SN Computer Science (2025)6:388
https://doi.org/10.1007/s42979-025-03904-7

ORIGINAL RESEARCH

Towards an Optimal Display of Superimposed Avatars for Motor Feedback

Florian Diller[1] · Thorben Frey[1] · Gerik Scheuermann[2] · Alexander Wiebel[1]

Received: 19 December 2024 / Accepted: 20 March 2025

Abstract

Visual motor feedback can support users doing exercises for fitness or rehabilitation. One of the most prevalent visual cues for motor feedback are superimposed human avatars. This work presents novel methods and valuable guidelines for optimizing the display of motor feedback using superimposed avatars. As a foundation, we explain how superimposed human skeleton-like avatars have to be registered according to the performed exercise. This not only helps to raise the quality of technology-supported feedback but also improves the general understanding of feedback provided by superimposed avatars. These foundational guidelines can be used as a base for implementation or further research. Given two registered avatars, we propose a novel method for real-time viewpoint selection. This method uses principal component analysis (PCA) to calculate viewpoints considering feedback (here superimposed avatars). The resulting camera movement is continuous, smooth, and faster than methods found in literature. This allows for real-time use or the viewpoint selection for superimposed avatars in exercise videos. A user study with 39 participants was conducted, verifying our basic assumptions and showing that our algorithm was preferred over the methods found in the literature. Together, the registration and view selection methods provide a powerful resource for optimizing the display of superimposed avatars. Furthermore, each of these methods can be used individually, providing additional value.

Keywords Motor feedback · Optimized display · Viewpoint selection · Avatar registration

Introduction

In modern society, skill training is crucial across various areas, including recreational sports, physical therapy, and professional environments. Enhancing the learning process through interactive technology is becoming increasingly important. Specifically, in motor skill training supported by mixed reality technologies, interactive visual corrective feedback has become particularly significant, as demonstrated in our previous work [1]. This feedback is designed to support individuals in performing specific body movements correctly, thereby reducing the need for constant supervision by professionals. Superimposed avatars, in various types, represent a particular prevalent feedback method. Proper execution of movements is vital in physiotherapy and physical exercise to achieve the desired benefits and prevent injuries. Additionally, the repetitive and controlled nature of movements in physiotherapy and strength training allows for precise feedback provision and the identification of common errors.

Despite their importance, current methods for viewpoint selection in human motion and action feedback do not adequately consider visual cues. Furthermore, many existing techniques are computationally intensive and unsuitable for real-time application. In contrast, this paper and its shorter version [2] identify key factors for optimal viewpoint selection for superimposed avatars used for motor feedback and propose an algorithm for this purpose using *principal component analysis* (PCA). In addition

✉ 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, Worms 67549, Rheinland-Pfalz, Germany

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

Published online: 15 April 2025

Table 1 Exercise examples with corresponding joints for each of the six degrees of freedom to match for an optimized registration as described in Sect. "Optimal Avatar Registration"

	Translation along			Rotation around		
	Vertical axis	Horizontal axes		Vertical axis	Horizontal axes	
Exercise	Up/Down	Left/Right	Forward/Back	Yaw	Pitch	Roll
Squat	Lowest joint	Pelvis	Pelvis	Pelvis	Fixed	Fixed
Leg raise	Lowest joint	Pelvis	Pelvis	Pelvis	Fixed	Fixed
Push-ups	Lowest joint	Pelvis	Pelvis	Pelvis	Fixed	Fixed
Warrior II	Lowest joint	Pelvis	Pelvis	Pelvis	Fixed	Fixed
Plank	Lowest joint	Pelvis	Pelvis	Pelvis	Fixed	Fixed
Dips	Hands	Neck	Neck	Neck	Fixed	Fixed
Pull-up	Hands	Neck	Neck	Neck	Fixed	Fixed
Jump	Pelvis	Pelvis	Pelvis	Pelvis	Fixed	Fixed

to [2], this work explores valuable aspects of avatar registration. Not only do these directly impact the viewpoint selection, but they also give us a further understanding of how to register superimposed avatars to facilitate an understanding of the feedback from the user (as illustrated in Fig. 1), independent of the rendering methods.

Related Work

To follow the structure of the methodology in Sect. "Methodology", we subdivided the related work section according to the twofold nature of the main contributions of the paper at hand.

Registration

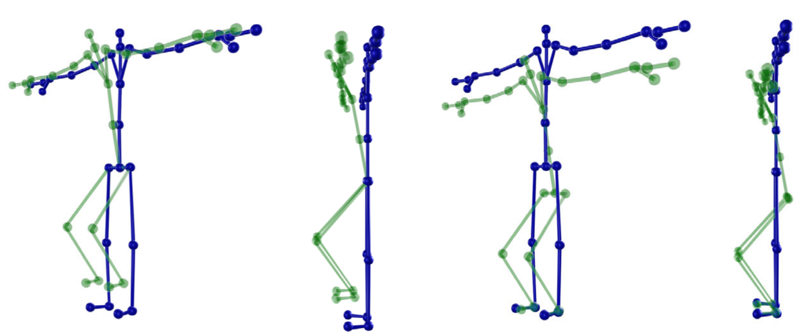
In the current literature, we find a plethora of both rigid and non-rigid registration methods, as analyzed by Tam et al. [3]. When registering an actual and target pose it is essential to use rigid registration, as we want to preserve potential deviations from the ideal form. These would be diminished by a non-rigid registration method, as it would deform the ideal avatar. Specifically for rigid registrations, there are several automatic methods, as Yaniv [4] and Bellekens et al. [5] stated.

In some cases, a simple registration based on a single joint is completely sufficient. For example, for optimizing posture guidance in VR, Hoang et al. [6] registered the target avatar based on the position and rotation of the pelvis. Instead, Anderson et al. [7] used only the pelvis' position while maintaining the orientation of the recorded target movement. Subsequently, the approach was evaluated, and users performed two ballet and two abstract movements with the proposed system. In contrast, Naour et al. [8] registered the avatars based on the position and rotation of the left foot. This way, the football-throwing motions of a learner were superimposed with those of an expert. Additionally, we find many approaches in the literature that register avatars but lack explanations for these methods.

View Direction

As demonstrated by Bouwmans et al. [9], robust PCA has numerous applications in the field of visual computing. For instance, Skaro et al. [10] introduced a method to reduce errors common in marker-based motion tracking.

The literature offers several methods for viewpoint selection regarding polygonal objects as surveyed by Bonaventura et al. [11]. However, regarding human actions or movements, there seem to be fewer approaches. For



(a) Superimposition registered at pelvis (b) Superimposition registered at feet

Fig. 1 Highlighting the importance of registration: When registered at the pelvis (**a**), the squatting target avatar (in green) seems to be floating. Registering at the feet seems to provide more intuitive feedback, connected to the environment

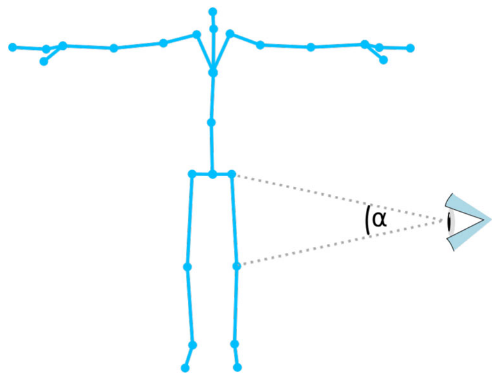


Fig. 2 Measure for the self-occlusion of the skeleton by Ishara et al. [17]: Joint Mutual Occlusion. Originally published in [2]

example, Rudoy et al. [12] developed a method to generate a three-dimensional volume from several successive frames to select the best physical camera for television broadcasts and similar applications. Whereas Kiciroglu et al. [13] proposed an algorithm predicting pose estimation accuracy, facilitating drone navigation to the calculated view point. Shi et al. [14] introduced an algorithm to determine the best viewpoint using *Kinematics Significance Based Saliency*, which orients figures and objects to show their most protruding features.

Wang et al. [15] utilized information theory and deep reinforcement learning to select a single viewpoint for action sequences. In parallel, Choi et al. [16] extracted key frames from motion data to create a sequence of stick figures representing the initial motion data.

