Teaching psychomotor skills using machine learning for error detection

Benjamin Paaßen[1][0000−0002−3899−2450] and Miloš Kravčík[1][0000−0003−1224−1250]

German Research Center for Artificial Intelligence, 10559 Berlin, Germany
{benjamin.paassen,milos.kravcik}@dfki.de

Abstract. Learning psychomotor skills is challenging because motion is fast, relies strongly on subconscious mechanisms, and instruction typically disrupts the activity. As such, learners would profit from mechanisms that can react swiftly, raise subtle mistakes to the conscious level, and do not disrupt the activity. In this paper, we sketch a machine learning-supported approach to provide feedback in two example scenarios: running, and interacting with a robot. For the running case, we provide an evaluation how motions can be compared to highlight deviations between student and expert motion.

Keywords: Psychomotor skills · Running · Human robot interaction · Dynamic movement primitives

1 Introduction

Teaching beneficial psychomotor skills – such as moving healthily or interacting with a robot companion – is challenging [3,6], because coaches need to infer mistakes from subtle clues in observable behavior, build a hypothesis regarding the underlying cause of the mistake, and verbalize an instruction that enables the learner to improve performance, even though the learner may not be conscious of the mistake or how to correct it [6]. Automatic mechanisms may be helpful to support instruction in such cases. An automatic feedback mechanism can perceive and analyze psychomotor activity with only split-second delay and, thus, provide feedback in almost real-time [3]. This permits learners to improve their psychomotor performance in a much faster loop: if they receive feedback, they can adapt during the activity and receive additional feedback immediately, whereas classic coaching would require to interrupt an activity, then receive verbal feedback, discuss it, and re-start the activity to check the improvement.

In this paper, we sketch a machine-learning-supported approach for feedback for two psychomotor activity scenarios, namely running and interacting with a robot. Our approach is intended as an inspiration for a general template that can be applied across a wide range of psychomotor skills. For the case of running, we provide a first analysis using dynamic movement primitives which shows that we can abstract from irrelevant variations in psychomotor data and hone in on deviations from expert demonstrations that may be indicative of mistakes.

2   Background and Related Work

To avoid injuries in running, a special wearable assistant was created, using an electrical muscle stimulation (EMS) device and an insole with force sensing resistors [2]. The results of the conducted study showed that EMS actuation significantly outperforms traditional coaching, which implies that this type of feedback can be beneficial for the motor learning of complex, repetitive movements.

Training of new skills by means of wearable technologies and augmented reality is supported by a conceptual reference framework, which enables capturing the expert's performance and provides various transfer mechanisms [5]. To capture an expert's experience with wearable sensors high-level tasks were mapped to low-level functions (including body posture, hand/arm gestures, biosignals, haptic feedback, and user location), which were decomposed to their associated sensors [9].

The Visual Inspection Tool (VIT) facilitates annotation of multimodal data as well as the processing and exploitation for learning purposes [1]. The VIT enables 1) triangulating multimodal data with video recordings; 2) segmenting the multimodal data into time-intervals and adding annotations to the time-intervals; 3) downloading the annotated dataset and using it for multimodal data analysis. The tool is part of the Multimodal Learning Analytics Pipeline.

To describe running motion, we rely on dynamic movement primitives (DMPs) [4]. DMPs describe a motion as a combination of two forces: First, a dampened spring system which counteracts any undesired disturbances over time and, second, a time-dependent force term which is fitted to the data. More specifically, the time dynamics of a DMP are described by the following equations.

$$\tau \cdot \dot{v}(t) = -\alpha \cdot \big(\beta \cdot x(t) + v(t)\big) + f(t),$$
$$\tau \cdot \dot{x}(t) = v(t), \tag{1}$$

where $x(t)$ models the location of a joint at time t, $v(t)$ the velocity of the joint at time t, $\tau \in \mathbb{R}^+$ is a hyperparameter determining the period length of the system, $\alpha \in \mathbb{R}^+$ and $\beta \in \mathbb{R}^+$ are hyperparameters determining how fast the dampened spring system counteracts disturbances, and $f(t)$ is the forcing term which models the specifics of our motion. In DMPs, this forcing term is always a linear combination $f(t) = \sum_{k=1}^{K} \Psi_k(t) \cdot w_k / \sum_{k=1}^{K} \Psi_k(t)$ of (learned) coefficients w_k with nonlinear basis functions Ψ_k. For this work, we use the following rhythmically repeating basis functions as suggested in [4].

