



SkillAR:omnipresent in-situ feedback for motor skill training using AR

Florian Diller¹ · Nico Henkel¹ · Gerik Scheuermann² · Alexander Wiebel¹

Received: 1 August 2024 / Accepted: 17 January 2025
© The Author(s) 2025

Abstract

We present a novel technique called SkillAR to display augmented reality feedback for motor skill learning on a head-mounted display (HMD). SkillAR allows the user to consider movement corrections independent of the head position. Therefore, the user can receive motor feedback comfortably without risking an incorrect exercise performance. Head-mounted displays represent versatile technologies for providing motor feedback regarding skill training. In contrast to room-mounted displays, HMDs are easily portable and wearable. That allows for in-situ feedback in many situations where this would otherwise not be possible. However, the spatial positioning of the 3D feedback is not trivial. On the one hand, the user needs to understand the relation between the body and the suggested corrections in the space. On the other hand, certain exercises demand a specific head positioning to minimize errors and injuries. Depending on the exercise and the type of feedback, these two aspects can be highly conflicting. The paper at hand presents a solution for augmented reality headsets, that provides continuous and omnipresent motor feedback comfortably while facilitating a correct exercise performance. In an in-subject user study with 39 participants, we could not detect a significant disadvantage to SkillAR compared to conventional feedback methods found in the literature regarding identification and execution time as well as identification accuracy, in addition to the apparent major advantages for user health and usability. Furthermore, we found that the users could identify the feedback on the HMD more accurately.

Keywords Augmented reality · Motor feedback · Head mounted display · Skill training

1 Introduction and background

According to the World Confederation for Physical Therapy, 1,962,741 physical therapists practiced worldwide in 2023. Europe in particular had the highest rate of physiotherapists with 13.5 per 10,000 people as documented by the World Physiotherapy (2023). This may be attributed to the fact that 20% of people globally suffer from chronic pain (see Treede et al. (2015)). With computer science advancing, technology

can increasingly support therapists and clients during therapy. For instance, superimposing virtual content onto the user's perception, as common in mixed reality, can assist during physical therapy or exercise to facilitate the correct execution of motion (as seen in for example the works of Brepohl and Leite (2023), Campo-Prieto et al. (2021), and Diller et al. (2024)).

Throughout the literature, we see the use of head-mounted displays (HMDs) to provide feedback for physical therapy and exercise. On the one hand, these technologies present many challenges to overcome. For example, the first-person perspective of HMDs limits the user's view of the user's body parts. On the other hand, HMDs offer solutions to problems that occur, when providing visual feedback during physical therapy and exercise. For instance, providing visual feedback in a manner that allows users to maintain a neutral head position is challenging. However, if the feedback can only be perceived on a wall-mounted display, users are forced to change their head position, and this — in the worst case — inhibits the execution of the correct movements. This can bear serious consequences, such as

✉ Florian Diller
diller@hs-worms.de

✉ Gerik Scheuermann
scheuermann@informatik.uni-leipzig.de

Alexander Wiebel
wiebel@hs-worms.de

¹ UX-Vis, Hochschule Worms, Erenburgerstraße 19,
67549 Worms, Rheinland-Pfalz, Germany

² BSV group, Universität Leipzig, Augustusplatz 10,
04109 Leipzig, Sachsen, Germany

becoming accustomed to incorrect exercise movement, possibly even leading to injury. For instance, when executing squats correctly, the spine is meant to stay straight. Feedback provided on a room-mounted display (RMD) can force the user to bend or twist the neck (and with it the spine) to see the feedback. Therefore, the exercise becomes uncomfortable, the execution can become incorrect, and thus the risk of injury can rise — especially when performing the exercise with free weights (e.g. barbell). Alternatively, if the exercise is executed correctly and securely, the feedback via RMD might not be in the user's field of view.

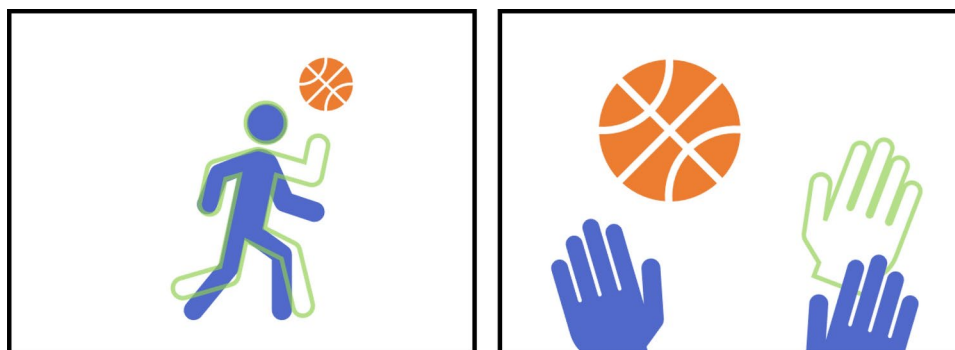
When the motion feedback is provided on an HMD, the user can execute the movements correctly, comfortably, and non-injuriously. Since the display is mounted on the head, the visual cues can adapt to the head position and the view direction and thus be omnipresent. In this paper, we present the novel approach *SkillAR* to provide such omnipresent, in-situ, and corrective feedback (as specified for example by Hattie and Timperley (2007), and Lysakowski and Walberg (1982)), which adapts to the user's movements, and hence can facilitate an organic and healthy exercise performance. Additionally, we present the results of a user study verifying that SkillAR has no additional disadvantages compared to a conventional screen.

In the following, we provide a brief overview of methods, that are common in recent research, but not directly linked to our approach.

1.1 Point of view in motor feedback

Considering motion feedback, the *point of view* plays an important role. A first-person perspective (Fig. 1, right) provides an immersive and natural viewpoint, as we usually experience the world from this angle. However, the foreshortening deriving from this perspective can make it difficult to perceive the spatial positioning of limbs. Furthermore, some body parts, like the back, can simply not be observed from this limited viewpoint. Optical see-through HMDs (as are used in this paper) always provide a natural first-person perspective in addition to any other superimposed feedback types.