Ishara et al. [17] calculated the optimal camera position for robot navigation using a camera mounted on top of the robot. Their approach involves calculating the so-called *Joint Mutual Occlusion* (JMO), which considers the angle α between two joints and the viewpoint, as shown in Fig. 2. The angles α_{nm} between joints n and m are summed and normalized, where $n, m \in N$ and $n \neq m$, with N representing the number of joints. This approach results in $\frac{N!}{2(N-2)!}$ calculations of α [18] and could therefore be computationally expensive.

Kwon et al. [19] proposed a method that results in a weighted sum of three metrics: *normalized limb length*, *normalized area of a 2-D bounding box*, and *normalized visible area of a 3-D bounding box*. In addition, they also presented an algorithm that avoids recalculating weights for each frame by summing the three metrics without weights. However, their algorithm, designed for static poses, requires recalculation for each frame in videos. Despite these methods automatically selecting camera positions for human poses, they are insufficient for visual feedback, as the skeleton can occlude the feedback, making it difficult to perceive, as analyzed by Nundy et al. [20] and discussed in Sect. "Introduction".

The approaches by Ishara et al. [17] and Kwon et al. [19] were compared to our method in a subsequent user study, as they were the only methods applicable to human figures with feedback. For more details, see Sect. "Viewpoint Selection Evaluation".

PCA is commonly used to reduce dimensionality in data sets for machine learning [21]. The principal components represent the main independent directions in which data points spread. For spatial data, three independent directions are involved. The first two principal components represent the main spread directions, while the third component provides a good viewing direction, perpendicular to the first two. This is equivalent to reducing the dimensionality from three to two, as the rendered image of 3D objects is displayed in two dimensions. Assa et al. [22] used this method to calculate camera paths. However, their use case differs significantly, as they compute camera paths that involve camera cuts, which we avoid due to the short duration of exercise repetitions where cuts can be disorienting. Additionally, the work of Assa et al. is action-based, while ours is feedback-based, requiring additional measures to ensure feedback visibility. Lastly, their approach is not real-time capable, being computationally intensive and requiring the entire motion sequence for computation.

Methodology

The main contribution of this paper is twofold: In section Sect. "Viewpoint Selection" the viewpoint selection for two superimposed avatars is explained, as can be found in the shorter version of this paper [2]. In addition to that, the registration, which directly impacts the viewpoint selection, is elaborated on in Sect. "Registration".

Registration

The registration of two exercises represents in most cases the foundation of visual feedback for motor skill training, in particular, superimposed avatars. When registering a superimposed actual and target avatar for a given exercise, we need to keep certain factors in mind to facilitate a better understanding of the scene for the user. For instance, the spatial relation of the avatars to the environment helps the user with orientation. Likewise, corrections for typical deviations have to be anticipated. In summary, a well-made registration can lead to a far more intuitive understanding of visual feedback.

The registration is directly linked to the other aspects of our feedback optimization. Not only does the registration directly influence the inputs of the PCA viewpoint calculation (i.e. Δ_n and $\vec{v}_{Fn}$ in Eq. 1 found in Sect. "Viewpoint

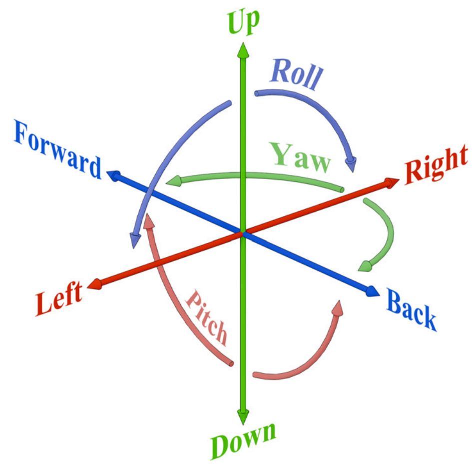


Fig. 3 Six degrees of freedom to define the placement of a rigid body in space

Calculation”), but it is a necessary prerequisite. Furthermore, our PCA-based viewpoint method as explained in Sect. “Viewpoint Calculation”, is designed to work with smaller deviations. Therefore, an appropriate registration method is important.

Optimal Avatar Registration

An optimal avatar registration is highly dependent on the specific exercise performed and even on the individual interpreting the visual feedback. However, there are a few key aspects that are crucial in helping users to comprehend feedback using two superimposed avatars. In the following, we present guidelines, that can be used to register avatars for a certain scenario or to implement optimal avatar registration independently.

One single joint is oftentimes sufficient for registration. Especially for use cases with a limited selection of exercises, it can lead to satisfying results, as described in Sect. “Registration”. However, to avoid irritating users it is important to connect the avatars with their spatial surroundings. For this purpose, we consider the six degrees of freedom (6DoF) for placing a rigid body in 3D space, as depicted in Fig. 3. We limited the 6DoF to four different categories, adequate to optimize registration:

- Vertical alignment (up/down)
- Horizontal alignment (left/right & forward/backward)
- Rotation about vertical axis (yaw)
- Rotation about horizontal axes (pitch & roll)

To provide the best avatar registrations, we match these 4 different degrees of freedom with the following characteristics from the fixed avatar:

Vertical alignment: As mentioned above, it is crucial to connect the target avatar to the environment. Otherwise, the user might be irritated, as the avatars seem to defy the rules of physics. Depending on the exercise, there are different connections to the surroundings: When doing exercises while standing like squats or lateral raises, the feet rest safely on the ground, thus connecting us with the environment. When hanging (e.g. pull-ups) or supporting the weight with the arms (e.g. dips) the hands connect us to the outer world. The user expects the target avatar to be connected to the environment the same way. Consequently, the connection to the environment represents a good measure of aligning the avatars. That means when the user is standing on the ground, the avatar squatting seems to do the same (see Fig. 1). We can further improve the registration if the lowest point is matched for horizontal alignment. This way, when the user lifts a foot, the avatar still stays on the ground. Especially concerning exercises on the ground, this step leads to a better understanding of what to do, even when standing in a neutral position.

Horizontal alignment: The torso represents a big and important part of the body and also contains the center of mass while standing. Therefore, a horizontal alignment of both the actual and target avatars proved to be most intuitive based on the torso. Slightly better results can be achieved when choosing the point closest to the limbs connecting to the environment (i.e. the pelvis for standing exercises, a point between the shoulders for hanging, etc.).

Rotation around vertical axis: The rotation around the vertical axis plays an important part in registering two avatars. In most cases, this rotation represents the orientation within the environment, which is arbitrary except when working with stationary equipment. This means the rotation around the vertical axis can be freely adapted without compromising the correct execution of an exercise. As mentioned in Sect. “Registration”, there are a few joints with which the registration can seem intuitive, the joints in the torso emerge as the best option. In particular, using the same joint that we chose for the horizontal alignment delivers good results when also used as a reference for vertical rotation.

Rotation around horizontal axes: During exercise performance, the rotation around the horizontal axes (i.e. pitching and rolling) is highly dependent on the exercise. Altering these rotations can lead to false interpretations of the feedback. Therefore, we see these rotations as exercise-inherent measures, which should not be altered (or only with great care). Consequently, the pitching and rolling rotations might best be carried over from the exercise recording or reference movement. Otherwise, the avatars seem too adaptive, possibly leading to wrong movement execution.

Scaling

When two avatars are superimposed, depending on the recorded individuals, the scale of the limbs is not necessarily the same. Scaling one avatar according to the total size of the other avatar only partially solves the problem. As the proportions (e.g. torso length compared to leg length) can differ between individuals, a uniform scaling can still result in inconsistencies. Therefore, if the recorded exercises compared potentially stem from different individuals, it is advisable to scale the target avatar bone-wise to match the actual live avatar. Here, it is important to ensure that the angles between the bones, i.e. at the joints, stay consistent. Otherwise, the scaling will falsify the exercise.