$$\Psi_k(t) = \exp\Big(h \cdot \big[\cos(2\pi \cdot \frac{t}{\tau} - c_k) - 1\big]\Big), \tag{2}$$

where $c_k \in [0, 2\pi]$ is the phase shift of the kth basis function and h regulates the width of each basis function.

The main strengths of DMPs are that we can fit the coefficients w_k to data via simple linear regression, and that we can replay a motion at arbitrary speed

(by adjusting τ) for arbitrary long times (by executing the system in Equation 1 for longer). DMPs have been particularly popular in robotics to mimic human demonstrations [4,7,8] but, to our knowledge, have not yet been applied to provide feedback to human trainees.

3   Methodology

The aim of our project is to support learners or trainees in developing specific psychomotor skills by means of immersive learning environments. The expected solutions will combine AI approaches processing multimodal data from suitable sensors by machine learning techniques in order to analyze performance and detect faults, and finally automatically generating individual feedback.

We consider two application cases: running and collaboration with a robot. From the learning perspective, they are different. In running (the concrete aim apparently depends on the target group and the objective, which we currently set as healthy running for a wide public) the trainee repeats (relatively simple) rhythmic movements again and again, but the movements of various parts of the body should be in harmony and follow certain rules, in order not to harm the body and perform effectively. It suggests a behavioristic approach of learning, when each error should be immediately reported to the person, in order to assign the message to the corresponding movement. So when the deviation in comparison with an optimal blueprint overcomes a threshold, a suitable feedback is given. For example, when the person does not lift the feet properly, an acoustic feedback is provided.

On the other hand, collaboration between a human and a robot consists of various actions on both sides, following a common aim. Here, the human actions are typically performed by hands (e.g. in an assembly process), but in this case the person needs to evaluate the current context and decide what to do next, e.g. whether a micro-aim has been achieved and one can proceed with the next step. Two types of skills are required – the ability to cooperate with the robot (e.g. learner nudges the robot on its empty arm) and the ability to fulfill the requested task (e.g. learner puts the lid on the box). This reminds cognitivistic learning approaches, where formative feedback (we distinguish corrective and reinforcing type) plays an important role, allowing trial and error. Therefore, more complex actions need to be assessed, usually considering whether a specific micro-aim has been achieved. Nevertheless, immediate feedback cannot be excluded either, especially in case of dangerous operations.

What both application scenarios have in common is the necessity of a summative feedback, which evaluates a whole (training) unit or a (work) session. Here, different phases or sequences of actions can be analyzed, showing which parts were managed well and where is a potential for improvements.

4   Implementation

The artificial intelligence in our project essentially has the following tasks that are important for the learning process:

- Modeling templates and movement patterns to guide learners: data sets are collected that represent expert performance in selected psychomotor processes. With this data, machine learning processes are trained for the use cases.
- Detection of mistakes in the execution of movements of the learners compared to an optimal blueprint. Any deviation above a threshold triggers feedback.
- Generating helpful feedback for learners: Detected errors must be processed in such a way that learners receive starting points for improving their processes, which they can cognitively process and implement psychomotorically.
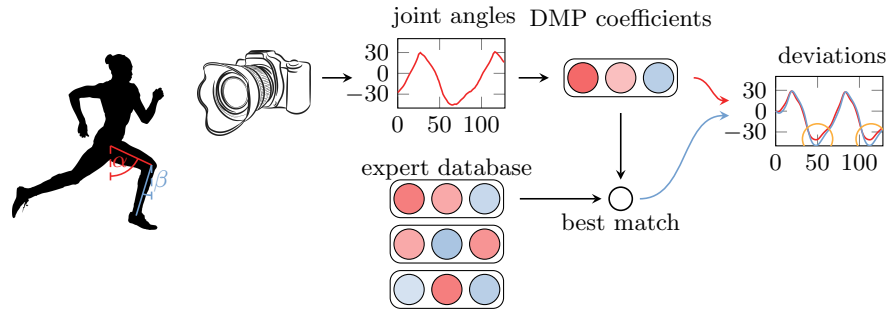

Fig. 1. An illustration of the proposed pipeline to provide feedback for the running case.