Fig. 1 A person exercising with a ball. Exocentric (left) and egocentric (right) view types with possible target movement (feedback) in transparent green and the actual movement in blue



Alternatively, a common way to display motion feedback is a third-person perspective (Fig. 1, left). This gives the user an overview of his or her body in space. The relation of all limbs becomes clear and deviations from the correct form can be adjusted accordingly. Therefore, Rymal and Ste-Marie (2009) discovered that an exocentric view of an exercise could help athletes improve their ability to mentally visualize the motion, which can lead to learning new skills or improving performance as found by White and Hardy (1998). Moreover, we discovered in a survey (2024) that a third-person view is predominantly used in current literature when providing visual feedback for motor skills.

1.2 Common visual feedback technologies

The most common tool to obtain a full-body view (i.e. third person / exocentric as seen in Fig. 1 and explained in Sect. 1.1) is the mirror. Ballet dancers and their instructors have utilized mirrors to acquire simple visual feedback since the 19th century as stated by Desmond (1997). Likewise, mirrors are to be found in fitness studios and physiotherapy studios all over the world. Furthermore, it is possible to enhance the natural feedback of the mirror with technology. Such enhancements often lead to smart mirrors, which can assist with correct exercise execution (e.g. Kim et al. 2020; Park et al. 2021). Moreover, some approaches use the mirror metaphor in a virtual reality (VR) setting to create an intuitive feedback system (e.g. Waltemate et al. 2016; Hülsmann et al. 2019). These so-called *virtual mirrors* are additionally often used to increase the sense of embodiment as seen in the works of Inoue and Kitazaki (2021), and González-Franco et al. (2010).

Equally important in recent research are conventional displays. Especially displays can provide visual feedback for motor skill training, in particular on mobile devices or computers. In some instances, RMDs mimic the function of mirrors. These approaches are commonly described as *augmented mirrors*. For example, Anderson et al. (2013) visualize an actual avatar and a target avatar on a room-scale

display in addition to a camera stream. Likewise, Trajkova and Cafaro

(2018) provide feedback as a mirrored camera stream with superimposed feedback.

2 Related work

Several approaches addressed the limitations of the first-person perspective in HMDs as described in Sect. 1.1 when providing motion feedback. For example, Chua et al. (2003) provided feedback for tai chi displaying several redundant exocentric feedback avatars around the user. This solved the visibility issue when moving the head, i.e. changing the view direction in the first-person perspective in VR. Additionally, the approach explored different feedback methods, like a single teacher or superimposed wireframe feedback.

Likewise, Han et al. (2017) utilized multiple coaches fixed in space and oriented circularly around the user. However, they transferred the idea into AR. In addition to the redundant instructors, a drone was automatically navigated to record the user. This video stream was then displayed in AR, mimicking a mirror. As a result, the user could see a coach executing the target movement in every horizontal direction juxtaposed with the mirror image. However, while this approach might be well matched for the use case of tai chi, feedback would not be visible in exercises where the view is naturally directed to the ground (e.g. planks, push-ups) or the ceiling (e.g. sit-ups, bench press).

Yan et al. (2015) solved the issue of a limited first-person perspective by juxtaposing the HMD video pass-through with the video stream of an external camera. This image was cropped at the silhouette to blend with the surroundings and create a cohesive mixed reality experience. Although the approach provided an exocentric feedback of the body, the functionality was limited by the position and view direction of the physical camera.

In contrast to the previously mentioned works, Kawasaki et al. (2010) included the perspective of another person by superimposing the user's view with the video see-through stream of an HMD. As a result, the user could see if his or her motions corresponded with the instructor's. This enabled skill transmission in a user study. A similar approach was taken by Kasahara et al. (2016), who also included another person's perspective in an HMD. In contrast to the previous work, the two perspectives were juxtaposed. The two approaches enhance the egocentric perspective, therefore they are well suited for skill training mostly involving the hands, like diabolo juggling (see Kawasaki et al. (2010)) or drawing (see Kasahara et al. (2016)). However, they lack an exocentric perspective and hence an overview of the body, making feedback for whole-body exercises impossible.

In addition, Ikeda et al. (2018) displayed stationary small-scale (1:4) avatars in real-time during the motion and true-scale models during playback feedback. As the system was built specifically for golf swings and players look downwards during swings, the small-scale avatar could be seen during the exercise performance. Afterward, when replaying a recorded exercise, the view was no longer restricted by the exercise. Consequently, the avatar was displayed in true-scale. Although the avatar could be seen well when performing a golf swing, there are some exercises, for which providing feedback with this system would prove difficult. In particular, difficulties arise for exercises where rotating the view direction on the horizontal plane or an almost vertical view direction is necessary.

Lastly, Hamanishi et al. (2019) developed a system that enables the user to view his or her motions from all sides. This was achieved by giving the motion-captured avatar a fixed direction in space. Consequently, the user can walk around it, viewing it comprehensively. The two participants of the qualitative user study seemed to receive the system well. However, it might not be sufficient for all forms of exercise. For example, the feedback might irritate users during exercises where the view direction changes a lot (e.g. sit-ups, torso rotations). Furthermore, the system did not provide a target movement for the user to imitate. Hence, the system does not qualify as *corrective feedback* (see Sect. 1), which is the focus of the work at hand.

3 Omnipresent feedback

This section will explain the SkillAR system in detail. First, Sect. 3.1 and Sect. 3.2 will provide information about the hardware used and the nature of the feedback displayed. This should give the reader an understanding of the whole system. Our main contributions lie in the combination of methods found in Sect. 3.3 to ensure the omnipresence of the feedback.

