ . Conversely, the exponential map  $\text{Exp}_p(\cdot) : \mathcal{T}_p\mathcal{M} \rightarrow \mathcal{M}$  allows to go back from the tangent space to the manifold. More details about exponential and logarithmic map computation on the  $\mathcal{S}^3$  manifold can be found in [18]. As the human pose space  $\mathcal{H}$  is the Cartesian product of several manifolds, corresponding exponential and logarithmic mapping are obtained by concatenating individual functions of each sub-manifold.

Using linear tangent spaces, we can compute approximated multivariate Gaussians on the human pose space. The mean  $\mu$  of  $N$  points  $p_i$  on the human pose space can be obtained using [19]:

$$\mu = \arg \min_p \sum_{i=1}^N d(\mu, p_i)^2, \quad (3)$$

where  $d(\mu, p)$  is the geodesic distance on the manifold which can be written using logarithmic map as  $d(\mu, p) = \|\text{Log}_\mu(p)\|$ . Such mean is called the Riemannian center of mass [19] and is obtained by an iterative process until no change. Once a mean point is computed, the covariance matrix  $\Sigma$  can be computed from points  $p_i$  projected into the tangent space at  $\mu$  using the logarithmic map  $\text{Log}_\mu(p_i)$ . We can then learn a Gaussian Mixture Model defined as:

$$p(x) = \sum_{k=1}^K \phi_k \mathcal{N}(x | \mu_k, \Sigma_k), \quad (5)$$

where  $x$  encodes both the human pose  $y_t$  and the timestamps  $t$ ,  $K$  is the number of Gaussians,  $\phi_k$  the weight of the  $k$ -th Gaussian,  $\mu_k$  the Riemannian center of mass of the  $k$ -th Gaussian computed on the manifold and  $\Sigma_k$  the covariance matrix of the  $k$ -th Gaussian. The parameters  $\phi_k$ ,  $\mu_k$  and  $\Sigma_k$  are learned using Expectation-Maximization on the human pose Space [20]. Figure 2b illustrates the human pose space  $\mathcal{H}$  with three Gaussians computed on tangent space at means. We note that expert demonstration are first temporally aligned using dynamic programming. This allows us to handle possible velocity variations among expert demonstrations.

3) *Ideal movement generation*: Once a model is learned for each exercise, we can generate an optimal sequence using Gaussian Mixture Regression (GMR) which approximates the sequence using a single Gaussian:

$$p(\hat{x}|t) \approx \mathcal{N}(\hat{\mu}, \hat{\Sigma}). \quad (6)$$

As described in [18], to remedy the non linearity of the manifold,  $\hat{\mu}$  is computed in an iterative process similarly to the Riemannian center of mass and  $\hat{\Sigma}$  is computed on the tangent space at  $\hat{\mu}$ . By evaluating  $\hat{x}$  for successive values of  $t$ , we obtain a generalized form of the ideal motion  $\hat{X}$ . This optimal motion sequence will be used as a reference to evaluate a motion sequence of a patient.

## B. Human motion assessment

In order to evaluate a test sequence of a patient and provide him feedback, we propose a multi-level analysis of the movement, as illustrated in Figure 3.

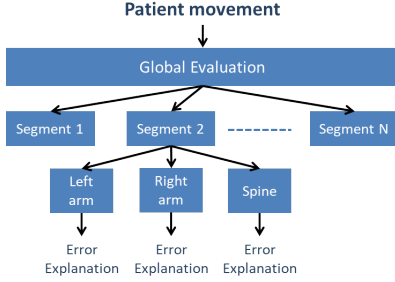


Fig. 3: Illustration of the multi-level analysis of patients' movement.

1) *Global evaluation*: To evaluate a patients' movement we first temporally aligned the motion sequence to the ideal movement computed above using dynamic programming. Then, we compute the log-likelihood that the given sequence  $X$  has been generated by the learned Gaussian Mixture Model of the corresponding exercise:

$$\ln(p(X|\phi, \mu, \sigma)) = \sum_{t=1}^T \ln\left(\sum_{k=1}^K \phi_k \mathcal{N}(x_t|\mu_k, \Sigma_k)\right). \quad (7)$$

As the log-likelihood value may not be significant for a patient or for a physiotherapist, we can use thresholds to translate the log-likelihood into percentage of success.