Registration Limitations

Avatar registration is a complex subject. The most generally applicable guidelines can be found above. However, it seems as if there is no optimal approach for every individual, situation, and use case. Especially when considering the details concerning the rotation around the horizontal axes, it seems in some cases there is a trade-off between the exercise's validity and perfect registration. Furthermore, scaling the target avatar eliminates most inconsistencies in a bone-wise fashion, yet the angles between the bones at the joints stay consistent. Consequently, given certain proportions and thus centers of mass of two individuals compared, the scaling could lead to the target exercise not being the most optimal. Although the methods mentioned above yield good results for most exercises, even when registering with a neutral standing position, hanging exercises are hard to perceive while standing. This results from missing reference points of the environment. Once the actual avatar is hanging, we can use the hands as reference points (in particular, horizontal alignment).

Examples

To establish an understanding of the methods explained in Sect. "Optimal Avatar Registration", we chose a few exercises with corresponding joints to match the six degrees of freedom and will discuss them in this section. The exercises can be seen in Fig. 4 and the joints used for alignment can be taken from Table 1. The examples were each chosen to be representative of a category of exercises. They are not meant to be complete, but rather a diverse collection, where finding examples close to many use cases is possible. Regarding registration, the example exercises could be sorted into 3 categories, as seen in Table 1, based on how the individual connects with the environment:

Connection to floor: Exercises that connect with the feet on the floor (like squats and warrior II), and those that involve lying down (like leg raises, plank, and push-ups) both profit from the same registration. Here, the pelvis was used to align the avatars on the horizontal plane. As a result, the torsos and therefore, the body's center were registered quite well. The avatars were vertically placed by aligning the lowest joint of each. Consequently, the avatars were intuitively placed on the floor, even if the movement of the actual avatar did not (yet) follow the target avatar (see Fig. 4 and Fig. 1). Concerning rotation, the pitch and roll rotations, as depicted in Fig. 3, were fixated as in the recorded exercise. As already stated in Sect. "Optimal Avatar Registration", the pitch and roll rotations seem inherent to the exercise, as altering them can falsify the execution. However, the yaw rotation was based on the same joint used for horizontal alignment: The pelvis. This aligns the main orientations of the avatars.

Connection to equipment: Alternatively, during an exercise, the individual could connect with the environment on a piece of equipment. For example, this is the case when doing dips or pull-ups where the individual is either supported or suspended by the arms, as seen in Fig. 4. As a joint for both, horizontal alignment as well as yaw alignment, the base of the neck (i.e. center between shoulders)

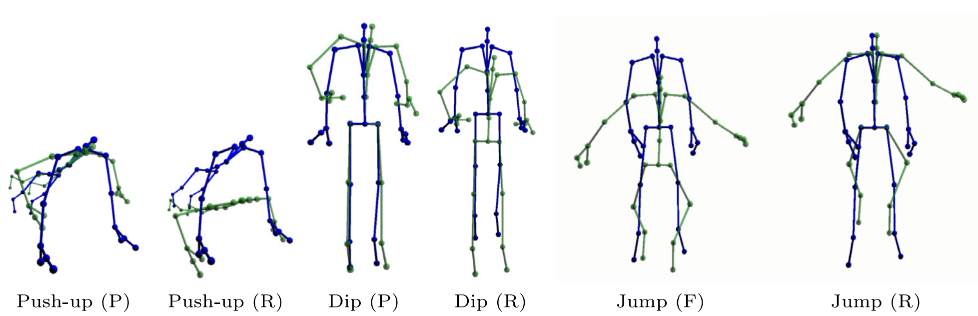


Fig. 4 Start and execution position of exercise examples registered according to the registration methods in Sect. "Optimal Avatar Registration" (R), the feet (F) and only the pelvis (P). The exercises and registration parameters can be found in Table 1

was chosen. Here, this provides a more intuitive registration point than the pelvis, as the arms connect to the environment. For the same reason, the hands were chosen as the base for a vertical alignment. Since the bars for dips or pull-ups represent the connection to the environment, registration at the hands appears stable and intuitive.

No connection to the environment: Lastly, some exercises require no connection to the environment (like jumping or swimming). Still, it is possible to register avatars in these cases. To do so, the pelvis is chosen as the center of the body, as seen in Fig. 4. The pitch and roll rotations are again fixated on the target avatar from how the exercise was recorded. The remaining alignments (i.e. up/down, left/right, forward/backward, yaw) are all done according to the pelvis.

Viewpoint Selection

To facilitate a user-friendly and intuitive viewpoint selection, several things have to be considered. We will discuss these in detail in Sect. "Perspective Considerations". With the foundational registration methods discussed in Sect. "Optimal Avatar Registration", it is possible to calculate an optimized viewpoint, as presented in Sect. Viewpoint Calculation".

Perspective Considerations

In the literature, concrete rules for generating good perspectives are lacking. However, based on user preferences and logical argumentation, several criteria and hints can be extracted to determine what facilitates a good viewpoint.

Polonsky et al. [23] identified seven measurable view descriptors but concluded that determining a universally good view of an object is challenging. None of the view descriptors alone provides a general measure of viewpoint quality. However, some clues are available for treating specific objects. For instance, Zusne [24] empirically demonstrated that humans prefer a frontal view of objects with eyes and a face.

When looking at motor feedback, here in particular superimposed avatars, the positions of the joints and the angles between limbs are particularly critical for understanding movements. However, as previously analyzed by Nundy et al. [20], angle perception is highly perspective-dependent. This is especially true in computer-rendered perspectives due to screen projection distortions, as shown in FigS. 5 and 6. While stereoscopic viewing (e.g., real-world or head-mounted displays) helps depth perception and angle interpretation, monoscopic rendering does not offer the same possibilities. Additionally, occlusion can impede understanding of the human pose, with self-occlusion of the avatar's limbs behind each other. Similarly, visual feedback cues can be obscured by the avatar itself or by other cues, as depicted in Fig. 6 and 7.

Given the lack of a general description for a good view, we need to define the characteristics of a good viewpoint for our specific use case. In our context, we frequently use the metaphor of a virtual camera, common in rendering, to describe the viewpoint and viewing direction. Following Zusne's [24] findings, we prefer an approximately frontal view of the human pose, meaning the virtual camera should be oriented towards the front of the pose rather than from behind. Additionally, the camera's up-vector should align with the world's up-vector to avoid viewer confusion, as this is the biologically typical perception for humans. Furthermore, we aim to minimize the occlusions of the avatars demonstrating the movement execution. Finally, in our use case, visual feedback is provided to correct movements or poses, and this feedback must be clearly visible. Therefore, feedback should not be occluded by the avatar or by itself and should be as perpendicular to the view direction as possible.

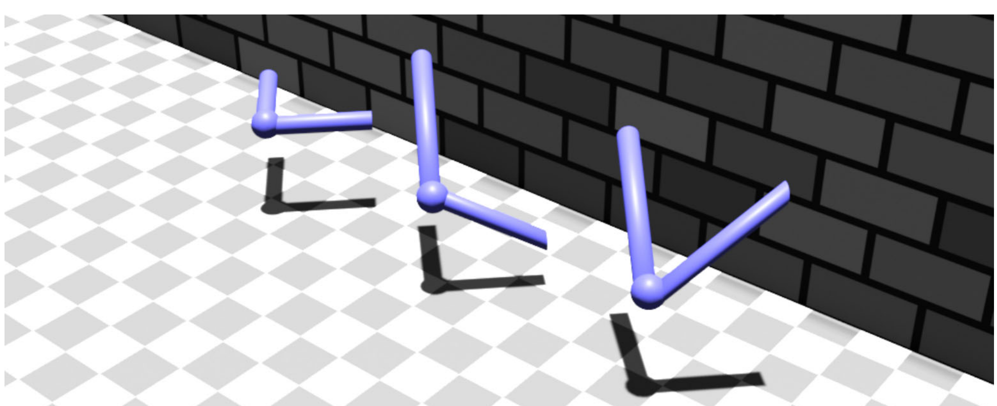


Fig. 5 Highlighting the importance of viewpoint selection: Three *different* joint angles produce *identical* shadows when projected onto the ground, implying they appear identical from an overhead perspective. Inspired by Nundy et al. [20]. Originally published in [2]

