# ARFit: Pose-based Exercise Feedback with Mobile AR

Sara Mandic
smandic@ucsb.edu
University of California
Santa Barbara, CA, USA

Rhys Tracy
University of California
Santa Barbara, CA, USA
rhystracy@ucsb.edu

Misha Sra
University of California
Santa Barbara, CA, USA
sra@ucsb.edu

## ABSTRACT

Fitness encompasses a diverse array of activities, including gym sessions, home workouts, and various other forms of physical exercise. The importance of proper muscle movement remains consistent across all these settings. Ensuring people do the physical movements correctly, avoid injury, gain insights and maintain motivation is traditionally accomplished with the help of an expert instructor. In the absence of an expert, the most common at-home training methods are books, videos, or apps but they provide limited feedback. In this work, we introduce ARFit, an augmented reality application that uses pose-tracking technology to capture user exercise movements and provide feedback to ensure accurate posture.

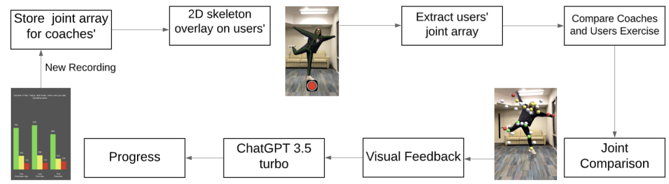


**Figure 1: System pipeline: starting with camera data that is processed to recover body pose, the system calculates errors relative to an expert's pose and provides detailed feedback on joints and movements using natural language through the integration of a LLM (ChatGPT 3.5).**

## CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Mixed / augmented reality**.

## KEYWORDS

augmented reality, fitness, exercise, feedback

**ACM Reference Format:**
Sara Mandic, Rhys Tracy, and Misha Sra. 2023. ARFit: Pose-based Exercise Feedback with Mobile AR. In *The 2023 ACM Symposium on Spatial User Interaction (SUI '23), October 13–15, 2023, Sydney, NSW, Australia.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3607822.3618008

## 1 INTRODUCTION

Working out at home is increasingly popular among fitness minded users due to convenience and lower cost. These workouts require correct execution of movements to be effective and to avoid injury [6, 11, 12, 17], for which feedback is critical. With newer AI technologies, several fitness solutions, including mobile apps, have emerged in both research (e.g., [7, 20]) and industry [1, 2]. These apps and devices assist users in their fitness activities, but the feedback they provide tends to focus on sensing and measuring performance such as hours spent exercising or heart rate during exercise. They do not offer feedback on a user's movements. Individualized feedback on body movements has been shown to enhance motivation, engagement, and satisfaction [18]. While the Mirror [15] can send recorded video to instructors for feedback, it is expensive and needs a subscription for class access, requires booking and payment for instructor-led training sessions, and is tied to a specific location.

In this work, we explore the design of ARFit, a mobile augmented reality (AR) app that offers an inexpensive method to provide automated natural language feedback on a user's body movements, anytime and anywhere. YouMove [3] is a related work comprised of a Kinect-based authoring system to record actions which are manually edited. A corresponding training system uses a half-silver mirror with graphic overlays for guidance and feedback. Other research has explored audio AR for providing feedback on skating movements [9], projection AR for strength training [13], and mirror AR [14] for visualizing the teacher's motion. In contrast, our system automatically compares trainee/expert actions, provides natural language feedback, and does not require additional hardware to help democratize access to AI-based fitness support for large audiences worldwide. While OneBody [10] offers first person perspective feedback in VR, feedback in our system is akin to looking in the mirror, commonly used in all gyms.

## 2 SYSTEM DESIGN AND IMPLEMENTATION

The system design includes three main components: (1) user and expert joint extraction, (2) joint comparison, and (3) feedback. We built our pipeline (Figure 1) for exercise feedback in AR on an iOS mobile device using the Unity game engine.

*Joint Extraction.* Prior to the execution of the application, the expert data is stored. We utilize PoseCamera, a Python SDK, to extract the 2D joints from the camera data. In order to ensure compatibility, we match the number of joints to ARFoundation, a Unity framework for AR development [19]. The processed data is serialized as an array of joint arrays into a JSON text file. For seamless integration, we store this file within Unity Resources, allowing it to be loaded and de-serialized in the AR app. To capture the user's exercise session, we use ReplayKit [5], allowing the user to record, save, and view their recordings on their mobile device. With ARFoundation and ARKit [4], we overlay a skeleton on the


*SUI '23, October 13–15, 2023, Sydney, NSW, Australia*

ACM ISBN 979-8-4007-0281-5/23/10.
https://doi.org/10.1145/3607822.3618008

user's video feed, to provide a visual representation of their body pose/joints and exercise movements. For each frame of the recorded exercise, we extract the pose joint positions and store them as an array of joint arrays for further analysis.