2) *Temporal segment analysis*: While providing a global score for the sequence allows the patient to have a global idea on his performance, it is also interesting to know which part of the exercise is not performed correctly. Thus, we propose an analysis based on temporal segments. We analyze the motion within a window of length  $W$  by computing the standard deviation among data included in the temporal window. The standard deviation within a window is computed on the human pose space as:

$$\sigma = \sqrt{\frac{1}{W} \sum_{t=1}^W d(\mu_W, y_t)^2}, \quad (8)$$

where  $\mu_W$  is the Riemannian center of mass of all skeleton poses of the sequence included in the window. By sliding the window along the sequence we can obtain the evolution of the motion along the exercise. When the value is low, it means that there is no motion. It corresponds to transition between two elementary motions. Thus, by detecting local minima, we are able to detect the transition frames and thus identify the different temporal segments. However, for some rehabilitation exercises, the holding postures that a patient should keep for a while are more important than transitions between them. With the the segmentation strategy proposed above, we focus our analysis on transitions which may not be suitable for some cases. That is why we propose a second

segmentation strategy by detecting key frames when the motion value crosses the threshold. With this strategy, we are able to differentiate transition movements and holding postures. According to the evaluated exercise, we can select the corresponding segmentation strategy. Figure 4 shows an example of exercise segmentation using our two strategies. Detected key frames representing boundaries between temporal segments are depicted in red color.

In addition, by detecting when motion value crosses the threshold for the first time, we can identify the beginning of the exercise. We select the beginning 10 frames before such crossing point. This allows us to compare sequences starting at the same time. The detected starting frames are depicted in green color in Figure 4.

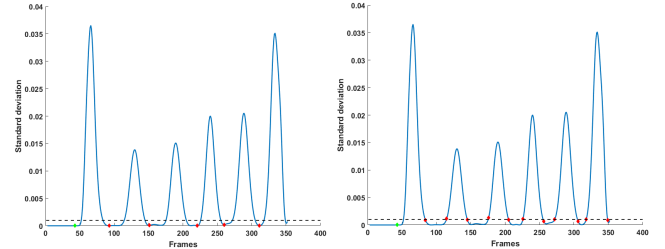


Fig. 4: Motion evolution for segmentation. (Left) First strategy focuses on elementary motion by detecting transitions when the motion becomes slower. (Right) Second strategy differentiates transition motions and holding postures. Red points correspond to boundaries between two temporal segments. Green points correspond to the beginning of the exercise. The black line corresponds to the threshold used to detect starting frames and boundaries in the second strategy.

3) *Body part analysis*: We also propose a local analysis to differentiate different body parts. This allows to identify which body part is more responsible of the error and gives more precise feedback to the patient. As the skeleton pose  $x_t$  is the concatenation of all joint features, we can compute the log-likelihood for data corresponding to joints of the desired body part to obtain a score for each body part. As shown in Figure 3, we differentiate the two arms and the spine.

4) *Error explanation*: Until now, the evaluation of a motion sequence only provides a global score of success. However in the case a movement is not correctly performed, it is very important to understand the reason why it has been detected as incorrect and propose a solution how to improve the movement. When an error is detected for a given body part, we propose to evaluate the log-likelihood of the position and orientation features of each joint of the body part to detect which joint and which feature are more responsible of the incorrect movement. Then we compare the corresponding feature values of the test sequence  $X$  with the ideal movement  $\hat{X}$ . Suppose that the minimum log-likelihood value is obtained for the position feature of joint  $j$  at time  $t$ , we compute the difference between the two corresponding position features  $\hat{P}_j$  and  $P_j$  as  $\delta = (P_j - \hat{P}_j) / \hat{\sigma}_{P_j}^2$ , with  $\hat{\sigma}_{P_j}$  the standard deviation obtained from the covariance matrix  $\hat{\Sigma}$  for the corresponding position feature  $\hat{P}_j$  of the ideal

movement. By analyzing the 3-dimensional vector  $\delta$ , we can detect which coordinate deviates the most from the ideal movement. This allows us to explain the error. For instance, if the  $y$  coordinate is higher and positive, it means that the joint is too high in comparison to the ideal movement.