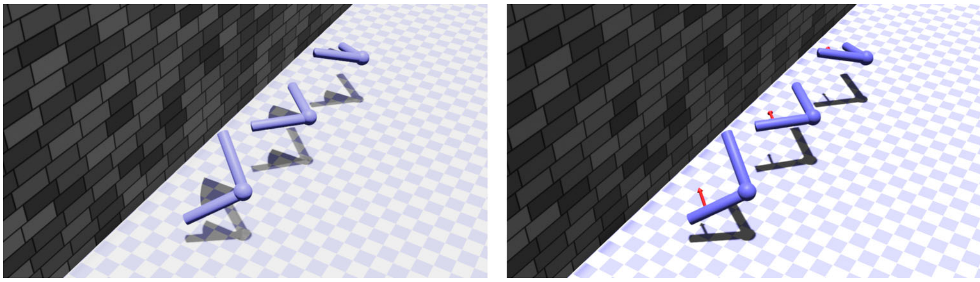


Fig. 6 Feedback for the same angle viewed from different perspectives. Two feedback cues: circular sector (left) and arrow (right). From left to right: Perfectly visible, visible, and hardly visible feedback. The shadows demonstrate that viewpoint affects not only the perception of the geometry but also of the feedback. Originally published in [2]

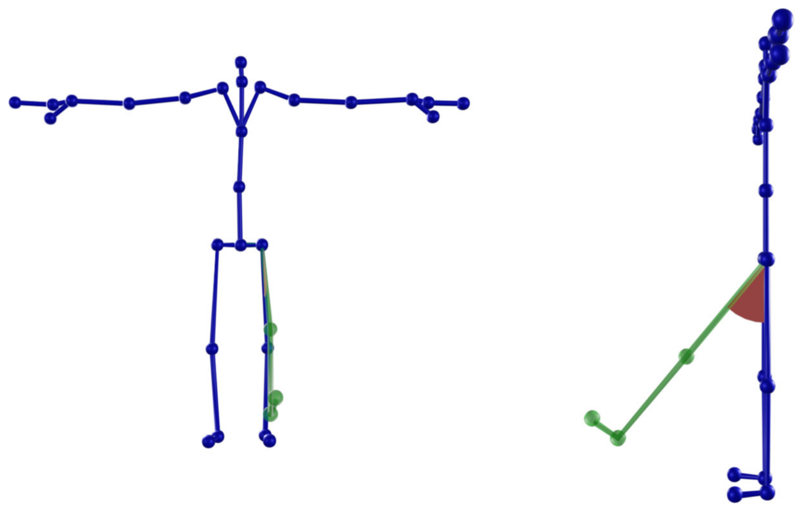


Fig. 7 Skeleton of a human pose with feedback from two perspectives. Two visual feedback cues are shown: A red angle sector and a superimposed avatar (here green skeleton). The feedback is hardly visible from the frontal perspective on the left. Inspired by [2]

When selecting perspectives for human motions and corresponding feedback, it is important to consider the dependencies of different body parts. Specifically, the limbs are hierarchically linked, meaning that moving the upper arm will cause the lower arm and hand to follow. Consequently, perspectives for such motions should ideally employ a *hierarchical drill-down mechanism* to prioritize viewing along the hierarchy.

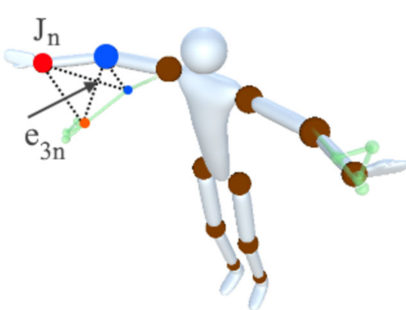


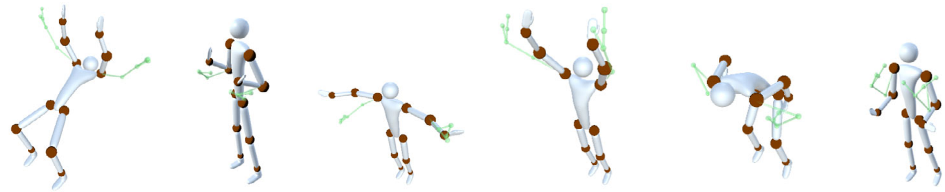
Fig. 8 If Joint J_n (in red) deviates from the target position, a PCA is conducted including the corresponding target joint (in orange) and their parents (in blue). The eigenvector $\vec{e}_{3n}$ then gives us an optimal view direction $\vec{v}_{Fn}$ of the feedback for J_n. The resulting view direction is orthogonal to the plane defined by the eigenvectors $\vec{e}_{1n}$ and $\vec{e}_{2n}$. This plane approximates the distribution of the considered joints and does not interpolate them. Originally published in [2]

Viewpoint Calculation

As presented in Sect. "View Direction", the existing literature does not yet provide an optimal viewpoint calculation for human motions with visual feedback suitable for skill learning. Most approaches are optimized for human actions, leading to the potential invisibility of feedback from an action-optimized viewpoint. In the following, we guide you through our computationally inexpensive method for calculating a viewpoint for human actions with feedback. Equation 1 shows the calculation of our view direction $\vec{v}_d$:

$$\vec{v}_d = w \cdot \vec{v}_S + \sum_{n=1}^{N} (\Delta_n - \delta_0) \cdot \vec{v}_{Fn} \tag{1}$$

Calculating $\vec{v}_d$ involves the following variables: w represents a weight to adjust the impact of the view between the whole skeleton and the feedback; $\vec{v}_S$ is the viewpoint optimized for all joint coordinates (i.e., the actual skeleton); N represents the number of joints that exceed a given



(a) Bench press (b) Curl A (c) Lateral raises (d) Shoulder press (e) Bend over row (f) Curl B

Fig. 9 Sample exercises with deviations as described in Sect. "Sample Exercises". Originally published in [2]

deviation threshold δ_0; Δ_n is the deviation of a joint J_n from the intended target position; the variable δ_0 is a constant threshold of the deviation; and $\vec{v}_{Fn}$ is the view direction optimized for the feedback corresponding to joint J_n, i.e. the deviating of J_n and its corresponding joints, as seen in Fig. 8. We do not consider rotations in the calculation, because they also inevitably cause deviations in distance.

Some motion capture systems provide recorded spatial data as three-dimensional joint coordinates (see Sect. "Experimental Setup for Exercise Recording" for our data acquisition method). When we conduct PCA over such a point cloud of joint coordinates, the first two eigenvectors $\vec{e}_{1S}$ and $\vec{e}_{2S}$ represent the two main spatial spread directions. The third eigenvector $\vec{e}_{3S} = \vec{v}_S$, perpendicular to the first two, consequently providing a well-suited view direction $\vec{v}$ for all joints, as explained in Sect. "Registration". In other words, the point cloud representing the whole skeleton is most spread out along the horizontal and vertical axes of the captured camera picture. Consequently, the view direction $\vec{v}_S$ is optimal for comprehending motions and poses. This method is also presented in Assa et al. [22].

Because the view direction should be optimized for corrective feedback corresponding to the deviations of the exercises, we must consider the deviating joints. For this purpose, we locally apply the above-mentioned PCA viewpoint calculation. The PCA is conducted with the actual and target joint coordinates and the corresponding parent joint coordinates as seen in Fig. 8 for joints $J_n, n \in [1..N]$, with deviations Δ_n, that exceed the deviation threshold δ_0. Consequently, the resulting eigenvector $\vec{e}_{3Fn} = \vec{v}_{Fn}$ is a suitable view direction for displaying joint J_n, its parent, the corresponding optimal joint position, and its parent. This is illustrated in Fig. 8, where the considered joint J_n is shown in red, the optimal joint position in orange, and the corresponding parent joints are depicted in blue.