3.1 System overview

In order to provide visual feedback for motions and in particular exercises, it is necessary to capture the limb positions in space over time. For this reason, we used a 3D camera, which extracted joint coordinates from a recorded point cloud. Consequently, a skeleton-like avatar as seen in Fig. 2 could be constructed on the HMD, showing the executed motion in real-time. In addition to the current motion, we superimposed a recorded movement, which represents an ideal execution of the movement. As a result, the user can now correct the motion following the provided feedback. In addition to superimposition, there are further possibilities

Fig. 2 Screenshot of the motor feedback provided by SkillAR via HMD

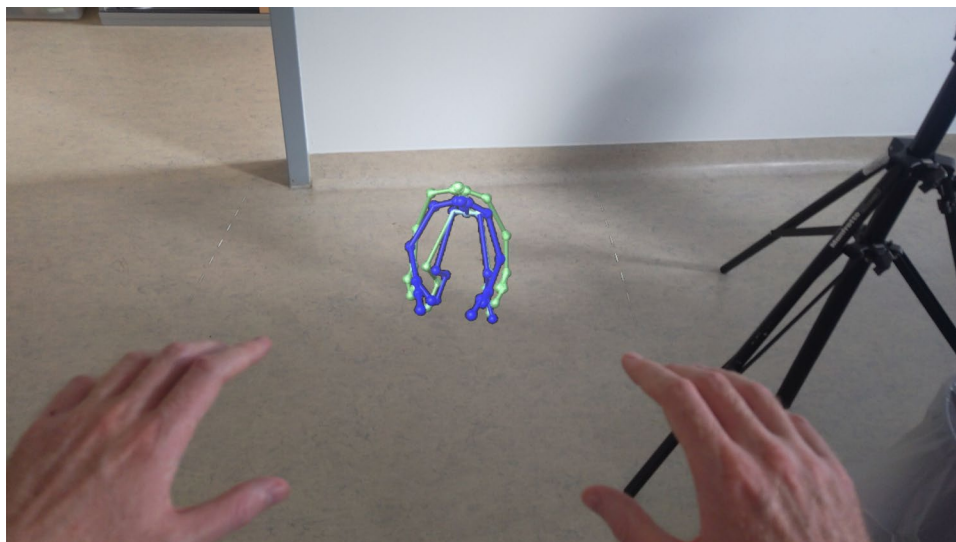
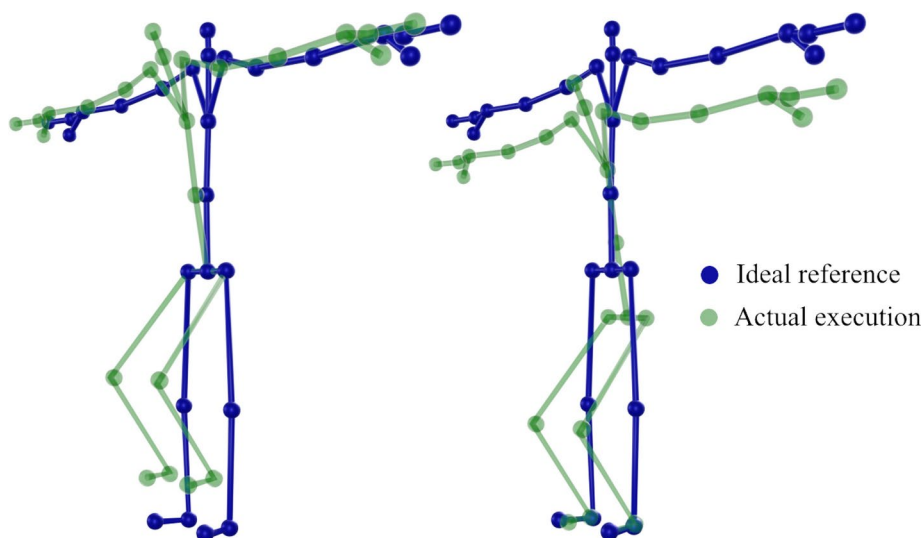


Fig. 3 The reference (in green) is performing the ideal exercise — in this case a squat — while the actual avatar (in blue) is neutrally standing. The result of registering at the pelvis (left) looks as if hovering and can potentially be irritating to the user. In this case, it might be preferable to register the avatars at the foot or the ground



to display comparative visuals as analyzed by L'Yi et al. (2021). Varying avatars and visual cues could impact how users perceive the feedback and therefore yield different results in a user study as conducted in Sect. 4. However, evaluating different visual cues and avatars would exceed the scope of this paper.

3.2 Avatar registration

To superimpose the actual and target pose avatars in relation to each other, in current literature, we often see that the pelvis is used for registration: For instance, Hoang et al. (2016) register at the pelvis based on position and rotation, while Anderson et al. (2013) only utilize the position. In contrast to that, Le Naour et al. (2019) register the target avatar based on the position and rotation of the left foot. Depending on where both avatars are registered, the visualization can be irritating for the user. For instance, when

a target avatar doing squats is registered using the pelvis position with a standing avatar, it seems as if the squatting avatar is hovering, as visualized in Fig. 3.

Consequently, in our work, we fixated the actual avatar and matched the horizontal position of the pelvises. For the vertical position of the target reference avatar, we matched the lowest feet position of both avatars, considering that it should be on the ground since we did not include exercises that involved jumping or hanging (see Fig. 4), as we wanted the exercises to be feasible for a large group of people. Since the only two foot positions in our case are the toe tips and the ankle, matching the toe tips was sufficient. When recording further joints in the foot, they all have to be considered. Regarding rotation, the orientation of the actual pelvis joint was transferred to the target pelvis joint. Additionally, the target avatar had to be scaled bone-wise to match the anatomy of the user. Otherwise, it would be impossible for the user to mimic poses recorded by someone else, as the limbs

Fig. 4 Example exercises (in green) with deviating execution (in blue). It was aimed to provide a diverse set of exercises and deviations (see also Sect. 4.4)