For orientation features the process is slightly different. As quaternion features  $O_j$  are on the  $\mathcal{S}^3$  space we first need to compute an intermediate difference  $v_d$  on the tangent space at  $\hat{O}_j$  as  $v_d = (\text{Log}_{\hat{O}_j}(O_j) - \hat{O}_j) / \hat{\sigma}_{\hat{O}_j}^2$ . Then,  $v_d$  is projected back to the manifold using exponential map, resulting in  $q_d$ . Finally  $\delta$  is computed as  $\delta = \phi(q_d) - \phi(\hat{O}_j)$ , where  $\phi(\cdot)$  is the function converting quaternion into Euler angles which are more suitable to explain orientation errors. For instance, if the  $x$  value of  $\delta$  is higher and negative, it means that the corresponding bone is too much leaned down.

We note that in order to be more robust to noise, we perform the error explanation process in a temporal window of size  $W$  around the frame  $t$  where the error is detected and report result based on the average through the window.

#### IV. EXPERIMENTAL RESULTS

We propose to evaluate our method on three rehabilitation exercises selected in collaboration with physiotherapists<sup>1</sup>. In the first exercise, the patient raises his arms horizontally and then rotates the trunk on both sides. In the second exercise, the patient should raise an arm upright then lean on the side. The same is the performed for the other arm. In the third exercise, the patient should lift his arms in front of him with the elbows bent and then spread the arms.

##### A. Dataset

Under the supervision of a physiotherapist, we collect a training database of two different subjects performing each exercise three times. Then we collect test data of a third subject performing the same exercises twice. In addition, this test subject performs incorrect exercises by simulating errors. For the first exercise, the arms are not enough raised. For the second exercise, the subject does not tilt the arm and keep it straight. In the third exercise, the arms are not enough raised. Therefore, for each exercise, we have both test data which are correct and incorrect<sup>2</sup>.

##### B. Global Assessment

First, we compute the global and body part scores for each test sequence. Results are reported in Table I. We can first observe that the obtained scores for correct exercises are much higher than those obtained for incorrect exercise. This shows that our method is able to detect when an exercise is not correctly performed. Moreover, as different subjects are used for training and testing, it shows that our method is independent of subjects. In addition, we can see that correct exercise 1 obtained lower scores than other correct exercises. This can be explained by the fact that when the subject is rotating the trunk, one arm is behind the chest and may be incorrectly detected by the Kinect sensor. Thus it can affect

the overall score but not as much as incorrect exercises. Finally, by observing the scores of the spine, we can see that a score of 100% is obtained for most cases. This is expected as the spine should always be straight. However, for the incorrect exercise 2, when the subject keeps his back straight and does not lean on the side, we can see that the score for the spine is 90%. This shows that we are able to detect the error. However, in comparison with other errors in the arms, the score is still high. This shows that errors in different body parts do not affect the score similarly. Adding some weights for each body parts could allow us to overcome this limitation. This will be part of our future work.

TABLE I: Evaluation results for all test sequences. Global scores and body part scores are reported for each sample.

	Global	Left arm	Spine	Right arm
Ex1 Correct	72%	75%	100%	63%
Ex1 Correct	82%	86%	100%	73%
Ex1 Error	<b>12%</b>	<b>31%</b>	100%	<b>19%</b>
Ex2 Correct	92%	98%	100%	96%
Ex2 Correct	98%	100%	100%	99%
Ex2 Error	<b>24%</b>	<b>37%</b>	<b>90%</b>	<b>49%</b>
Ex3 Correct	95%	90%	100%	100%
Ex3 Correct	93%	99%	100%	90%
Ex3 Error	<b>27%</b>	<b>47%</b>	100%	<b>40%</b>