In Eq. 1, the factor Δ_n (minus the threshold δ_0) of $\vec{v}_{Fn}$ increases the influence of joints depending on their deviation. This automatically considers a hierarchical drill-down mechanism (see Sect. "Perspective Considerations"), since lower hierarchy joints (farther from the body center) usually have a higher absolute deviation, as they are impacted by the deviations of the higher hierarchy joints (closer to the body center), adhering to the intercept theorem. The subtraction of the threshold δ_0 ensures continuous camera movement, so that the impact of deviating joints continuously increases or sets in from zero. The sum of all $\vec{v}_{Fn}$ represents a feedback-optimized view direction for all joints exceeding the deviation threshold. This could be seen as the calculation of the mean of all $\vec{v}_{Fn}$ without the division. The division is unnecessary, as the length of the view direction vector is irrelevant.

The view direction optimized for the skeleton $\vec{v}_S$ is weighted with the constant w to adjust optimization between the skeleton and feedback. Values of $\delta_0 = 50$ and $w = 3\delta_0 = 150$ showed the best empirical outcomes for our use case. This holds several implications:

- The eigenvectors resulting from the PCA, and therefore the view direction vectors, are normalized, i.e. they have a length of 1. In the virtual 3D space, we applied a scale of $1\ unit = 1\ mm$. Consequently, the deviation threshold δ_0 corresponds to 50 mm.
- To have the same impact as the skeleton-optimized view direction $\vec{v}_S$, the feedback-optimized view direction $\vec{v}_{Fn}$ of a single joint would need to have a deviation of 200 mm, consisting of a 50 mm minimal threshold plus 150 mm of weight.
- The deviations of several joints together can exceed 150 mm (plus threshold) to have the same impact on the resulting view direction as the skeleton as a whole.
- If multiple joints do not exceed the 50 mm minimal threshold, the skeleton has 100% impact, thus the viewpoint is optimized for just the skeleton.
- Because we consider the absolute deviation (instead of relative to the parent), the deviations of lower hierarchy joints and their parent joints are dependent. This results in a hierarchical drill-down mechanism, as explained in Sect. "Perspective Considerations", where the joints closer to the body center have a higher impact on the view direction.

To calculate the viewpoint for the virtual camera, we subtract the normalized view direction $\vec{v}_d$ from the location of the focus point, which will be centered in the rendered frame (in our case, the joint representing the pelvis location, since it is a good representation of the body's center). The distance to the focused point can be set by multiplying a constant. The digital distance corresponding to 2 m yielded the best results for us, as all exercises were in frame at this distance. However, this highly depends on the settings (e.g., focal length) of the virtual camera chosen.

If $\vec{e}$ is an eigenvector, $c \cdot \vec{e}$ is also an eigenvector, for all $c \neq 0$ [25]. Consequently, $-\vec{v}_d$, the flipped eigenvector of $\vec{v}_d$, is also a viable view direction. Initially, we select the direction resulting in a more frontal view of the avatar, since this is the predominantly preferred view [24]. For every subsequent frame, we select the direction (from $\vec{v}_d$ and $-\vec{v}_d$) with a smaller angular difference from the previous frame's direction, ensuring smooth camera motion.

Although the third eigenvector of the PCA follows a smooth path, the view direction (i.e. camera) tends to rotate around the avatar, contradicting Zusne's [24] findings that a frontal view is commendable. To resolve this, we projected view angles from behind to the frontal plane, bypassing the predominantly small number of frames that feature a view from behind and showing a view from the side. Consequently, the projection affects the camera view only slightly and briefly.

Existing view selection approaches often focus on solving an optimization problem, iterating over a limited number of potential viewpoints, and choosing the one with the best score. This can result in a high number of costly iterations or erratic camera motion if the number of potential viewpoints is too small. Additionally, the best-scoring viewpoints in consecutive frames might be far from each other, resulting in inconsistent camera movements. Our method, however, provides continuous camera movement, as none of the mathematical operations in Eq. 1 compromise consistency, and the PCA computations are conducted for continuously moving point clouds.

Empirically, no cases were encountered where a null vector arose from the calculations above. We also assessed stability regarding the PCA, as it is mathematically possible, that the camera view flips if the second and third eigenvectors have approximately the same length and deviate just slightly. In our experiments, this never occurred.

Viewpoint Selection Evaluation

To verify the viewpoint selection method described in Sect. "Viewpoint Calculation", a user study was conducted. The necessary preliminaries, the study design, and the subsequent evaluation methods are found in the following.

Participants

For the user study, 39 individuals were recruited from an academic environment. These were predominantly computer science students between the ages of 20 and 30. More than half of the participants reported frequently exercising and considering movement-related aspects, giving ratings of four or higher on a five-point scale. This shows that the participants were somewhat acquainted with similar exercises and their execution. By comparison, a much smaller number of physiotherapy clients were represented in the study. More than half of the participants reported receiving

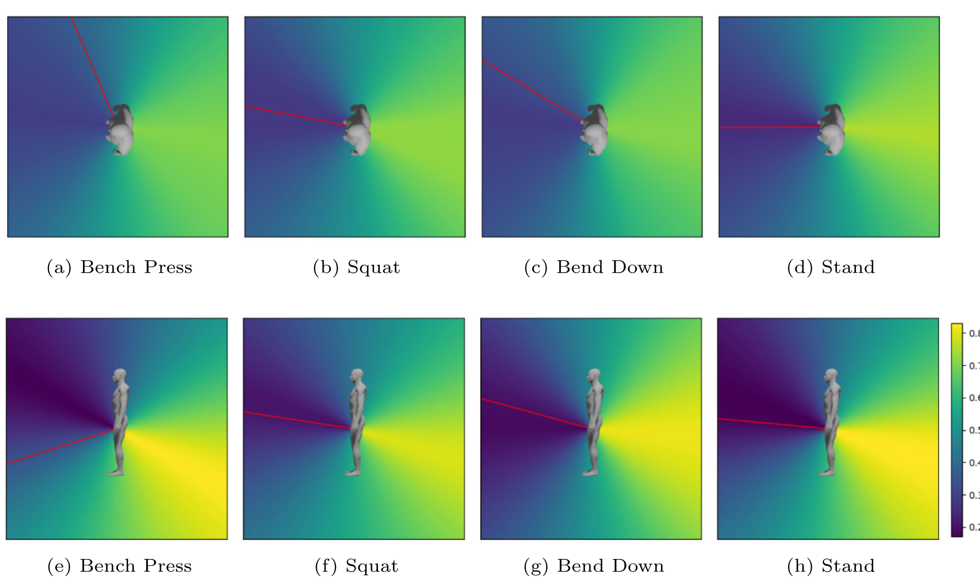


Fig. 10 *View Selection Error* (VSE) for different viewing angles from the top (**a–d**) and side (**e–h**) using the benchmark of Dutagaci et al. [30]. The red line represents the view direction selected by our method. The human figure only shows spatial orientation and does not represent the executed movements. Originally published in [2]

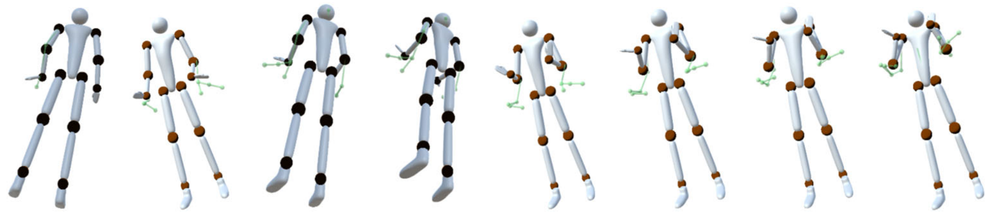


Fig. 11 Image sequence, taken from a video of a biceps curl exercise with deviation. The viewpoint is optimized by the *Joint Mutual Occlusion* algorithm by Ishara et al. [17]. Originally published in [2]

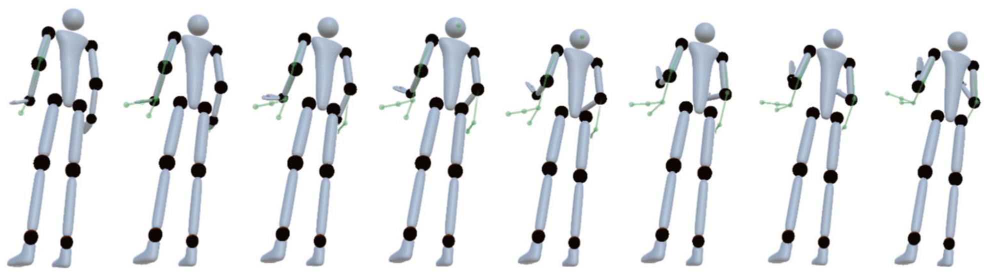


Fig. 12 Image sequence, taken from a video of a biceps curl exercise with deviation. The viewpoint is optimized by the algorithm by Kwon et al. [19]. Originally published in [2]

physiotherapy with the lowest frequency. Color vision deficiency did not affect our user study because the tasks required participants to recognize shapes rather than colors, as our focus was on perspective.