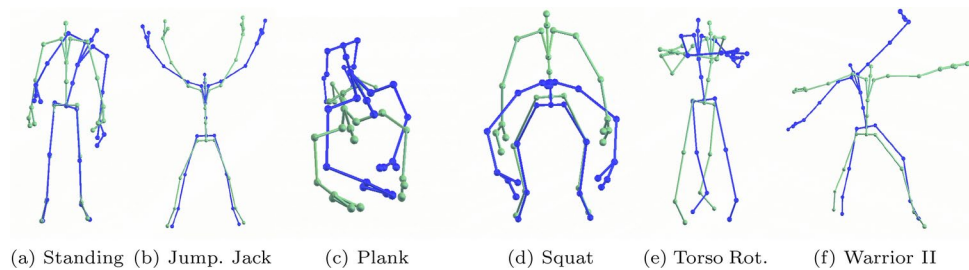
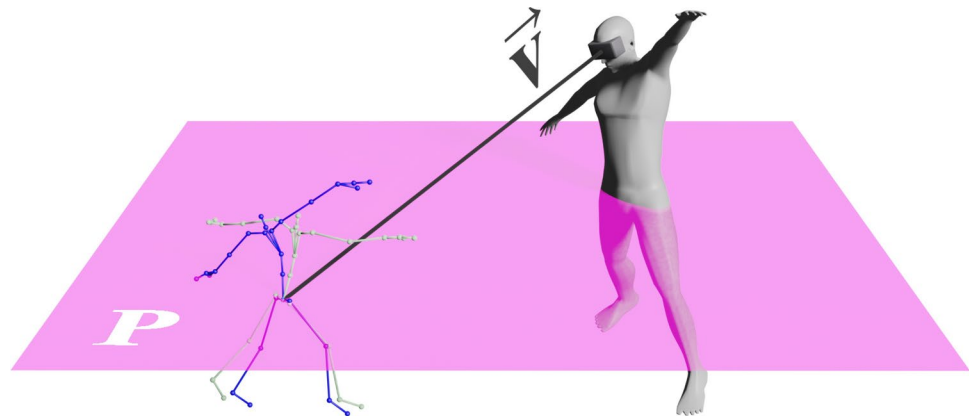


Fig. 5 The feedback avatar is positioned such that its pelvis lies on the intersection of the view direction \vec{v} and the horizontal plane through the user's pelvis P



cannot be stretched or compressed. If the same individual records the ideal and actual execution, this step can be skipped. Furthermore, the avatars were mirrored horizontally, as this makes it easier to take the displayed pose and increases embodiment as stated, for example, by Raffè and Garcia (2018).

3.3 Feedback transformation

Because SkillAR intends to provide omnipresent in-situ feedback, avatar transformation in space is crucial. Utilizing the orientation of the HMD in space and the information we have of the body and surroundings, we can transform the feedback to adapt to the head movement during different exercises and to facilitate a better understanding of the feedback itself. Internal pilot studies have shown various important aspects that we will illustrate in the following.

Placement: The stereoscopic display of the HMD makes it possible for the user to perceive the avatars in space. Consequently, placing the avatars in relation to the actual surroundings of the users helps enhance the interpretation of the feedback. For this purpose, we use the view direction \vec{v} of the HMD to place the avatar. An intuitive way of placing would be to position the avatars' feet at the intersection of \vec{v} with the floor plane. However, this would often lead to the avatar only taking up the upper half of the user's field of view. Instead, the avatars' pelvis represents predominantly a good approximation of the body center. Yet, we can not place the avatar such that its pelvis is at the height of the floor, as the feedback would appear too low (partially below

ground) and hence be irritating. Instead, we constructed a plane P horizontally through the pelvis. The avatars (registered as described in Sect. 3.2) were then positioned by placing the pelvises at the intersection of \vec{v} and P as seen in Fig. 5. Using the pelvis orientation as an indicator for the forward direction of the avatar, the feedback was rotated around the vertical axis, so it faced the virtual camera and thus, the user. This supports the mirror metaphor and hence user guidance as described in Sect. 3.2. The steps above ensure that the feedback is visible in any view direction, i.e. omnipresent.

Snapping: While moving — especially during exercises — the head movement can be unstable, leading to an erratic feedback placement. In an internal pilot study, the visual feedback consequently irritated users. Therefore SkillAR only moves the feedback if the intersection of \vec{v} with P exceeds a certain distance Δ from the current avatar position as seen in Fig. 6. The feedback is moved smoothly to its new position, which leads to a far more stable system behavior similar to snapping.

Scaling: Moreover, the size of the feedback is scaled depending on the distance to the virtual camera to optimally utilize the HMD's field of view. This is illustrated in Fig. 6. Likewise, the snapping threshold Δ is scaled (also visible in Fig. 6) to facilitate a consistent user interaction. Scaling the avatar results in slightly contradicting depth cues: The vergence changes with the distance but the size stays constant due to the scaling. Additionally, the avatars are still appropriately located in the scene, but they might no longer connect with the real floor plane. However, it is to be said,

Fig. 6 The avatars move horizontally if the view direction deviates more than Δ from the center of the visualization, as indicated by the circles. The visualization is scaled up in the distance and down in proximity to the camera. The constant size always matches the field of view of the HMD and irritates the user less as if it were constantly changing

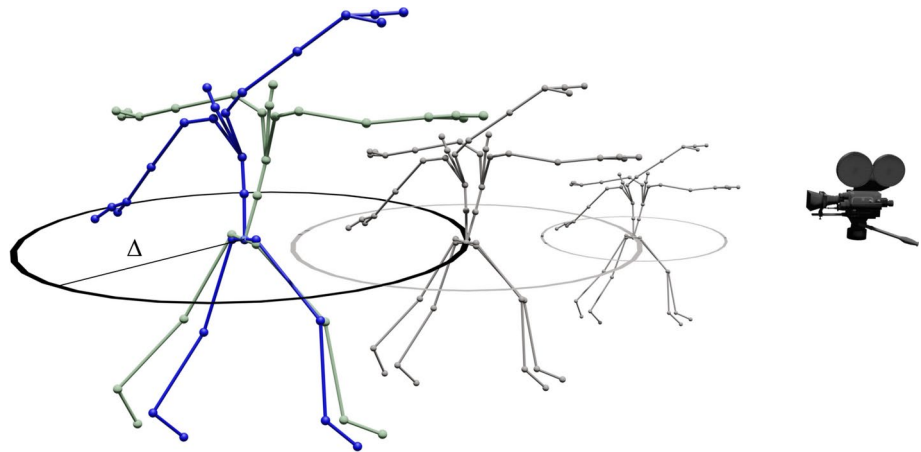
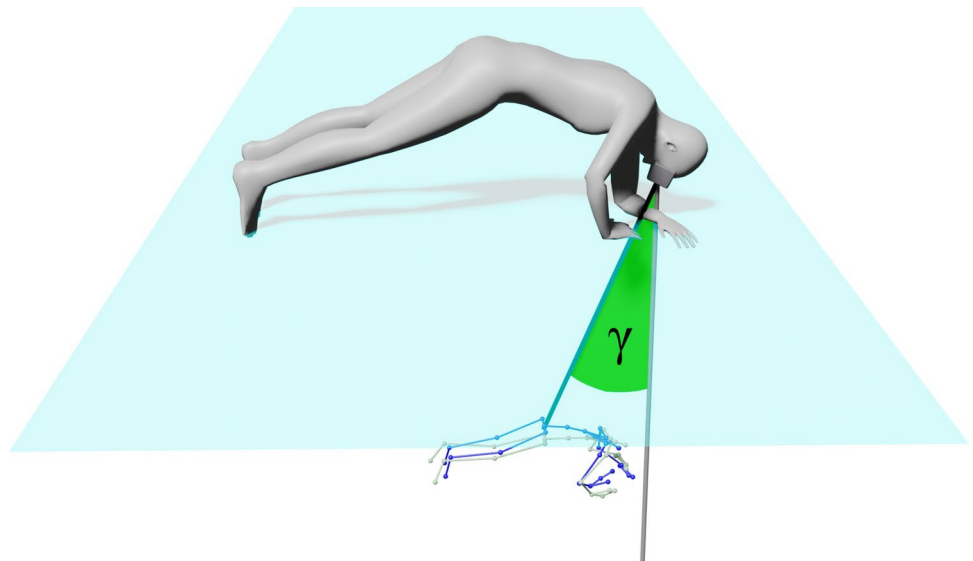