##### C. Assessing each temporal segment

In a second time, we compute scores for each temporal segment according to the two segmentation strategies. Figure 5 illustrates results for the incorrect exercise 3. By using the first segmentation strategy where only transition motions are considered, we can see that all temporal segments are detected as incorrect but the first one and last one obtained higher scores. For the first segment, this is because the movement of raising arms is correct most of the time except near the end as the arms are not enough raised. As a result the mean score of this temporal segment is quite low. The same applies for the last segment where the subject lowers his arms. Only the beginning of the segments is incorrect as it starts with the arms too low. In comparison, if we use the second segmentation strategy, the first and last segments are not detected as incorrect. This is because the second strategy differentiates the movement and the holding postures. This shows that for this third exercise, the second strategy is more suitable. It allows to detect the error on the holding postures and not on the previous transition motion which is correct.

Finally, we show two examples of explanation of detected errors for exercises 2 and 3. These examples are illustrated in Figure 6. Explanation sentences are build automatically according to the evaluation of the patient's movements. In our rehabilitation scenario, such sentences are sent to the Poppy robot which enunciates them using a Text To Speech system.

These experimental results show that our method is able to assess patients' exercises and provide corresponding feedback indicating where and why the movement is not correct.

<sup>1</sup>Videos of exercises are available at: [www.keraal.enstb.org/exercises.html](http://www.keraal.enstb.org/exercises.html)

<sup>2</sup>Videos available at: [www.keraal.enstb.org/incorrectexercises.html](http://www.keraal.enstb.org/incorrectexercises.html)

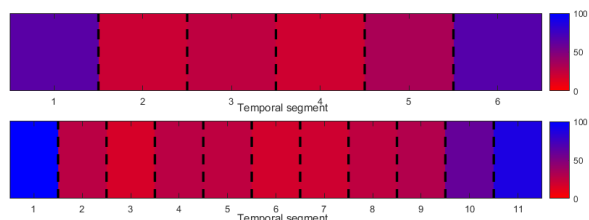


Fig. 5: Temporal evaluation. Scores are computed for each temporal segment according to the first strategy (top) and the second strategy (bottom). Red colors mean low scores, while blue colors represent correct scores. Black dot lines represents boundaries between temporal segments.

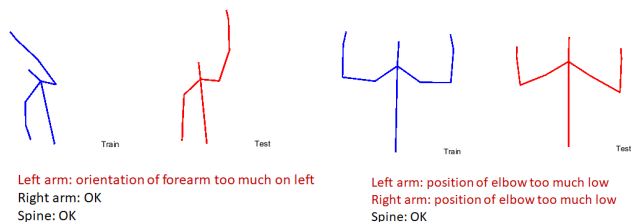


Fig. 6: Error explanation for exercises 2 (left) and 3 (right).

## V. CONCLUSIONS

We have proposed a human motion analysis method for physical exercises assessment and compared two segmentation strategies providing temporally analysis of patients' performances. Our method is able to learn an ideal movement from position and orientation features of expert demonstrations. It allows to consider variations among demonstrations and identify which part of the exercise is important. In addition, our multi-level analysis allows to provide detailed feedback of detected errors including the body parts, the temporal segment and how to correct the error. Evaluation on three different rehabilitation exercises targeting low back pain demonstrated promising results.

However, these are only preliminary results as the evaluation is performed on three rehabilitation exercises with healthy subjects. In that sense we will evaluate our method within the whole integrated robot coach system in a clinical environment with patients suffering from low-back pain. These experiments will allow us to evaluate our method in a real scenario with a larger set of users. Moreover, the acceptability of the Poppy robot as a robot coach will also be analyzed. In addition, as mentioned above, we plan to investigate methods to convert human motion to robot motion to make our robot automatically learn exercises by observing expert demonstrations. Finally, we plan to extend our method to be able to consider cases when a patient cannot perform correctly an exercise due to physical limitation or pain. With the current method his evaluation would be always incorrect. As a result, this could affect his motivation. To overcome this point, an adaptive model will be investigated to consider personal characteristic of patients.

