Augmented Intelligence in Tutoring Systems: A Case Study in Real-Time Pose Tracking to Enhance the Self-learning of Fitness Exercises

Nghia Duong-Trung(✉), Hitesh Kotte, and Miloš Kravčík

Educational Technology Lab, German Research Center for Artificial Intelligence (DFKI), Alt-Moabit 91C, 10559 Berlin, Germany
{nghia_trung.duong,hitesh.kotte,milos.kravcik}@dfki.de

Abstract. In technology enhanced learning is development of psychomotor skills an area with a lot of potential, which is enabled by rapid improvements of sensors and wearable devices, combined with artificial intelligence. Here we focus on fitness exercises and present a novel approach based on computer vision techniques to track the practitioner's pose and provide automatically real-time feedback for improvement, based on the input from an expert trainer. Taking into account the gathered data and ground-truth poses, the proposed pipeline can learn actively from a professional trainer demonstrating an exercise in front of a camera or passively from a recorded video. In our experiment, we used professional fitness exercise videos as the ground truth and measured the performance of five inexperienced participants. The results show positive responses from participants, indicating the feasibility of the proposed approach as well as suggestions for its further improvement.

Keywords: Real-time Pose Tracking · Sustainable Learning · Psycho-motor Skills · Fitness Exercises

1 Introduction

Augmented intelligence (AuI) is a rapidly evolving technology that combines human intelligence with machine intelligence to enhance performance and productivity in various fields, including sports and fitness [2]. In tutoring systems for sport and fitness, augmented intelligence refers to a collaboration between humans and artificial intelligence (AI), where AI assists humans in decision-making and problem-solving [11]. The systems can provide individuals with a highly personalized coaching experience tailored to their needs and goals [4]. By analyzing individuals' performance data, these systems can adjust coaching recommendations to help them improve in areas where they struggle. The use of AuI in tutoring systems [5] holds a lot of potential to improve the effectiveness, safety, and enjoyment of sports and fitness activities. Nevertheless, various

O. Viberg et al. (Eds.): EC-TEL 2023, LNCS 14200, pp. 705–710, 2023.
https://doi.org/10.1007/978-3-031-42682-7_65

sports are primarily done in a do-it-yourself process, which can be unsafe [3]. People watch trainer demonstrations or instructional videos, then perform the exercise while imitating what they have observed. A good plan and guidance are essential to achieve their fitness goals.

In the rest of the paper we first briefly introduce the related work. Then we describe several challenges in fitness training. The core of the paper is the presentation of our experiment and its results, including a short qualitative evaluation. Finally we summarize our findings.

2 Related Work

Early research illustrated the potential of AI techniques in sports, particularly in weight training, to measure the virtual displacement and force determinants while performing various exercises [7]. Li *et al.* introduced spatio-temporal skeleton encoding to recognize and evaluate the gym action in an AI fitness system [6]. Intensive data collection and model training was conducted, but unfortunately, they are unavailable to reproduce. Ueta [10] proposed expensive augmented reality devices and an impractical virtual-real interaction to correct the fitness movement. Conceptual and theoretical models for learning psychomotor skills have emerged from preliminary studies [1] where further work should focus on automatic, individual feedback generation. Multiple methods were investigated to efficiently represent and predict human motion, e.g., a machine learning-supported approach to provide feedback in two example scenarios: running and interacting with a robot. For the first scenario, an assessment was provided of how movements can be compared to show discrepancies between student and expert movement [9]. Some psychomotor tasks require students to perform a specific sequence of poses and movements. A natural teaching scheme for such studies is to match a person's performance with a teacher's demonstration. However, this requires strategies to match the teacher's demonstration of each movement to the person's attempts and to recognize differences between demonstration and attempt. To automatically identify student attempts for poses with only one correct teacher demonstration, the investigated methods included relevance learning, prototype networks, and attentional mechanisms to achieve a robust, few-shot approach that can be generalized to all students. An experiment [8] showed that prototypical networks work best with an attention mechanism.

3 Self-practicing and Feedback in Fitness

Self-practicing of fitness exercises can be challenging for several reasons. First, it may not be easy to maintain proper posture and technique without the guidance of a coach or trainer. Poor posture can lead to injuries and reduced effectiveness of the exercise. Second, staying motivated and tracking progress without external accountability and feedback is also difficult. Third, it is necessary to personalize a workout program to meet individual needs and goals. In the process, a coach or an instructor performs a particular exercise before a trainee.

Then, the participant tries to repeat the movement by imitating what has been observed. Next, by monitoring the trainee's performance, the trainer can provide feedback or perform the exercise again. While having a trainer present during fitness practice can be helpful in terms of guidance, several potential issues can arise. Trainees may become overly reliant on the trainer and need their presence to motivate themselves or maintain a consistent fitness routine. Additionally, hiring a trainer can be expensive, and if the instructor is always present, the cost of their services can add up quickly. If the trainer is always leading the workout, the client may not have the opportunity to try different exercises or workout routines, which can lead to boredom and decreased motivation. It is also essential for clients to perform their workouts and make adjustments based on their individual needs and goals.