Fig. 7 The feedback transitions to a *top-down mode*, in which the avatars are shown below, if the angle γ between the normal of the floor and the view direction is smaller than 20°



that our priority is to provide omnipresent in-situ feedback rather than creating a perfectly immersive experience. None of the users commented about the above-mentioned topics (see Sect. 6.1).

Top-Down Mode: When looking down, the feedback positioning at pelvis height becomes hard to perceive as it appears too close. This would prevent the user from perceiving feedback when doing exercises that involve a downward head orientation like planks or push-ups. For this reason, when the angle γ between the view direction \vec{v} of the HMD and the vertical axis lies below 20° (see in Fig. 7), SkillAR displays the feedback in constant distance below the user (underneath the physical floor plane). In this top-down view, the avatars are facing the same way as the user and are not mirrored. This provides an overview of exercises facing down, allowing for an AR-supported execution of, for example, push-ups.

Free Mode: Likewise, the feedback positioning mode changes when looking upwards. In this case, the avatars can not be placed at the intersection of \vec{v} and P as they no

longer intersect. Consequently, the avatar positioning is then strictly bound to the view direction. The distance to the camera is constant and the avatar still faces the user. While the avatar's relation to its environment is less clear in this mode, it allows for performing exercises with an upward-facing view direction like sit-ups or bench press.

4 Evaluation

SkillAR provides major advantages for viewing motor feedback: Users can view the ideal exercises while performing exercises without compromising comfort, good form, or safety. We conducted a user study to verify that SkillAR has no additional major disadvantages compared to a conventional RMD. This study was reviewed by the ethics committee and the data protection official of the University of Applied Sciences Worms and carried out following the appropriate guidelines and regulations.

4.1 Participants

For the user study, 32 individuals were recruited from an academic environment. Their age ranged from 22 to 61 ($M = 35.4$, $Mdn = 30.5$, $SD = 11.1$). Their gender was evenly distributed among males and females with 16 (50.0%) individuals each. They rated their prior XR experience on a scale from one to five, with one equaling no previous engagement in XR and five being an XR expert. The rating averaged out at 1.8 ($M = 1.8$, $Mdn = 1.75$, $SD = 0.9$) and their physical activity at 3.1 ($M = 3.1$, $Mdn = 3$, $SD = 0.9$) (one again representing no physical activity and five the maximum). As the visual capabilities of individuals play an important role in the study, the number of participants wearing glasses ($n = 17$, i.e. 53.1%) was documented as well as visual impairments: Three participants (9.4%) reported to have limited or missing spatial perception. One participant exhibited both red-green and blue-yellow color vision deficiency. Regarding gender, the participants' demographics match the population of Europe and the world well. The median age of the participants lies between the median age of Europe (42.2) and the world (30.4) as estimated by the UN in 2023 (2023). However, the educational background is expected to be higher than average in our user group, as the individuals were recruited in an academic environment.

4.2 Apparatus

A Microsoft Azure Kinect 3D camera provided the spatial positions of the joints. It was set up on a tripod about 2.5 m from the subject to ensure that everything was in frame. The visualization and experimental applications were developed in Unity. To display the augmented reality sections of the study a Microsoft HoloLens 2 was utilized. Additionally, to ensure the legibility on the HoloLens, the blinds in the room were closed to provide consistent lighting conditions for the participants. Internal pilot studies showed, that it is hard to perceive the visualization in intensely illuminated environments.

It was ensured the participants understood the avatars, the visual feedback, and the study tasks beforehand. The study took a total of about 20 min with 5 min of introduction. Every task demanded approximately the same duration. Exercises could be skipped or aborted at any time due to health concerns.

4.3 Tasks

In the **comparison** task, the user was asked to identify errors represented by an actual and a target avatar as seen in Fig. 4. The answers were documented and categorized, regarding the accuracy of the addressed body part and correction.

With both — the body part and the correction — being able to be either correct or incorrect, this led to 2×2 categories. It was possible for answers to fall into several categories, as sometimes participants made contradicting remarks or answered both correctly and incorrectly sequentially. Additionally, the time to answer was measured. This task was carried out using the HMD and the RMD.

The **imitation** task involved the users mimicking a pose they saw on the HMD or RMD. After adapting the pose, they gave a command, and the distances of joints, and the time elapsed since the display started showing the pose, were recorded. In between the exercises, no feedback was visible and the users could retake a neutral position and recover from the previous exercise. This should prevent impacting the following exercise performances in any way (e.g. adding time, because the participant has to stand up after an exercise on the ground). This task was carried out using the HMD and the RMD. The supplemental material provides a video of the juxtaposed exocentric and egocentric perspectives of a user imitating poses.