*Joint Comparison.* To enable a comparison between the expert and user skeletons, we normalize their alignment and scale. By aligning the root node of the expert's skeleton with that of the user and adjusting the translation and scaling of expert joints, we ensure consistent hip-width distance and root-to-head node distance for both, yielding equal horizontal and vertical scaling of the user and expert skeletons. This alignment facilitates overlaying the two skeletons for comparison while granting users the freedom in recording their exercise videos from near or far. After normalization, each recorded frame is assessed by calculating the Euclidean distance between the corresponding joints of the user and the expert. Color-coded labels based on an empirically defined error threshold to indicate the correctness of each joint.

*Feedback.* Once the exercise is complete, the user can receive four diverse forms of feedback, including visual skeleton overlays, human-like feedback generated by ChatGPT 3.5 turbo [16], and progress updates. This multimodal feedback is designed to enhance the user's comprehension of the exercise and provide valuable insights for improvement. First, they can choose to replay the exercise video or watch it frame-by-frame, comparing their motions to that of the expert. The left image in Figure 2 depicts a single frame where the joints are color-coded based on their correctness. Second, they are provided an exercise summary for each video that consists of a skeleton with joints color-coded based on the most prevalent joint coloring over the entire exercise. Third, the user is provided with natural language feedback. We manually extracted key parts of the exercise motion (a future version could use an AI model [8]) and sent the joint coordinate arrays of the user and expert to ChatGPT to receive personalized feedback in natural language mimicking expert human feedback. An example response for an incorrect jumping jack from ChatGPT says: "Comparing the leg positions, the expert's skeleton shows a wider leg stance during the jumping jack, with the legs positioned further apart. In your skeleton, the legs appear to be closer together." Lastly, the user is also able to see their aggregate progress on their exercise over a three-day period. The overall feedback consists of the count of green, yellow, and red joints for a specific exercise over a span of three days. Additionally, text feedback informs users of their progress and allows them to track improvements or setbacks compared to previous exercises (Figure 2). The joint-based progress feedback enables the user to identify the most incorrect joint in their upper and lower body and observe the progress of that joint over a three-day period. This also includes detailed feedback on a specific joint. This allows users to track their development over time providing a sense of accomplishment and motivation as they see their advancements and understand their progress towards achieving their goals.

## 3 LIMITATIONS AND FUTURE WORK

Our work establishes a foundation for leveraging mobile AR to provide exercise feedback. While we use ChatGPT to enhance the natural language feedback, a limitation of LLMs is their tendency to hallucinate, requiring a secondary validation method of the generated feedback. Our future version will provide realtime feedback as the user is exercising and will integrate ChatGPT4 for improved functionality on the specifics of feedback. Another future plan is to integrate with existing health devices, such as smartwatches, to broaden data collection and further personalize the user experience. By using deep learning to automatically identify and segment exercises, we plan to eliminate the need for manual selection of exercises, making the system mimic a workout session with a fitness trainer. Lastly, we strive to improve pose tracking, addressing challenges related to different viewpoints between user and expert data, background interference and limb overlap with multiple users in the camera's field of view that currently result in incorrect reconstruction. Beyond technical updates in future versions, we plan to conduct a user study to understand if our app is helpful and get feedback on the design of feedback modalities.

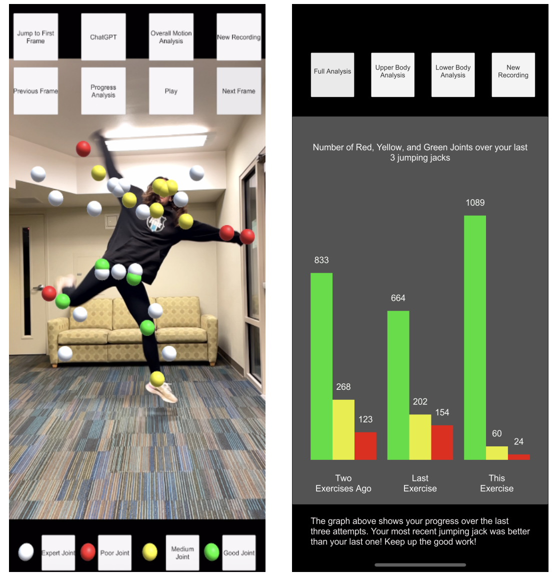


**Figure 2: Left: User performing a jumping jack sees their incorrect joints in yellow and red. Right: Progress over a 3-day period for jumping jack performance.**

## 4 CONCLUSION

We introduced an AR app that uses human pose tracking to provide feedback on exercise movements that are often missing when using videos and apps. The workout summary, frame-by-frame comparison, natural language feedback, and progress feature allows the users to evaluate their motions, track progress, and perform the exercise correctly. This can reduce the risk of injury, save money on personal trainers, increase the user's confidence, and allow for continuous improvement.

