Multi-Level Analysis of Motor Actions as a Basis for Effective Coaching in Virtual Reality

Felix Hülsmann, Cornelia Frank, Thomas Schack, Stefan Kopp, Mario Botsch

Cluster of Excellence Cognitive Interaction Technology (CITEC)
Bielefeld University

Abstract. In order to effectively support motor learning in Virtual Reality, real-time analysis of motor actions performed by the athlete is essential. Most recent work in this area rather focuses on feedback strategies, and not primarily on systematic analysis of the motor action to be learnt. Aiming at a high-level understanding of the performed motor action, we introduce a two-level approach. On the one hand, we focus on a hierarchical motor performance analysis performed online in a VR environment. On the other hand, we introduce an analysis of cognitive representation as a complement for a thorough analysis of motor action.

1 Introduction

In recent years, several approaches to Virtual Reality (VR) motor learning have been presented. For instance Rector et al. [3] introduced a virtual yoga coach for visually impaired people. The authors followed a rule-based approach to define desired yoga poses. Rules were extracted from literature and interviews with experts. However, only desired optimal poses were specified, which did not take into account a hierarchical representation of motion: only simple features (joint angles calculated from joint positions) were considered. To analyze feedback strategies, Sigrist et al. [6] conducted an experiment investigating the impact of haptic, auditive, and visual feedback on motor learning using a VR rowing simulator. The simulator provides cues on performance, depending on temporal and spatial deviations of the optimal oar blade movement. Here, only the resulting motion of the oar is analyzed, not the actual motor performance of the athlete. Another coaching system was developed by Tang et al. [7] with a focus on dance training: Their analysis uses a block matching algorithm to compare the athlete's motion in terms of joint angles to pre-recorded optimal performances. Thus, a comparison of online movement and the desired optimal movement is possible without having to define any rule by hand. However, the system does not respect multi-layer representations, e.g., typical error patterns: It only determines which type of dance movement is executed and then calculates an accumulated error value for the performance. No fine grained analysis and interpretation of non-optimal parts of the performance is done.

Overall, these approaches perform analysis of motor actions with respect to only a subset of particular aspects of motor actions: Often only simple features of the motion (e.g., joint angles) are considered and other important features like speed etc. cannot be integrated into the model. As a detailed analysis is helpful for further steps, e.g., giving helpful feedback, this is one of the gaps we aim to diminish in this work: We develop a suitable representation and analysis of (a) online motor performance, but also with respect to (b) an offline aquisition of the athlete's cognitive representation of the of the action as presented in [4]. In our scenario, an athlete, who is tracked by an OptiTrack Prime13W system, is placed inside a two-sided CAVE (stereo projection area on floor and front). Here, she has to perform a motor action (e.g., squats) in front of a virtual mirror. The mirror reflects the athlete's motion mapped on a virtual avatar. As exemplary actions we use squat performances. The approaches discussed in the following can also be extended to further types of motor actions.

2 Methodology and Realization

Analysis of Motor Performance: Our real-time performance analysis is able to combine an extendable set of features into a hierarchical representation. It is based on rules that describe the desired motion. This kind of analysis is highly efficient and allows a direct interpretation of the results in terms of performance flaws. Motion tracked by the OptiTrack system is transferred to our framework, where we represent it as a temporal sequence of a set of features (this could be, e.g., joint angles, positions, symmetry etc.). These features can be calculated using the raw motion capture data. The sequence is split into single repetitions of motor actions (e.g., squats), connected by arbitrary transition movements. Each repetition is a combination of simpler sub-actions, denoted *Movement Primitives* (MP). This could be, e.g., the going-down and the going-up stages during a squat. Also, additional MPs, like an is-down stage, can be defined if required for the performance analysis. For each type of action and MP, a list of relevant features is manually specified. Then, key-postures for the MPs are defined using manual analysis of recorded video and/or tracking data. For the squat, this can be performed — among others — via observing symmetric key angles of knees. To detect a single action, the system has to detect a posture similar enough to one of these key-postures. The analyzer is a state machine: As soon as a posture similar enough to a key-posture describing a valid state transition is detected, the analyzer switches its state and waits for the next key-posture. If the observed motion frame does not belong to the current state nor to an allowed state transition, the system returns to the starting state: The performed movement does not depict any known action/MP or the performance has been aborted. The state of the analyzer reflects the current action and MP (see Figure 1). This representation has the advantage to allow focusing on style patterns on an extensible feature set which are especially relevant for single parts of the action.

To detect erroneous performances, a list of Prototypical Style Patterns (PSP) is defined for every action, describing movement styles considered as undesired. A PSP is defined using at least one rule, which describes, e.g., the violation

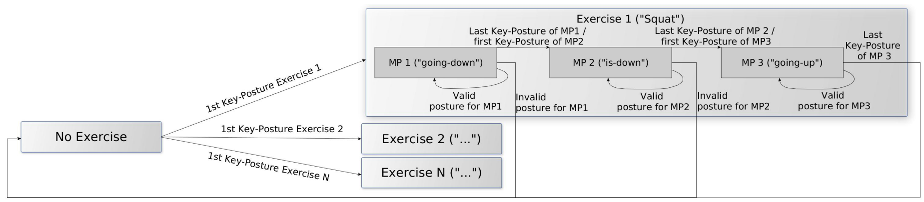

Fig. 1. State machine used to determine current action and Movement Primitive.