Experimental Setup for Exercise Recording

The poses and motions found in this work were recorded using a Microsoft Azure Kinect 3D camera [26]. Its computer vision capabilities can provide spatial coordinates of several joints of the human body. Here, the term *joint* is rather defined as biological landmarks than referring to the medical definition [26].

The following conditions achieved optimal positioning of the subject in our case: The camera was mounted at about 140 cm with the help of a tripod. It was placed about 280 cm away from the subject. The height of the subject was about 190 cm. This setup allowed for stable tracking and captured all poses within the frame. According to our use case, the joints of the eyes, ears, and nose were discarded as we found that these were too imprecise and irrelevant to our use case. This left us with 26 joints. For further information on the visualization of avatars, see Sect. "Exercise Visualization".

Subsequently, a diverse set of example exercises was developed including various exercise and deviation combinations. We then superimposed each of these exercises to the corresponding counterpart with deviation from the correct form (see Sect. "Sample Exercises"). The spatial registration was unambiguous, as the majority of joints were nearly identical and both executions were recorded from the same individual. Consequently, many registration methods would yield similar results. For further information regarding registration, covering more complex cases, see Sect. "Registration". The methods used to create an overlay of two exercises matching temporally exceed the scope of this paper. Throughout the literature we often see *Dynamic Time Warping* fulfilling that role (e.g., [27, 28], and [29]).

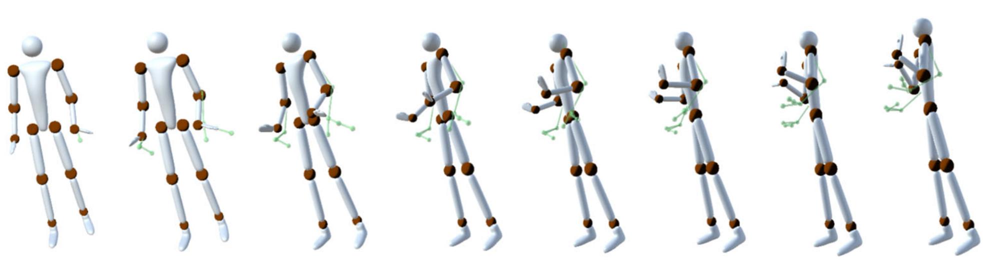


Fig. 13 Image sequence, taken from a video of a biceps curl exercise with deviation. The viewpoint is optimized by our algorithm. Originally published in [2]

Table 2 Results of user study. Distribution of viewpoint selection methods chosen by the participants. Originally published in [2]

Method	Bench press	Biceps curl A	Lateral raises	Shoulder press	Bend over row	Biceps curl B	Total Percentage
Neutral	19	15	3	15	6	18	32.48 %
JMO	1	1	6	0	25	2	14.96 %
Kwon	14	7	3	3	6	8	17.52 %
Ours	5	16	27	21	2	11	**35.04 %**

Exercise Visualization

An abstract avatar was used to visualize the actual motion, and for the target motion, a skeleton is displayed as seen in Fig. 9. The visualization of the skeleton displayed in green corresponds to the joints recorded by the 3D camera [26] as mentioned in Sect. "Experimental Setup for Exercise Recording". Two different avatar visualizations were used to help users distinguish the actual movement from the target movement. In addition, users with color vision deficiency are supported, as the differences between the avatars are distinguished by shape, not by color. The abstract avatar occludes itself and its background to a greater extent and visualizes fewer joint positions than the skeleton, as the fingertips and thumbs are integrated into the hand. Yet, for viewpoint optimization, all joints are considered in the calculations. The visualizations in this paper are simply used for demonstrative purposes and are not the subject of our research. The focus of this work is the viewpoint selection, where the form of visualization plays a subordinate role.

Sample Exercises

To assess our method, as described in Sect. "User Study Design", and compare it with those described in existing literature, we selected four static poses to establish basic assumptions (see Sect. "Perspective Considerations") and six dynamic exercises, each with specific deviations from the ideal form. The deviations were chosen to be common mistakes for the considered exercises. We aimed to select a wide variety of exercises and deviations to evaluate the methods exhaustively. As a result, poses and exercises were selected, so that different movement and feedback directions are represented. For instance, during lateral raises, the arms are moved laterally away from the body, whereas in a biceps curl, the arms move in front of the body (see Fig. 9). We also included an exercise with different deviations (biceps curls A and B).

Selecting a viewpoint for videos can be considered as choosing a continuous viewpoint for each static pose in the individual frames. To verify the underlying assumptions of viewpoint quality (see Sect. "Perspective Considerations"), we chose four representative static poses: standing (standard anatomical position), squatting, bending down, and bench press. To learn more about how the user study was conducted, please refer to Sect. "User Study Design".

The following six exercises were chosen, including deviations (see Fig. 9 for visualization of the exercises): bench press (deviation: Arms too wide), lateral raises (deviation: Arms asymmetrical), bend over row (deviation: Elbows tucked in), shoulder press (deviation: Arms asymmetrical), biceps curl A (deviation: Repetition only half executed), and biceps curl B (deviation: Elbows do not stay stable). The exercises and their deviations were recorded at the same position, and performed by the same individual. Therefore, it was possible to use the absolute position as a registration method. However, the commendable methods described in Sect. "Registration" would yield the same results, even when performed with different-sized individuals at varying locations.

User Study Design

The user study was structured into three tasks. The participants were presented with two tasks and a set of structured questions. These are explained in the following.

Viewpoint Selection: We intended to confirm the underlying assumptions of user preferences for the views as explained in Sect. "Perspective Considerations". For this purpose, we asked users to choose the viewpoint for static poses. Feedback was not displayed during this task, as we wanted to verify the established assumptions. As viewpoint selection for videos selects a viewpoint for a static pose in each frame, this should give us insights into user preference and how our algorithm performs compared to that. Furthermore, it is unfeasible for users to select a camera path in real-time. Therefore, only with static poses, user evaluation is even possible. This also enables our method to be compared to the current literature (see Sect. "View Direction").

A skeleton-like avatar successively showed four static poses of exercises: bench press, squat, bend over row, and standing (for more information, see Sect. "Sample

Exercises”). The users could adjust the viewing angle pose-wise by clicking and dragging the mouse. A skybox around the avatar supported orientation in virtual 3D. After confirming, the viewpoint was registered and stored for analysis.

Viewpoint Comparison: To evaluate the performance of our view selection algorithm, we showed four looped videos of exercise repetitions with the corresponding correction feedback randomly juxtaposed. The viewpoint in each video was selected by a different method. This way, six different exercises with deviations were successively shown, as explained in Sect. “Sample Exercises”.

The different methods used for viewpoint selection are:

- Ishara et al. [17], who chose the viewpoint according to the *JMO*, the biggest sum of angles between all joints and the potential viewpoint (see Fig. 2).
- The method of Kwon et al. [19] involves the sum of displayed limb lengths, a 2D, and 3D bounding box. As their best resulting method is computationally intensive and not capable of real-time, we chose their second-best algorithm variant without weights.
- Our algorithm, as described in Sect. “Viewpoint Calculation”.
- To compare the methods to a neutral position, we included a viewpoint as it is used in *isometric projection* (rotated 45° horizontally and 35.264° vertically).

For more detail on the methods mentioned in this section, see Sect. “View Direction”.