Finally, the participants answered structured **questions** regarding their demographic including age, affiliation, XR experience, physical activity, optical deficiencies, gender, and potential comments.

4.4 Study design

The study aimed to compare SkillAR to conventional RMDs using an in-subject design measuring time and accuracy during the comparison and identification tasks. A few border case exercises were selected, representing various body positions in space. In addition, deviations were chosen to be evaluated during the comparison task (see Sect. 4.3). These deviations require corrections in different directions: Bend vertically, tilt side-wards, raise or lower limbs, twist, etc. Furthermore, the following exercises were each represented by a still pose: Standing, jumping jack, plank, squat, torso rotations, and warrior II as found in yoga. These exercises and the corresponding deviations can be seen in Fig. 4. The order of exercises, tasks, and devices was randomized to eliminate any learning bias. The intent of the study was not to evaluate how well an exercise was executed, but how well the feedback could be comprehended.

5 Results

We suspected a dependency between RMD and HMD regarding the results of the same tasks. Additionally, we had no assumption which device would allow for more accuracy or a faster task completion time. Consequently, a dependent two-tailed t-test seemed most appropriate to evaluate

Table 1 Time measurements for each task in the user study

Device	Comparison		Imitation	
	RMD	SkillAR	RMD	SkillAR
Mean	18.81 s	19.01 s	21.42 s	22.95 s
Median	14.70 s	17.95 s	19.94 s	22.43 s
Standard Deviation	11.44 s	8.17 s	7.78 s	7.80 s

The values representing faster completion times and smaller standard deviation are found in bold for easier reference

Table 2 Accuracy for each task in the user study

Device	Comparison		Imitation	
	RMD	SkillAR	RMD	SkillAR
Mean	75.27 %	92.71 %	10.71 cm	11.14 cm
Median	83.33 %	100.00 %	10.17 cm	10.82 cm
Standard Deviation	11.88 %	12.46 %	2.56 cm	3.33 cm

The values representing higher accuracy and smaller standard deviation are found in bold for easier reference

the results. The results of each individual were averaged to ensure comparability. One evaluation of the comparison task was discarded, as it was not suitable for evaluation due to technical difficulties leading to an incomplete recording. For the t-test, we assumed a significance level of $\alpha = 0.05$.

5.1 Completion time analysis

In each of the study tasks, the completion time was measured. In particular, the time from the start of displaying the exercise to identifying the exercise or adopting the displayed pose. When imitating, this was always starting from a neutral standing pose.

Regarding the comparison task, SkillAR exhibits an insignificantly ($p = 0.92$) longer time to assess the feedback while featuring a lower standard deviation, as seen in Table 1. Likewise, the users required an insignificantly ($p = 0.38$) longer time to mimic the poses displayed with SkillAR compared to the RMD.

5.2 Accuracy analysis

In the comparison task, the percentage of right answers (right body part and right correction) was recorded per user per device. When measuring the accuracy of the pose imitation, we measured the distance of each joint to its counterpart in the ideal pose. These were then averaged to find the mean deviation of the user pose per device.

Assessing the accuracy in the comparison task between the RMD and the SkillAR, we see that the users could identify the corrections significantly ($p = 6 \cdot 10^{-7}$) more precisely with SkillAR in relation to a conventional RMD. As in this case, the p-value is much smaller than α , it seems likely that SkillAR facilitates a more precise interpretation of the given feedback. When imitating the poses, the

precision was insignificantly ($p = 0.37$) higher using the RMD compared to SkillAR.

6 Discussion

In the statistical analysis, we could not verify a significant difference between SkillAR and RMD regarding the comparison time as well as time and accuracy when mimicking poses. The large standard deviation compared to the mean suggests that an even more controlled study with more participants and a more precise camera could be profitable. Further, the measurements of the two analyzed methods seem quite similar. It might be profitable to analyze the equivalence of the methods in future studies. However, our evaluation shows that the participants identify the right feedback significantly more often with SkillAR compared to an RMD. The mean difference is large compared to the standard deviation. We attribute this mainly to the fact, that it is possible to perceive a 3D image of the feedback with the HMD: Due to the two displays in the HMD, a stereoscopic image is created. This represents an additional depth cue compared to the screen. The adaptive nature and the positioning of the feedback in the room could further enhance the interpretation of the scene. Lastly, the rotation of the avatar to the user as described in Sect. 3.3 reduces motion parallax, but it still could be a factor when the feedback is moved away and towards the user.

Additionally, our approach provides the major advantage of an independent viewing direction. This facilitates in-situ feedback in many situations where it otherwise would be inconvenient with a conventional RMD or impossible. Furthermore, the independent head direction enables the user to perform exercises optimally.

6.1 User comments

After the study, the users were asked if they had comments concerning the system. Eleven participants stated that it was easier to perceive the feedback and execute the poses with SkillAR. This happened without being prompted to compare the conventional RMD and SkillAR. The remarks match our findings when evaluating the accuracy with which the feedback was identified with SkillAR. Similarly, this could be attributed to the additional depth perception in AR. In contrast, only one person preferred the feedback on the RMD, as they disliked the feedback adapting to the head movement altogether in SkillAR. Moreover, some users reported understanding the poses not until using SkillAR. This was never observed the other way around and also matches the previous findings.

Additionally, three people reported that they recognized a discrepancy between how they remembered to do the exercise right and what they saw. One of these participants reported neglecting the experience with the exercise and relied completely on the visualization. This emphasizes the carefulness with which exercise feedback must be considered. If the visualization can override motor perception, it is crucial to ensure the feedback is safe for the user.

Furthermore, the participants with limited spatial perception or color vision deficiency reported that they had no problems regarding their deficiencies when using the system.

6.2 Observable behaviors

In addition to the answers and comments of the users, there were interesting behaviors that could be observed multiple times during the study. For instance, where possible, some participants used the HMD like a conventional RMD, always looking at the same spot, although the view direction could be arbitrary. Additionally, many users did not utilize the *top-down viewing mode* for the planks as explained in Sect. 3.3. These observations led us to the conclusion that our system could be leveraged even more with more extensive instructions, explaining the full functionalities of SkillAR. Furthermore, this could lead to interesting new research regarding the reason for these behaviors and the possible consequences, especially regarding deviations from the ideal execution.