## ACKNOWLEDGMENTS

We thank Zechen Ma for contributions to the initial version of the app and Arthur Caetano for feedback on the project and writing.

## REFERENCES


[1] [n. d.]. Alfa AI. https://www.alfa-ai.com/. Accessed: 07/03/2022.

[2] [n. d.]. Gymfitty. https://www.gymfitty.com/. Accessed: 06/29/2022.

[3] Fraser Anderson, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2013. YouMove: enhancing movement training with an augmented reality mirror. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*. 311–320.

[4] Apple. 2023. Apple AR. https://developer.apple.com/augmented-reality. (Accessed on 07/16/2023).

[5] Apple. 2023. Apple ReplayKit. https://developer.apple.com/documentations/replaykit. (Accessed on 07/16/2023).

[6] Thomas R Baechle and Roger W Earle. 2014. *Fitness weight training*. Human kinetics.

[7] Don Samitha Elvitigala, Denys JC Matthies, Löic David, Chamod Weerasinghe, and Suranga Nanayakkara. 2019. GymSoles: Improving Squats and Dead-Lifts by Visualizing the User's Center of Pressure. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[8] Mihai Fieraru, Mihai Zanfir, Silviu Cristian Pirlea, Vlad Olaru, and Cristian Sminchisescu. 2021. Aifit: Automatic 3d human-interpretable feedback models for fitness training. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 9919–9928.

[9] Andrew Godbout and Jeffrey Edwin Boyd. 2012. Rhythmic Sonic Feedback for Speed Skating by Real-Time Movement Synchronization. *International Journal of Computer Science in Sport (International Association of Computer Science in Sport)* 11, 3 (2012).

[10] Thuong N Hoang, Martin Reinoso, Frank Vetere, and Egemen Tanin. 2016. Onebody: remote posture guidance system using first person view in virtual environment. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. 1–10.

[11] Nicola J Hodges and A Mark Williams. 2012. Skill acquisition in sport: Research, theory and practice. (2012).

[12] Honorata Jakubowska. 2017. *Skill Transmission, Sport and Tacit Knowledge: A Sociological Perspective*. Routledge.

[13] Raine Kajastila, Leo Holsti, and Perttu Hämäläinen. 2016. The augmented climbing wall: High-exertion proximity interaction on a wall-sized interactive surface. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 758–769.

[14] Itaru Kuramoto, Yukari Nishimura, Keiko Yamamoto, Yu Shibuya, and Yoshihiro Tsujino. 2013. Visualizing velocity and acceleration on augmented practice mirror self-learning support system of physical motion. In *2013 Second IIAI International Conference on Advanced Applied Informatics*. IEEE, 365–368.

[15] Mirror. 2023. Mirror. https://mirror.co/. (Accessed on 07/16/2023).

[16] OpenAI. 2023. ChatGPT. https://openai.com/blog/introducing-chatgpt-and-whisper-apis. (Accessed on 07/16/2023).

[17] Michael H Stone, Kyle C Pierce, William A Sands, and Meg E Stone. 2006. Weightlifting: A brief overview. *Strength and Conditioning Journal* 28, 1 (2006), 50.

[18] Laia Turmo Vidal, Hui Zhu, and Abraham Riego-Delgado. 2020. BodyLights: Open-Ended Augmented Feedback to Support Training Towards a Correct Exercise Execution. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[19] Unity. 2023. Unity ARFoundations. https://unity.com/unity/features/arfoundation. (Accessed on 07/16/2023).

[20] Darragh Whelan, Martin O'Reilly, Tomás Ward, Eamonn Delahunt, and Brian Caulfield. 2016. Evaluating performance of the lunge exercise with multiple and individual inertial measurement units. In *Pervasive Health 2016: 10th EAI International Conference on Pervasive Computing Technologies for Healthcare, Cancun, Mexico, 16-19 May 2016*. ACM.