Questionnaire: Finally, the third task asked participants to provide more details about their prior experience with the topic and to share their opinions. The first four questions were answered using a Likert scale, while the last two were answered with free-text responses:

- How often do you exercise?
- How often are you involved in strength training?
- How often do you receive physiotherapy?
- How often do you consider movements?
- What options would you have liked to see?
- What stood out to you?

Viewpoint Benchmark

The viewpoints, chosen in the viewpoint selection task of the user study, were evaluated using the benchmark presented by Dutagaci et al. [30]. They provided a method to evaluate potential viewpoints and compare them to a selection of views chosen by users. The calculation of what Dutagaci et al. call the *View Selection Error* (VSE) can be found in Eq. 2. The VSE represents a number between 0 and 1, where low values signify a high discrepancy between the viewpoints in question and the ones chosen by the users.

$$VSE = \frac{1}{M \cdot \pi \cdot r} \sum_{m=1}^{M} GD_m \tag{2}$$

In Eq. 2, GD_m represents the geodesic distance of the potential viewpoint to each user-chosen viewpoint $m \in M$. The variable M represents the total number of participants (i.e., the number of viewpoints to consider). The viewpoints are expected to be on a sphere (viewpoint sphere) around the focused object. The radius of said sphere (i.e. the distance of each viewpoint to the focused object) is represented by r. To visualize the user-selected viewpoints, the viewpoint vectors were projected on the median and transverse planes. Subsequently, we plotted the VSE by comparing each direction around the center as a potential viewpoint. As a result, the *View Selection Error* is displayed angle-wise in the median and frontal plane around the body using the Viridis colormap [31] in Fig. 10. Here, blue areas represent a low VSE and therefore, an overall low distance to the user-selected view directions. In contrast, view directions that were avoided by the participants are shown by yellow areas.

Results

In Sect. “Viewpoint Selection”, we will discuss the performance of each algorithm’s viewpoint selection relative to the user-selected viewpoints, using the method described in Sect. “Viewpoint Benchmark”. Subsequently, in Sect. “Method Analysis” the performance of the above-mentioned methods in optimizing viewpoints for the same exercise will be discussed, based on image sequences extracted from the videos. Lastly, Sect. ’Viewpoint Comparison” concludes the user study results of the viewpoint comparison. The results of the questionnaire are found in Sect. User Study Design”, where they specify the participants, and in Sect. 6, where the free-text answers are discussed.

Viewpoint Selection

In Fig. 10, a low view selection error is represented by blue areas. Therefore, viewpoints in these areas aligned well with the user selection. Yellow areas were chosen less. The red line represents the viewpoint chosen by our method for the static pose without movement. Looking at Fig. 10, we see that our method calculated viewpoints predominantly lying in the blue regions, i.e., in regions preferred by users. Similarly, it becomes apparent when analyzing the view selection error mean over the four exercises, that, in comparison, our algorithm fits the selection of the users best with an average VSE of 0.3467. The static isometric-like view performed second best with an average VSE of 0.347 followed by JMO with 0.4825 and the method of Kwon et al. with 0.5497.

Method Analysis

To interpret the comparison of methods in Sect. “Viewpoint Comparison”, it is essential to understand the viewpoints each method provides and how these viewpoints change over time.

JMO [17]**:** The JMO algorithm predominantly produced an adequate overview of the human body. A major drawback was that, when applied to videos, the algorithm erratically switched between viewpoints that were significantly distant from each other. This can be perceived in Fig. 11. Consequently, the feedback often was difficult to comprehend, as the algorithm was not designed to work with visual cues or videos. Additionally, several viewpoints were selected from below and behind, although participants preferred perspectives from the front and slightly above (see Sect. “Viewpoint Selection”).

Kwon et al. [19]**:** As can be seen in Fig. 12, the algorithm of Kwon et al. seemed to predominantly produce views from behind in our application. As elaborated in Sect. “Viewpoint Selection”, this is a view that is mostly avoided by users. Additionally, the algorithm sometimes selected views from below, similar to the JMO algorithm mentioned earlier. The algorithm by Kwon et al. offered a much more stable perspective than JMO. However, the feedback was often challenging to see.

Ours: Our algorithm consistently transitioned between an optimal viewpoint for the neutral position and the contracted position with deviations, as illustrated in Fig. 13. When feedback was present, it was displayed clearly and with perceivable emphasis on it. However, in some exercises, the rapid exercise execution caused a conflict between the neutral and the feedback-optimized viewpoint. This resulted in quick camera movements, which some users found irritating.

Viewpoint Comparison

Table 2 shows the user choice distribution of the viewpoint comparison. Our algorithm was most prevalent with 35.04 % of votes, the isometric-like position was chosen second most with 32.48 %, followed by Kwon et al. [19] with 17.52 % and lastly JMO [17] with 14.96 %.

Occasionally camera positions from behind were provided by the methods of Kwon et al. [19] and Ishara et al. [17]. Additionally, they produced an unsteady camera movement, because they jumped between far-distant viewpoints and generally had just a limited amount of viewpoints available. In contrast, the static isometric-like viewpoint produced surprisingly good results, although it lacked an adaptation for movement or feedback. The primary advantage of the isometric-like viewpoint over the other methods was its stability. Our method provided an adequate view of the neutral positions of the exercises. Furthermore, it produced a continuous camera movement toward a feedback-oriented viewpoint with increasing deviation. However, the camera movement showing the bench press and bend-over row exercises was occasionally rapid.

Insights and Discussion

Looking at Fig. 10, it becomes evident that the participants preferred a frontal view direction. This aligns with the statement made by Zusne [24], that humans desire frontal views (see Sect. “Perspective Considerations”), and verifies these requirements for our application. Additionally, it was observed that participants preferred a viewpoint from slightly above.

Our algorithm performs significantly less well for some specific exercises. This can be attributed to the consistently smooth, though occasionally rapid, camera movement. In particular, the camera moved rapidly during the bench press and bend-over row exercises. As stated in Sect. “Viewpoint Calculation”, our algorithm generally prevents inconsistent camera movement, though rapid camera motions may still occasionally occur.

The most prevalent statement made by the participants regarded the camera movement consistency. Specifically, users were irritated by movements that were too rapid or erratic. This observation matches the findings by Assa et al. [22] concerning camera paths. Additionally, many participants indicated that having multiple camera perspectives would help them to understand the poses and feedback. This is especially interesting for future work and when applying suggested methods. Additionally, some users desired the option to select no method, as they felt none of the suggested perspectives were adequate. This implies that there are possible improvements to our algorithm and that human viewpoint preferences might need further assessment. Lastly, some users struggled to interpret poses without relation to the environment. Primarily, this concerned the bench press exercise, where a virtual bench might help users interpret the avatar posture. Therefore, incorporating the surrounding environment could enhance understanding, especially for exercises involving equipment such as weights, benches, and pull-up bars. However, additionally rendered equipment could occlude the avatar or visual cues and therefore hinder the perception of the provided feedback. In conclusion, stable frontal views satisfied users the most. Regarding viewpoint preference, we could not identify any difference regarding gender or age.

The spatial registration (see Sect. "Registration") for our use case was trivial, since the superimposed exercises were recorded at the same position with the same individual. Other circumstances, like varying individuals or different registration methods, can yield fundamentally different results in terms of feedback appearance. However, the view selection methods, as presented in Sect. "Viewpoint Calculation", would still find a valid viewpoint. Depending on what registration methods were chosen, the view selection could be skewed toward the feedback deviation. If there are other registration methods chosen, it might be necessary to adapt the constants in the viewpoint selection calculation (in particular, w and δ_0 in Eq. 1) to retrieve the desired perspectives.

Conclusion

The paper at hand provides novel insights on how to optimize the display of superimposed avatars. As we can see in current literature, the superimposition of avatars plays an increasingly important role. As an accessible and intuitive method of providing and receiving motor feedback, it is widespread in both mixed reality and traditional feedback technologies.

The consideration of avatar registration is inevitable when attempting to optimize the display of superimposed avatars. In the literature, avatars are often registered by aligning the position and/or rotation of a single joint. For specific use cases, this can be adequate. However, to ensure that users can easily understand a wide range of exercises, a more detailed approach must be taken. We offer valuable insights on how certain exercises could be optimally registered based on the performed exercise. Without the claim for completeness, we offer fundamental guidelines as the basis for application development or further research. Furthermore, we provide concrete examples to help with comprehension and potential implementation. While we deliver a fundamental framework for avatar registration, specific use cases have to be further explored.