Moreover, when adopting a pose, the users predominantly preferred to correct one limb after another — in some cases, the lower and upper body separately. This could be important to consider in the future when providing visual feedback for several deviations at once. Additionally, the registration of the avatars as explained in Sect. 3.2 did not seem optimal in some cases. For example, the users tried to adapt their pose by rotating in space, although the avatar was always facing the (virtual) camera. We authored a paper that is currently under review regarding the optimal universal registration of two avatars.

6.3 Limitations

Although our system provides major advantages for in-situ motor feedback, some limitations must be mentioned. For instance, some exercises or sports are difficult to combine with an HMD. In particular, exercises where the head has to rest on the floor might be uncomfortable or impossible to do, due to the headset interfering. Similarly, the same is true for exercises that involve contact with the head (e.g. pulling on the head for neck stretching). Additionally, some sports might not be suited for an augmentation

with an HMD or require technology tailored to the situation. For example, swimmers might require a special waterproof HMD to receive such feedback. Likewise, sports that require a helmet — like American football or ice hockey — could require a special HMD built into the helmet. Due to the limited robustness featured by standard HMDs, they might not be fit for a sport with such forces applied to the head area. Furthermore, the 3D camera we used to record the users' joint positions in space is not precise enough to detect very delicate movements. Therefore, it would be beneficial to assess the topic with a more precise camera in the future. Lastly, rotating the avatars to face the virtual camera as described in Sect. 3.3 reduces motion parallax as a depth cue. This could limit the interpretation of the feedback. Especially individuals with limited spatial perception could be impaired by this, as they might not be able to profit from other depth cues.

7 Conclusion

The paper at hand presents a novel method to provide omnipresent in-situ motor feedback. It does not only provide feedback where it would be otherwise not possible, but it also enables the user to perform exercises more correctly, more comfortably, and in a non-injurious manner. Additionally, we conducted a user study including 32 participants. We could not detect a significant difference between SkillAR and conventional RMD in terms of identification, imitation time, and imitation accuracy. Moreover, the study showed that the users could identify the errors more accurately when receiving feedback via SkillAR. Thus, we can conclude that our method bears major advantages in contrast to feedback via a conventional RMD, as exercise execution and comfort are not compromised and feedback can be provided independently from the head position.

8 Future work

As discussed in Sect. 6.2, there are novel insights regarding avatar registration exceeding the methods explained in Sect. 3.2. It might be beneficial for user acceptance to implement these methods into SkillAR. Additionally, repeating the experiments using a more controlled environment could lead to the detection of minor effects that were beyond the scope of this paper.

Moreover, the findings of this provide a foundation for further research. Future approaches could explore the potential of the system with a more comprehensive user introduction, explaining to them all the capabilities and using it in a more exhaustive training scenario. With the fundamentals

assessed, it will be valuable to assess the system performance with further poses or when doing dynamic exercises.

Additionally, it might be interesting to extend SkillAR with automatic viewpoint selection methods like described by us in prior work (2024). This could further facilitate feedback understanding. For example, in top-down mode, where the comprehension of the vertical posture of the spine is mostly limited to stereoscopic depth cues.

9 Additional information

As supplemental material, we provide a video of an individual imitating poses as it was carried out in the user study, which is found below.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10055-025-01108-1>.

Acknowledgements The mixed reality part of this work was supported by ProFIL - Programm zur Förderung des Forschungspersonals, Infrastruktur und forschendem Lernen of the University of Applied Sciences Worms. The SkillXR project was supported by the HAWdirekt funding program of the Ministry of Science and Health Rhineland-Palatinate. The authors want to thank Jörg Stimmel and Jana Binninger for the inspiration. The authors appreciate the time and effort all the participants could provide for the user study.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- World Physiotherapy: Annual Membership Census - Global Report (2023). <https://world.physio/sites/default/files/2024-01/AMC2023-Global.pdf>
- Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, Cohen M, Evers S, Finnerup NB, First MB, Giamberardino MA, Kaasa S, Kosek E, Lavand'homme P, Nicholas M, Perrot S, Scholz J, Schug S, Smith BH, Svensson P, Vlaeyen JWS, Wang SJ (2015) A classification of chronic pain for ICD-11. *Pain* 156(6):1003–1007. <https://doi.org/10.1097/j.pain.000000000000160>