4 Experiments

4.1 Fitness Exercises by McFIT GmbH

For our purposes we used the McFIT's YouTube channel for home fitness. The videos are led by experienced fitness trainers who guide viewers through each exercise, providing clear instructions and modifications for different fitness levels. We selected videos corresponding to 15 fitness exercises and analysed them (see Table 1).

4.2 Experiment Setup

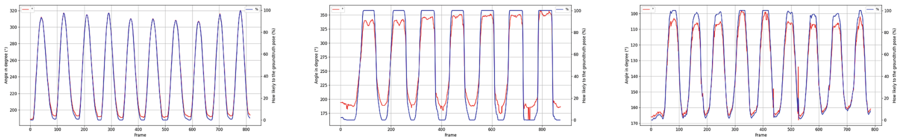
Given three key points $u(x_u, y_u), v(x_v, y_v), p(x_p, y_p)$, the joint angle $\theta(u, v, p)$ (in degree) between two rays formed by three mentioned points is calculated as follows: $\theta(u, v, p) = \frac{180(\phi(y_p - y_v, x_p - x_v) - \phi(y_u - y_v, x_u - x_v))}{\pi}$. where the angle $\phi(y, x)$ (in radian) between the ray from the origin to the point (x, y) and the positive x-axis in the Cartesian plane is calculated as follows: $\phi(y, x) = arctan(\frac{y}{x})$, if $x > 0$; $\phi(y, x) = \frac{\pi}{2} - \arctan(\frac{x}{y})$, if $y > 0$; $\phi(y, x) = \arctan(\frac{y}{x}) \pm \pi$, if $x < 0$; $\phi(y, x) = -\frac{\pi}{2} - \arctan(\frac{x}{y})$, if $y < 0$. The setup allowed for effective and efficient data collection and preparation for subsequent analysis using machine learning algorithms. A McFIT studio was chosen as the location to conduct experiments. Five voluntary participants, between the ages of 25 and 30, were asked to mimic the exercise shown to them while receiving real-time feedback and pose tracking through the AI system. The feedback provided by the AI system included information on the correctness of the exercise posture, the positioning of the body and limbs, and the movement range. The results showed that using real-time feedback and pose tracking through the AI system significantly improved exercise mimicry. Participants were able to correct their exercise posture, body and limb positioning, and movement range in response to the feedback provided by the AI system. We decompose our real-time fitness tutoring system into three phases: Keypoints Detection, Pose Tracking, and Output and Feedback, see Algorithm 1.

Algorithm 1. Pipeline for Self-Learning of Fitness Exercises.

Keypoints Detection Phase:
  Get input from recorded video or real-time streaming.
  Detect $\{x_b, y_b, w_b, h_b, \text{cf}, k_x^0, k_y^0, k_{cf}^0, k_x^1, k_y^1, k_{cf}^1, \ldots, k_x^{16}, k_y^{16}, k_{cf}^{16}\}$ in each frame.

Pose Tracking Phase:
  Get predefined keypoints combination and range of joint angles.
  Calculate θ and ϕ.
  Convert θ to 0-100% range using one-dimensional linear interpolation.

Output and Feedback Phase:
  Display the trainee's performance based on a predefined difficulty level.



(a) Barbell Shrugs (b) Lateral Pulls (c) Lunges

Fig. 1. Performance tracking of different exercises.

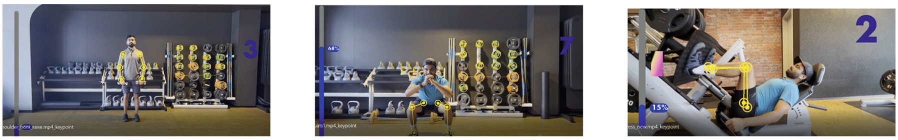
Table 1. Keypoint combination for different sports exercises.

Exercises	Keypoints			
	5 | 7 | 9	11 | 13 | 15	6 | 8 | 10	12 | 14 | 16
Barbell Shrugs	[192°,320°]	-	[192°,320°]	-
Bicep Curls	[010°,150°]		[010°,150°]	
Pushups	[210°,280°]		[210°,280°]	
Shoulder Lateral Raise	[171°,194°]		[171°,194°]	
Shoulder Front Raise	[145°,290°]		[145°,290°]	
Shoulder Press	[026°,180°]		[026°,180°]	
Chest Press Rope	[186°,199°]		[186°,199°]	
Upper Chest Press	[177°,285°]		[177°,285°]	
Lateral Pulls	[191°,327°]		[191°,327°]	
Pullups	[142°,321°]		[142°,321°]	
Lunges	-	[142°,321°]	-	[153°,066°]
Squats		[220°,280°]		[220°,280°]
Horizontal Leg Press		[157°,086°]		[157°,086°]
Vertical Leg Press		[252°,170°]		[252°,170°]
Chest Press		[151°,108°]		[151°,108°]

