DEPARTMENT OF INFORMATICS

TECHNICAL UNIVERSITY MUNICH

Bachelor’s Thesis in Informatics: Games Engineering

Learning Assistance for Sports Activities and Physical Therapy based on Body Tracking

Daniel Jazz Young

DEPARTMENT OF INFORMATICS

TECHNICAL UNIVERSITY MUNICH

Bachelor's Thesis in Informatics: Games Engineering

Learning Assistance for Sports Activities and Physical Therapy based on Body Tracking

Lernassistenz für Sportliche Aktivitäten und Physiotherapie mithilfe von Körpertracking

Author:	Daniel Jazz Young
Supervisor:	Gudrun Klinker
Advisor:	Sandro Weber
Submission Date:	11/15/2019

I hereby confirm that this bachelor's thesis in Informatics: Games Engineering is my own work and I have documented all sources and material used.

Munich, 15.11.2019 Daniel Jazz Young

Acknowledgments

I would like to thank the Technical University Munich for providing me with an outstanding education and Prof. Gudrun Klinker and Sandro Weber for accepting my thesis and providing aid when needed. I would also like to show gratitude towards my father, whose continuous support and guidance has helped me persevere through many arduous situations.

Abstract

Exercising has been proven to grant mental and physical health benefits, but in practice, these can only be reaped by performing the movement in a safe and healthy manner. The purpose of this thesis is to discuss the design and results of a user study in which the participants have been given different learning paraphernalia - one of which is an application that simulates exercises with the help of a three-dimensional avatar. By evaluating the data gained from the user study, it is possible to determine the value and efficiency of a three-dimensional learning tool. In the end, it became clear that the application has its advantages and disadvantages but has shown itself to be a valuable learning tool due to its flexibility and popularity.

Table Of Contents

Acknowledgments

Abstract

1 Introduction

A software engineer's onus is not simply developing or maintaining a system, but striving for optimization and change. An open and adaptable mindset allows developers to branch out into other fields of study and apply their problem-solving mechanisms there. The nature of said optimizations can range from performance upgrades to increased precision of acquired and transmitted data or quality of life enhancements.

Digitalization has increased the necessity and pushed the boundaries of specialized software. Increasing demand calls for increasing supply and diversity. One frontier that has been significantly expanded upon during the last decade is that of three-dimensional (3D) applications. These applications have their own set of alterable rules and dimensions, allowing the simulation and modification of digital objects that can be fictitious or based on real-world objects or entities.

Architectural students, for example, use 3D simulation to measure and design structures [Cla+02, Che+09]. Medical branches model organs [Suz+98] or teeth [Jod+14] in an attempt to plan procedures and military institutions prepare trainees with serious games, there referred to as real-world scenarios, such as live firing simulations [Bha+16].

These applications are practical and cheap in comparison, making them an efficient, effective and reversible solution.

The sports field has delved into 3D applications as well - for instance by using virtual reality to enhance a player's performance through predictive assistance [Kat+06].

This research paper discusses the usefulness of another 3D tool for sports education and physical therapy by conducting a user study in which the participants of varying experience levels compare the application with conventional instruction methods.

The application allows a user to view recorded training animations performed by a 3D avatar to develop an understanding of specific exercise movements. Instructors can also choose to add more animations by equipping trackers and recording the movement by themselves.For comprehensibility, the paper is divided into six parts. First, it discusses the theory behind the studies' core aspects, then introduces the utilized hardware and software. Afterward, it describes the user studies' concept and design before moving on and going into the details of the programs' implementation. The next part expounds the practical portion of the user study and finally evaluates the results. To conclude the bachelor's thesis, possible future enhancements and alternative applications are explored.

This bachelor's thesis' goal is to determine if the 3D application qualifies as a substitute or upgrade to traditional learning media for sports activities and physical therapy.