Let us illustrate these tasks for the running example. We first need to collect expert demonstrations that cover a wide range of reasonable and healthy styles of running. These demonstrations need to be recorded via motion capture devices to retrieve joint angle information that abstracts from the specific body configuration. Further, we need to abstract from running tempo and phase shift.

During a learning episode, we record a learner's current running, convert it into the same representation, and compare the latest cycle of the learner's running to the most similar expert demonstration, resulting in a measure of deviation. If the deviation exceeds a threshold, we provide auditive, tactile, or visual feedback, e.g. coloring the deviating limb in a virtual avatar of the runner [3]. Figure 1 displays a summary of our pipeline. To represent running, we use rhythmic dynamic movement primitives over the joint angle representation, which intrinsically abstracts from body shape and tempo, and is fast enough to be applied on-the-fly for each new cycle of running.

5   Evaluation

We evaluate our proposed representation for the running case on a data set of 33 runs of three runners recorded in the Carnegie Mellon Motion Capture Lab[1]. Our aim in this experiment is to abstract from particularities of a specific run and to identify the runner by comparing a run against other runs in the data base. Each run is represented by Euler angles of the right and left femur as well as the right and left tibia. We, then, derive a summary representation of each run by training a rhythmic dynamic movement primitive (DMP). We used $K = 12$ basis functions as this was sufficient to achieve a local minimum in the reconstruction error. As hyperparameters we chose $c_k = 2\pi \cdot \frac{k}{K}$, $h = \log(0.1)/(\cos(2\pi/K) - 1)$, $\alpha = 32$, and $\beta = \alpha/4$, as recommended by [4]. The period length τ for the DMP was chosen automatically to minimize the auto-regressive error (i.e. how similar the signal is to itself after shifting by τ frames). Finally, we normalized against phase-shifts by permuting the basis functions such that the distance to the first run in the data set was minimized. Figure 2 illustrates the effect of our representation on the data. The left plot displays the raw data, the right plot the reconstructed signal by our DMP representation after normalizing the period length and the phase shift. Color indicates the runner. We observe that all runs are much better aligned in the right plot and that it is easy to distinguish the running style in blue from the running style in red and orange via its peaks at frames 20 and 80.
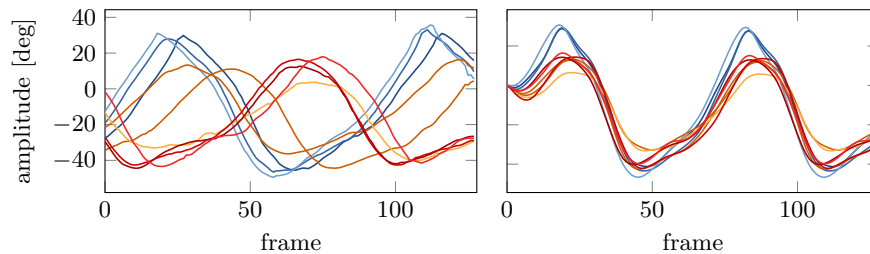
Fig. 2. The raw data (left) for the rx angle of the left femur and the DMP reconstruction (right). Runners are distinguished by hue, runs of the same runner by intensity.

Next, we evaluate the accuracy of identifying the runner in a leave-one-out crossvalidation across all runs. We use a κ-nearest neighbor classifier as implemented in sklearn[2] with the number of neighbors κ varying from one to five. To select nearest neighbors, we use the Euclidean distance on the DMP coefficients. As baselines, we also consider the Euclidean distance on the first 128 frames of

---

[1] http://mocap.cs.cmu.edu/info.php

[2] https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.
KNeighborsClassifier.html

the raw signal (because all signals where at least 128 frames long) and dynamic time warping as implemented in the edist package[3].