Avatar registration, important by itself, additionally represents a major factor of influence on viewpoint selection, another essential topic for optimal display of superimposed avatars. We introduce a new method for selecting viewpoints for motor feedback, such as superimposed avatars, among other options. Not only is this method computationally faster than the methods found in the literature, but evaluation in the context of a user study showed that participants preferred our method over other methods found in the literature. Nevertheless, there still is a lot of potential for further research regarding viewpoint selection, as our method seems better, but not optimal.

Our registration and viewpoint selection methods combined can adequately optimize the display of superimposed avatars. However, the individual methods provide value as well, as they can be utilized separately from each other.

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.

Author Contributions All authors significantly contributed to this work. Florian Diller and Alexander Wiebel conceptualized the work, conducted the user study, and interpreted the data. Florian Diller and Thorben Frey carried out the implementation. Florian Diller and Alexander Wiebel prepared this manuscript. Florian Diller, Alexander Wiebel, and Gerik Scheuermann reviewed the manuscript, ensured its integrity, and approved it for publication.

Funding Open Access funding enabled and organized by Projekt DEAL. 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.

Data Availability Not applicable.

Declarations

Competing of Interest The authors declare that they have no conflict of interest.

Research Involving Human and/or Animals Not applicable.

Informed Consent Informed consent was obtained from all individual participants included in the study. There was no personal data gathered.

References

1. Diller F, Scheuermann G, Wiebel A. Visual cue based corrective feedback for motor skill training in mixed reality: a survey. IEEE Trans Vis Comput Graph. 2022;30:1–14. https://doi.org/10.1109/TVCG.2022.3227999.
2. Diller F, Wiebel A, Scheuermann G. Automatic viewpoint selection for interactive motor feedback using principal component analysis, 2024;350–361. https://doi.org/10.5220/0012308700003660 . INSTICC
3. Tam GK, Cheng Z-Q, Lai Y-K, Langbein FC, Liu Y, Marshall D, Martin RR, Sun X-F, Rosin PL. Registration of 3d point clouds and meshes: a survey from rigid to nonrigid. IEEE Trans Visualization Comput Graph. 2012;19(7):1199–217. https://doi.org/10.1109/TVCG.2012.310.
4. Yaniv Z. Rigid registration. Image-Guided Interventions: Technology Review and Clinical Applications, 2008;159–192 https://doi.org/10.1146/annurev-bioeng-070909-105249
5. Bellekens B, Spruyt V, Berkvens R, Weyn M. A survey of rigid 3d pointcloud registration algorithms. In: AMBIENT 2014: the Fourth International Conference on Ambient Computing, Applications, Services and Technologies, August 24-28, 2014, Rome, Italy, 2014;8–13.
6. Hoang TN, 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
7. 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
8. Le Naour T, Hamon L, Bresciani J-P. Superimposing 3d virtual self + expert modeling for motor learning: Application to the throw in american football. Frontiers in ICT 2019;6. https://doi.org/10.3389/fict.2019.00016
9. Bouwmans T, Javed S, Zhang H, Lin Z, Otazo R. On the applications of robust pca in image and video processing. Proc IEEE. 2018;106(8):1427–57. https://doi.org/10.1109/JPROC.2018.2853589.
10. Skaro J, Hazelwood SJ, Klisch SM. Knee angles after crosstalk correction with principal component analysis in Gait and cycling. J Biomech Eng. 2021;143(5): 054501. https://doi.org/10.1115/1.4049809.
11. Bonaventura X, Feixas M, Sbert M, Chuang L, Wallraven C. A survey of viewpoint selection methods for polygonal models. Entropy. 2018;20(5):370. https://doi.org/10.3390/e20050370.
12. Rudoy D, Zelnik-Manor L. Viewpoint Selection for Human Actions. Int J Comput Vis. 2011;97(3):243–54. https://doi.org/10.1007/s11263-011-0484-5.
13. Kiciroglu S, Rhodin H, Sinha SN, Salzmann M, Fua P. ActiveMoCap: Optimized Viewpoint Selection for Active Human Motion Capture. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2020 (2020) https://doi.org/10.1109/CVPR42600.2020.00018
14. Shi Z, Yu L, El-Latif A, Le D, Niu X. A kinematics significance based skeleton map for rapid viewpoint selection. Int J Digital Content Technol Appl. 2012;6(1):31–40.
15. Wang M, Guo S, Liao M, He D, Chang J, Zhang J. Action snapshot with single pose and viewpoint. Vis Comput. 2019;35(4):507–20. https://doi.org/10.1007/s00371-018-1479-9.
16. Choi MG, Yang K, Igarashi T, Mitani J, Lee J. Retrieval and visualization of human motion data via stick figures. Comput Graph Forum. 2012;31(7):2057–65. https://doi.org/10.1111/j.1467-8659.2012.03198.x.
17. Ishara K, Lee I, Brinkworth R. Mobile Robotic Active View Planning for Physiotherapy and Physical Exercise Guidance. 2015 IEEE 7th International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), 2015;130–136. https://doi.org/10.1109/iccis.2015.7274609
18. Charalambides CA. Enumerative Combinatorics. Discrete Mathematics and Its Applications, p. 62. Taylor & Francis, Oxfordshire (2002). https://books.google.de/books?id=PDMGA-v5G54C
19. Kwon B, Huh J, Lee K, Lee S. Optimal camera point selection toward the most preferable view of 3-D human pose. IEEE Trans Syst Man Cybernet. 2020;52(1):533–53. https://doi.org/10.1109/tsmc.2020.3004338.
20. Nundy S, Lotto B, Coppola D, Shimpi A, Purves D. Why are angles misperceived? Proc Natl Acad Sci USA. 2000. https://doi.org/10.1073/pnas.97.10.5592.
21. Sorzano COS, Vargas J, Pascual-Montano A. A survey of dimensionality reduction techniques. ArXiv14032877 Cs Q-Bio Stat (2014) https://doi.org/10.48550/arXiv.1403.2877
22. Assa J, Cohen-Or D, Yeh I-C, Lee T-Y. Motion overview of human actions. ACM Trans Graph. 2008;27(5):1–10. https://doi.org/10.1145/1409060.1409068.
23. Polonsky O, Patané G, Biasotti S, Gotsman C, Spagnuolo M. What's in an Image? Vis Comput. 2005. https://doi.org/10.1007/s00371-005-0326-y.
24. Zusne L. Visual perception of form. New York: Academic Press; 1970.
25. Borisenko AI, Tarapov IE. Vector and tensor analysis with applications. New York: Dover Publications Inc.; 1979. p. 109.
26. Microsoft Development Team: Azure Kinect DK documentation. https://learn.microsoft.com/en-us/azure/kinect-dk/, Accessed: 2023-3-30 (2018)
27. Su C-J. Personal Rehabilitation Exercise Assistant with Kinect and Dynamic Time Warping. Int J Inform Educ Technol 2013;448–454. https://doi.org/10.7763/ijiet.2013.v3.316
28. Antón D, Goñi A, Illarramendi A. Exercise recognition for Kinect-based Telerehabilitation. Methods Inform Med. 2015;54(02):145–55. https://doi.org/10.3414/me13-01-0109.
29. Saenz-de-Urturi Z, Soto BG-Z. Kinect-based virtual game for the elderly that detects incorrect body postures in real time. Sensors. 2016;16(5):704. https://doi.org/10.3390/s16050704.
30. Dutagaci H, Cheung CP, Godil A. A benchmark for best view selection of 3D objects. Proceedings of the ACM workshop on 3D object retrieval - 3DOR '10, 45–50 (2010) https://doi.org/10.1145/1877808.1877819
31. Nuñez JR, Anderton CR, Renslow RS. Optimizing colormaps with consideration for color vision deficiency to enable accurate interpretation of scientific data. PLOS ONE. 2018;13(7):1–14. https://doi.org/10.1371/journal.pone.0199239.

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