2 Theory

A user research or user study “focuses on understanding user behaviors, needs, and motivations through observation techniques, task analysis, and other feedback methodologies” [Usa19]. Such a study is indispensible for this thesis' goal, as any tool which is developed for humans needs to have user input to gauge its worth.
In order to determine the optimal execution of the user study, the theory behind the core aspects has to be established first. Broadly speaking, it can be broken down into three large topics: the virtual abstraction of real-world objects, the concept of visual learning and that of three-dimensional movement analysis.

2.1 Virtual Abstraction

From a computer scientist's perspective, anything can be represented as some form of data. The next step lies in assessing what exactly has to be expressed and in what manner the object or phenomenon should be depicted. This depends on which properties are relevant for the abstraction's final purpose.

One option is keeping the data as a strictly binary representation. This means expressing an object in terms of zeros and ones, which is the way that computers perform calculations or store data at the basic level. Another higher-level option is to express the data as structures such as arrays and vectors or constructing objects in object-oriented programming. By doing so, these structures are used as containers to organize data or digitally associate an object with specific parameters.
If needed, the abstraction can appeal directly to a human's senses, such as transforming data into wavelengths and playing them through speakers or providing haptic feedback [Sal+95].
By using a medium with a display, individual pixels can be altered to illustrate graphs and images. A common example is the Utah Teapot (Figure 2.1), a 3D model usually referenced in computer graphics. It is a benchmark created by researcher Martin Newell in 1975 which revolutionized 3D modeling [Tor06]. For the sake of

this thesis, visual abstraction provides the most benefit on account of visual learning being a primary subject of the user study.

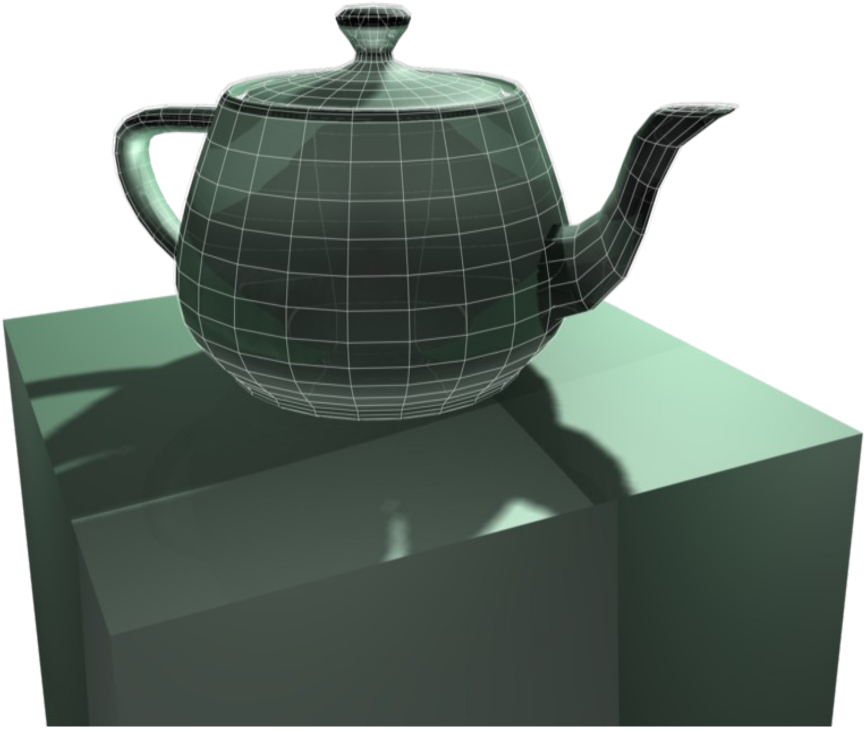

Figure 2.1: The 3D model of the Utah teapot. Hay Kranen/CC BY-SA 3.0

2.2 Visual Learning

Research has shown that humans naturally possess the capacity to learn new skills and ideas through observation. This behavior shows at a young age, where toddlers slowly grasp the concept of locomotion and control of their bodies. Learning through replication and repetition can benefit neural development and social competence. This is explored further by Albert Bandura in his take on social cognitive theory, in which he explains that humans tend to replicate the actions and intentions of others through observation during the learning process [Ban04].