- Brepohl PCA, Leite H (2023) Virtual reality applied to physiotherapy: a review of current knowledge. *Virt Real* 27(1):71–95. <https://doi.org/10.1007/s10055-022-00654-2>
- Campo-Prieto P, Cancela JM, Rodríguez-Fuentes G (2021) Immersive virtual reality as physical therapy in older adults: present or future (systematic review). *Virt Real* 25(3):801–817. <https://doi.org/10.1007/s10055-020-00495-x>
- Diller F, Scheuermann G, Wiebel A (2024) Visual cue based corrective feedback for motor skill training in mixed reality: A survey. *IEEE Trans Visual Comput Gr* 30(7):3121–3134. <https://doi.org/10.1109/TVCG.2022.3227999>. (Online as early access since 2022)
- Hattie J, Timperley H (2007) The power of feedback. *Review of Educational Research* 77(1):81–112. <https://doi.org/10.3102/003465430298487>
- Lysakowski RS, Walberg HJ (1982) Instructional effects of cues, participation, and corrective feedback: a quantitative synthesis. *Am Educ Res J* 19(4):559–572. <https://doi.org/10.3102/00028312019004559>
- Rymal AM, Ste-Marie DM (2009) Does self-modeling affect imagery ability or vividness? *J Imag Res Sport Phys Act*. <https://doi.org/10.2202/1932-0191.1035>
- White A, Hardy L (1998) An in-depth analysis of the uses of imagery by high-level slalom canoeists and artistic gymnasts. *Sport Psychol* 12(4):387–403. <https://doi.org/10.1123/tsp.12.4.387>
- Desmond, J.C.: *Meaning in Motion: New Cultural Studies of Dance*. Duke University Press, Durham (1997). <https://doi.org/10.2307/j.ctv11sn2b4.18>
- Kim, S., Seo, D., Lee, S., Kim, Y., Kang, H.W., Choi, Y.-S., Jung, J.-W.: Real-Time Motion Feedback System based on Smart Mirror Vision. 2020 Joint 11th international conference on soft computing and intelligent systems and 21st international symposium on advanced intelligent systems (SCIS-ISIS) 00, 1–4 (2020) <https://doi.org/10.1109/scisisis50064.2020.9322752>
- Park HS, Lee GA, Seo B-K, Billinghurst M (2021) User experience design for a smart-mirror-based personalized training system. *Multimedia Tools Appl* 80(20):31159–31181. <https://doi.org/10.1007/s11042-020-10148-5>
- Waltemate, T., Senna, I., Hülsmann, F., Rohde, M., Kopp, S., Ernst, M., Botsch, M.: The impact of latency on perceptual judgments and motor performance in closed-loop interaction in virtual reality. *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, 27–35 (2016) <https://doi.org/10.1145/2993369.2993381>
- Hülsmann F, Frank C, Senna I, Ernst MO, Schack T, Botsch M (2019) Superimposed skilled performance in a virtual mirror improves motor performance and cognitive representation of a full body motor action. *Front Robot AI* 6:43. <https://doi.org/10.3389/frobt.2019.00043>
- Inoue Y, Kitazaki M (2021) Virtual mirror and beyond: the psychological basis for avatar embodiment via a mirror. *J Robot Mech* 33(5):1004–1012
- González-Franco, M., Pérez-Marcos, D., Spanlang, B., Slater, M.: The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In: 2010 IEEE Virtual Reality Conference (VR), pp. 111–114 (2010). <https://doi.org/10.1109/VR.2010.5444805>
- Anderson, F., Grossman, T., Matejka, J., Fitzmaurice, G.: Youmove: enhancing movement training with an augmented reality mirror. In: *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*. UIST '13, pp. 311–320. Association for Computing Machinery, New York, NY, USA (2013). <https://doi.org/10.1145/2501988.2502045>
- Trajkova M, Cafaro F (2018) Takes Tutu to ballet: designing visual and verbal feedback for augmented mirrors. *Proc ACM Interactive, Mobile, Wearable Ubiquitous Technol* 2(1):1–30. <https://doi.org/10.1145/3191770>

- Chua, P.T., Crivella, R., Daly, B., Hu, N., Schaaf, R., Ventura, D., Camill, T., Hodgins, J., Pausch, R.: Training for Physical Tasks in Virtual Environments: Tai Chi. *IEEE Virtual Reality*, 2003. Proceedings, 87–94 (2003) <https://doi.org/10.1109/vr.2003.1191125>
- Han, P.-H., Chen, Y.-S., Zhong, Y., Wang, H.-L., Hung, Y.-P.: My Tai-Chi Coaches: An Augmented-Learning Tool for Practicing Tai-Chi Chuan. Proceedings of the 8th Augmented Human International Conference, 1–4 (2017) <https://doi.org/10.1145/3041164.3041194>
- Yan, S., Ding, G., Guan, Z., Sun, N., Li, H., Zhang, L.: OutsideMe: augmenting dancer's external self-image by using a mixed reality system. proceedings of the 33rd annual ACM conference extended abstracts on human factors in computing systems, 965–970 (2015) <https://doi.org/10.1145/2702613.2732759>
- Kawasaki, H., Iizuka, H., Okamoto, S., Ando, H., Maeda, T.: Collaboration and skill transmission by first-person perspective view sharing system. 19th international symposium in robot and human interactive communication, 125–131 (2010) <https://doi.org/10.1109/roman.2010.5598668>
- Kasahara, S., Ando, M., Suganuma, K., Rekimoto, J.: Parallel Eyes: Exploring Human Capability and Behaviors with Paralleled First Person View Sharing. Proceedings of the 2016 CHI conference on human factors in computing systems, 1561–1572 (2016) <https://doi.org/10.1145/2858036.2858495>
- Ikeda, A., Hwang, D.H., Koike, H.: AR based self-sports learning system using decayed dynamic time warping algorithm. eurographics symposium on virtual environments (2018) (2018) <https://doi.org/10.2312/egve.20181330>
- Hamanishi, N., Miyaki, T., Rekimoto, J.: Assisting viewpoint to understand own posture as an avatar in-situation. Proceedings of the 5th International ACM In-Cooperation HCI and UX Conference on - CHIuXiD'19, 1–8 (2019) <https://doi.org/10.1145/3328243.3328244>
- L'Yi S, Jo J, Seo J (2021) Comparative layouts revisited: Design space, guidelines, and future directions. *IEEE Trans Visual Comput Gr* 27(2):1525–1535. <https://doi.org/10.1109/TVCG.2020.3030419>
- Hoang, T.N., Reinoso, M., Vetere, F., Tanin, E.: Onebody: Remote posture guidance system using first person view in virtual environment. In: Proceedings of the 9th Nordic Conference on Human-Computer Interaction. NordiCHI '16. Association for Computing Machinery, New York, NY, USA (2016). <https://doi.org/10.1145/2971485.2971521>
- Le Naour T, Hamon L, Bresciani J-P (2019) Superimposing 3d virtual self + expert modeling for motor learning: Application to the throw in American football. *Front ICT*. <https://doi.org/10.3389/ftict.2019.00016>
- Raffe, W.L., Garcia, J.A.: Combining skeletal tracking and virtual reality for game-based fall prevention training for the elderly. In: 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH), pp. 1–7 (2018). <https://doi.org/10.1109/SeGAH.2018.8401371>
- United Nations, Department of Economic and Social Affairs, Population Division: World Population Prospects 2022, Online Edition. Accessed 2024/12/16 (2023). <https://population.un.org/wpp/Download/Standard/>
- Diller, F., Wiebel, A., Scheuermann, G.: Automatic viewpoint selection for interactive motor feedback using principle component analysis. In: VISIGRAPP: GRAPP, HUCAPP, IVAPP, pp. 350–361 (2024). <https://doi.org/10.5220/0012308700003660>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.