4.3 Experiment Results

First, we propose the ideal range of joint angles in 15 different exercises; see Table 1. Those values are the model's hyperparameters which can be configured by fitness experts, individually adjusting it for each participant. Another vital piece of information is performance tracking, where a trainee can review his execution or later consult feedback from the trainer. Figure 1 shows the measured time series of selected exercises among participants: The visualization of the joint angles, in degree (°), for each repetition and a comparison with the ground-truth pose, in percentage (%). The early performance is characterized by variable displacement fluctuation compared to relatively smooth plots after many executions. Our screenshots of real-time monitoring show two graphical indicators in different contexts (Fig. 2): 1. the percentage bar (on the left) indicating the matching of the executing joint angles with the ground-truth pose, 2. the number (on the right) counting correct movements. Participants observe their performance in real-time on the experiment laptop in the fitness room. We provide the experimental codes at https://github.com/duongtrung/ECTEL2023.

After conducting a month-long series of experiments, we administered a questionnaire to our participants to assess the effectiveness of our proposed system. The questionnaire consisted of three key inquiries: 1. In your opinion, does this pose tracking system, which offers real-time visual feedback, have the potential to be appealing and beneficial for trainees? 2. What do you perceive as the primary advantages of this system? 3. What improvements do you believe are necessary to enhance its performance? Remarkably, all participants responded affirmatively to the first question. Regarding the second question, numerous noteworthy points were raised. Participants highlighted the significance of posturing correction, visual assistance, and instant feedback during workouts. They also emphasized that the system's ability to reduce dependency on personal trainers was an advantage worth noting. Furthermore, participants made several suggestions for improvement, including a more user-friendly interface, voice command functionality, visual workout plans, and performance logging features.



(a) Shoulder Front Raise (b) Squats (c) Leg Press

Fig. 2. Several screenshots of our participants. Image copyright by the authors.

5 Conclusion

The proposed fitness tutoring system suggests an intuitive approach based on computer vision techniques to track the practitioner's pose during fitness exercises and provide instant feedback for improvement, even without a trainer or expert. The three-phase pipeline can actively or passively learn from a professional trainer demonstrating an exercise or a recorded video, respectively, and is powered by the state-of-the-art pose estimation model and human topology-based tracking technique. The demonstration using professional fitness exercise videos as the ground truth showed a positive reaction of participants, indicating the feasibility and potential of the proposed method for similar sport exercises.

Acknowledgment. The authors would like to thank the German Federal Ministry of Education and Research (BMBF) for their kind support within the project *Multimodal Immersive Learning with Artificial Intelligence for Psychomotor Skills* (MILKI-PSY) under the project ID 16DHB4014.

References

1. Di Mitri, D., Schneider, J., Limbu, B., Mat Sanusi, K.A., Klemke, R.: Multimodal learning experience for deliberate practice. In: Giannakos, M., Spikol, D., Di Mitri, D., Sharma, K., Ochoa, X., Hammad, R. (eds) The Multimodal Learning Analytics Handbook. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-08076-0_8
2. Fajrianti, E.D., et al.: Application of augmented intelligence technology with human body tracking for human anatomy education. IJIET: Int. J. Inf. Educ. Technol. **12**(6), 476–484 (2022)
3. Farrokhi, A., Farahbakhsh, R., Rezazadeh, J., Minerva, R.: Application of internet of things and artificial intelligence for smart fitness: a survey. Comput. Netw. **189**, 107859 (2021)
4. Graßmann, C., Schermuly, C.C.: Coaching with artificial intelligence: concepts and capabilities. Hum. Resour. Dev. Rev. **20**(1), 106–126 (2021)
5. Kim, J., Davis, T., Hong, L.: Augmented intelligence: enhancing human decision making. In: Albert, M.V., Lin, L., Spector, M.J., Dunn, L.S. (eds) Bridging Human Intelligence and Artificial Intelligence. Educational Communications and Technology: Issues and Innovations. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-84729-6_10
6. Li, J., Cui, H., Guo, T., Hu, Q., Shen, Y.: Efficient fitness action analysis based on spatio-temporal feature encoding. In: 2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), pp. 1–6. IEEE (2020)
7. Novatchkov, H., Baca, A.: Artificial intelligence in sports on the example of weight training. J. Sports Sci. Med. **12**(1), 27 (2013)
8. Paaßen, B., Baumgartner, T., Geisen, M., Riedl, N., Kravčík, M.: Few-shot keypose detection for learning of psychomotor skills. In: Proceedings of the 2nd International Workshop on Multimodal Immersive Learning Systems (MILeS 2022) (2022)
9. Paaßen, B., Kravcık, M.: Teaching psychomotor skills using machine learning for error detection. In: Proceedings of the 1st International Workshop on Multimodal Immersive Learning Systems (MILeS 2021), pp. 8–14 (2021)
10. Ueta, M.: Improving mirror fitness through augmented reality technology. In: 2022 3rd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), pp. 186–194. IEEE (2022)
11. Venkatachalam, P., Ray, S.: How do context-aware artificial intelligence algorithms used in fitness recommender systems? A literature review and research agenda. Int. J. Inf. Manag. Data Insights **2**(2), 100139 (2022)