Table 1. Mean nearest neighbor classification accuracy in a leave-one-out crossvalidation for varying number of neighbors κ.

metric	$\kappa = 1$	$\kappa = 2$	$\kappa = 3$	$\kappa = 4$	$\kappa = 5$
Euclidean (raw)	0.67	0.73	0.55	0.58	0.55
DTW	0.88	0.88	0.85	0.88	0.76
Euclidean (DMP)	0.85	0.91	0.88	0.88	0.88

Table 1 shows the classification accuracies. Our proposed DMP representation performs best for all κ except $\kappa = 1$, where DTW is better. Additionally, our DMP representation is computationally more efficient because DTW requires a quadratic complexity in the signal length which may become infeasible for long signals. Our DMP representation is linear in the signal length.

6   Conclusion

In this paper, we sketched an approach to support learning psychomotor skills using machine learning. In particular, we proposed a pipeline that compares a learner's activity to blueprints from experts and provides feedback whenever deviations are detected that exceed a threshold. To make this approach viable, our comparison needs to abstract from irrelevant factors such as body shape, tempo, or phase shifts. We evaluated dynamic movement primitives for the case of running and did achieve a representation that was abstract enough to identify the runner from running behavior. In future work, we wish to implement our pipeline fully for running and for a robot interaction scenario and evaluate its effectiveness in supporting human learners.

References

1. Di Mitri, D., Schneider, J., Klemke, R., Specht, M., Drachsler, H.: Read between the lines: An annotation tool for multimodal data for learning. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge. pp. 51–60 (2019). https://doi.org/10.1145/3303772.3303776
2. Hassan, M., Daiber, F., Wiehr, F., Kosmalla, F., Krüger, A.: Footstriker: An ems-based foot strike assistant for running. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **1**(1), 1–18 (2017). https://doi.org/10.1145/3053332

---

[3] `https://pypi.org/project/edist/`

3. Hülsmann, F., Frank, C., Senna, I., Ernst, M.O., Schack, T., Botsch, M.: Superimposed skilled performance in a virtual mirror improves motor performance and cognitive representation of a full body motor action. Frontiers in Robotics and AI **6**, 43 (2019). https://doi.org/10.3389/frobt.2019.00043

4. Ijspeert, A.J., Nakanishi, J., Hoffmann, H., Pastor, P., Schaal, S.: Dynamical movement primitives: learning attractor models for motor behaviors. Neural computation **25**(2), 328–373 (2013). https://doi.org/10.1162/NECO_a_00393

5. Limbu, B., Fominykh, M., Klemke, R., Specht, M., Wild, F.: Supporting training of expertise with wearable technologies: The WEKIT reference framework. In: Mobile and ubiquitous learning, pp. 157–175. Springer (2018). https://doi.org/10.1007/978-981-10-6144-8_10

6. Magill, R.A., Anderson, D.I.: The roles and uses of augmented feedback in motor skill acquisition, pp. 3–21. Routledge, New York, NY,USA (2012)

7. Nakanishi, J., Morimoto, J., Endo, G., Cheng, G., Schaal, S., Kawato, M.: Learning from demonstration and adaptation of biped locomotion. Robotics and Autonomous Systems **47**(2), 79–91 (2004). https://doi.org/10.1016/j.robot.2004.03.003

8. Schaal, S., Mohajerian, P., Ijspeert, A.: Dynamics systems vs. optimal control — a unifying view. In: Cisek, P., Drew, T., Kalaska, J.F. (eds.) Computational Neuroscience: Theoretical Insights into Brain Function, Progress in Brain Research, vol. 165, pp. 425–445. Elsevier (2007). https://doi.org/10.1016/S0079-6123(06)65027-9

9. Sharma, P., Klemke, R., Wild, F.: Experience capturing with wearable technology in the WEKIT project. In: Buchem, I., Klamma, R., Wild, F. (eds.) Perspectives on Wearable Enhanced Learning (WELL), pp. 297–311. Springer (2019). https://doi.org/10.1007/978-3-319-64301-4_14