## REFERENCES

- [1] A. Thépaut, S. M. Nguyen, and C. Lohr, "Télé-rééducation fonctionnelle dans le cadre du projet VITAAL," *Actualités en médecine physique et de réadaptation*, no. 1, pp. 18–19, Mar. 2017. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01514898>
- [2] W. S. G. on the Burden of Musculoskeletal Conditions at the Start of the New Millennium *et al.*, "The burden of musculoskeletal conditions at the start of the new millennium?" *World Health Organization technical report series*, vol. 919, p. i, 2003.
- [3] P. Kent and P. Kjaer, "The efficacy of targeted interventions for modifiable psychosocial risk factors of persistent nonspecific low back pain—a systematic review," *Manual therapy*, vol. 17, no. 5, pp. 385–401, 2012.
- [4] S. M. Nguyen, P. Tanguy, and O. Rémy-Néris, "Computational Architecture of a Robot Coach for Physical Exercises in Kinesthetic Rehabilitation," in *RO-MAN 2016 : IEEE International Symposium on Human and Robot Interactive Communication*, 2016, pp. 1–6.
- [5] S. Schneider and F. Kümmert, "Exercising with a humanoid companion is more effective than exercising alone," in *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*. IEEE, 2016, pp. 495–501.
- [6] D. Ramgoolam, E. Russell, and A. B. Williams, "Towards a social and mobile humanoid exercise coach," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 2014, pp. 274–275.
- [7] M. Lapeyre, "Poppy: open-source, 3d printed and fully-modular robotic platform for science, art and education," Ph.D. dissertation, Université de Bordeaux, 2014.
- [8] B. Gorer, A. A. Salah, and H. L. Akn, "An autonomous robotic exercise tutor for elderly people," *Autonomous Robots*, vol. 41, no. 3, pp. 657–678, 7 2017.
- [9] T. Obo, C. K. Loo, and N. Kubota, "Imitation learning for daily exercise support with robot partner," in *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on*. IEEE, 2015, pp. 752–757.
- [10] Y. Tao, H. Hu, and H. Zhou, "Integration of vision and inertial sensors for 3d arm motion tracking in home-based rehabilitation," *The International Journal of Robotics Research*, vol. 26, no. 6, pp. 607–624, 2007.
- [11] Q. Wang, G. Kurillo, F. Ofli, and R. Bajcsy, "Remote health coaching system and human motion data analysis for physical therapy with microsoft kinect," *arXiv preprint arXiv:1512.06492*, 2015.
- [12] R. Vemulapalli, F. Arrate, and R. Chellappa, "Human action recognition by representing 3d skeletons as points in a lie group," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 588–595.
- [13] M. Devanne, S. Berretti, P. Pala, H. Wannous, M. Daoudi, and A. Del Bimbo, "Motion segment decomposition of rgb-d sequences for human behavior understanding," *Pattern Recognition*, vol. 61, pp. 222–233, 2017.
- [14] S. Calinon, F. Guenter, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 2, pp. 286–298, 2007.
- [15] M. Devanne, A. Wannous, S. Berretti, P. Pala, M. Daoudi, and A. Del Bimbo, "Learning shape variations of motion trajectories for gait analysis," in *International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico, 2016.
- [16] J. F.-S. Lin and D. Kulic, "Online segmentation of human motion for automated rehabilitation exercise analysis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 168–180, 2014.
- [17] J. Jost, *Riemannian Geometry and Geometric Analysis*, ser. Springer Universitat texts. Springer, 2005. [Online]. Available: <https://books.google.fr/books?id=uVTB5c35F0C>
- [18] M. Zeestraten, I. Havoutis, J. Silvério, S. Calinon, and D. G. Caldwell, "An approach for imitation learning on Riemannian manifolds," *IEEE Robotics and Automation Letters (RA-L)*, vol. 2, no. 3, pp. 1240–1247, June 2017.
- [19] H. Karcher, "Riemannian center of mass and mollifier smoothing," *Comm. on Pure and Applied Math.*, vol. 30, pp. 509–541, 1977.
- [20] E. Simo-Serra, C. Torras, and F. Moreno-Noguer, "3d human pose tracking priors using geodesic mixture models," *International Journal of Computer Vision*, pp. 1–21, 2016.