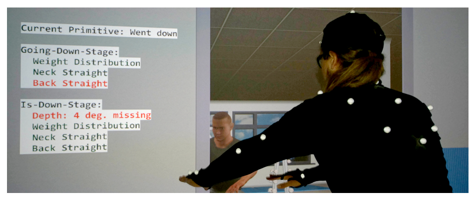

Fig. 2. The athlete is in the Is-Down-Stage. Here, he did not yet reach optimal depth (marked in red). During going down, PSP "Straight Back" was violated.

of specified feature constraints. Each rule returns a quantitative error value for the incoming motion. PSPs and rules are developed based on literature and information received from experts. Here, also observations from recorded data of correctly or incorrectly performed actions were taken into account. One example for a PSP in the context of squat is, e.g., "Incorrect Weight Distribution": One indicator for this pattern is that the knees are in front of the toes. Furthermore, this pattern can be detected via observing the angle of the shin or the ankle. Also using center of mass of upper and lower body may be conceivable. Finally, the highest level of this hierarchical motion representation is the action. Lower levels are the MPs and PSPs, and the lowest level are the features itself, e.g., joint angles, speed or positions at a given time. For a typical setup, our analysis needs approximately 1ms per motion frame. Figure 2 shows the system in action.

Analysis of Cognitive Representation: Analysis of cognitive representation is performed using structural-dimensional analysis (SDA-M) including a splitting task of a given set of basic action concepts (BACs), a hierarchical cluster analysis, and an analysis of invariance [4]. As such, this method provides psychometric data regarding the relations and groupings of BACs of a complex action (here: the squat) in long-term memory. For instance, the BACs "straight posture" and "feet shoulder-width" functionally relate to the preparatory phase of the squat, while the BACs "move bottom backwards" and "bend legs" functionally relate to the main phase. By determining the degree of BAC order formation in memory (i.e., relation of BACs in terms of functional movement phases), we can derive information both on the skill level of the athlete (cf. [5]) as well as on the athlete's progress of learning (cf. [1]). In addition, errors can be inferred by comparing an athlete's representation to a reference, e.g., an expert representation.

3 Discussion and Perspective

This paper presented a multi-level analysis of motor actions as a basis for an intelligent coaching space in the field of motor learning. We started with our online analysis of motor performance followed by an offline analysis of the athlete's cognitive representation. The online analysis of motor performance is highly efficient and easily interpretable. While Rector et al. [3] already successfully applied a rule-based approach to detect desired joint angles, our approach goes beyond: We respect the hierarchical properties of motion and allow considering a nearly arbitrary and extendable set of features. Thus we do not only wait for desired angles and return the current deviation, but we provide an online interpretation of performed erroneous motion in terms of style pattern. The analysis of the cognitive representation serves as a second source of information in order to learn about the individuals perceptual cognitive prerequisites for subsequent feedback and coaching strategies. As such, the cognitive analysis adds an additional high-level layer to the hierarchy, built-on by the online analysis of motor performance. In addition, coupling PSPs from motor performance analysis with groupings and relations of BACs from the analysis of the athlete's cognitive representation allows to link the level of cognitive representation to the level of motor performance [2]. Our next steps aim at investigating the potential of this combination as a basis for VR coaching.

Acknowledgments: This work was supported by the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG).

References

1. Frank, C., Land, W.M., Schack, T.: Mental representation and learning: the influence of practice on the development of mental representation structure in complex action. Psychology of Sport and Exercise 14(3), 353–361 (2013)
2. Land, W.M., Volchenkov, D., Bläsing, B.E., Schack, T.: From action representation to action execution: exploring the links between cognitive and biomechanical levels of motor control. Frontiers in Computational Neuroscience 7, 1–14 (2013)
3. Rector, K., Bennett, C.L., Kientz, J.A.: Eyes-free yoga: an exergame using depth cameras for blind & low vision exercise. In: Proc. of the 15th Int. ACM SIGACCESS Conf. on Computers and Accessibility. p. 12. ACM (2013)
4. Schack, T.: Measuring mental representations. Measurement in sport and exercise psychology, Human Kinetics (2012)
5. Schack, T., Mechsner, F.: Representation of motor skills in human long-term memory. Neuroscience letters 391(3), 77–81 (2006)
6. Sigrist, R., Rauter, G., Marchal-Crespo, L., Riener, R., Wolf, P.: Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning. Experimental brain research 233(3), 909–925 (2015)
7. Tang, J.K., Chan, J.C., Leung, H.: Interactive dancing game with real-time recognition of continuous dance moves from 3d human motion capture. In: Proc. of Int. Conf. on Ubiquitous Information Management and Communication. p. 50 (2011)