Even though adults possess much greater learning capabilities regarding abstract and theoretical concepts, the instinct of observational learning is still present.

Visualization can help with demonstrating theoretical notions such as height, weight, movement or interaction between objects. This is of great significance for this thesis, since sport in particular involves movements or exercises that can be explained verbally, but are much harder, and potentially more dangerous, to attempt in practice without visual examples or professional training [Ham94].

2.3 Three-dimensional Movement Analysis

The crux of this bachelor's thesis' user study. In sports, bar hands-on personal training, the most commonly used teaching methods consist of images and videos. These traditional media, however, come with significant drawbacks.
A set of images for a particular exercise usually consists of one image depicting the initial position and another showing the final position. Optionally, this set can be accompanied by an explanatory caption or text. The problem with still images is that they can not properly show the transitional movement from point A to point B, especially if the movement is complex and works multiple muscle groups.

Videos prove to be a good way of alleviating this issue, allowing a viewer to follow each step of a repetition. They are also easy and cheap to produce, provided the creator forego any additional editing. The limiting flaw of video material comes from the direction from which it is recorded. Every part of the movement is visible from one perspective at a given point in time, meaning a viewer can only see the exercise from a fixed angle. Rotating the camera around during the movement is not a solution, as the changed angle still doesn't allow the viewer to see the exercise from the same perspective at a prior point in time.

On the basis thereof, a 3D program is developed for this user study to determine if it is a suitable alternative or improvement to said methods. The program in question allows the recording of exercises which are then simulated by a virtual avatar, allowing a viewer to replay the movement indefinitely and rotate the camera around

360 degrees to see every part of the exercise during every timestep.

Prior to the 3D program's implementation, the method of achieving a precise replication of humanoid movement has to be determined. Other research in this field deals with a similar issue. For tracking, there are generally two options to consider, a marker-based tracking or a markerless approach.

Markers can be simple identifiers which a computer can, with the support of a camera, recognize and trace. They can also be pieces of hardware with sensors that are affixed to a body or object which communicate with a processing unit.

Markerless tracking is achieved by using a camera that recognizes the movements of objects and entities without any identifying marks and applies them to virtual models. Since the body moves in three-dimensional space, a standard camera is not feasible, as it cannot accurately determine the position in space. Multiple cameras can capture an object from different angles, providing a room-scale depth perception. The drawbacks are high installation and maintenance costs, greater space requirements and slower processing speeds [Shi+14].

Instead, depth-sensing cameras are the preferred solution. Depth-sensing cameras project beams of infrared light (IR) that collide with an object, creating a pattern. The Kinect[1], for example, measures the time-of-flight to create a 3D image that changes with the object's movement. Researchers have used the depth-sensing capability of the Kinect to great effect, using it for modeling and motion capture [Shi+14, Yoo+13, Zen+12]. A common problem seen in the studies' results are the movement jitters from noise and the lack of occlusion handling, which are not prevalent issues while using a 360-degree marker-based tracking method.

By virtue of this user studies' results depending on the replicated movements looking as close to the original as possible, a technique with higher data precision has to be chosen. Even though the markerless approach requires a less complicated setup, it struggles with smoothly following movements, especially complicated ones. Hence, a marker- or tracker based motion capture is the most obvious approach.

---

[1] https://developer.microsoft.com/en-us/windows/kinect

3 Fundamentals

The following section introduces the special hardware and software used in this thesis and describes their attributes and uses. For the sake of this chapter's listings, standard equipment is disregarded.

3.1 Hardware

3.1.1 HTC Vive

The HTC Vive is a virtual reality system consisting of an HTC Vive headset, two HTC Vive controllers and two HTC Vive base stations by default as seen in figure 3.1. The headset uses one organic light-emitting diode (OLED) panel per eye, each with a display resolution of 1080x1200 pixels. Its outer shell contains sensors which detect the base stations' infrared pulses at 60 pulses per second. The headset also includes an accelerometer, a gyroscope and a proximity sensor [Htc19].

Figure 3.1: Basic components of the HTC Vive setup.

The HTC Vive controllers are wireless input accessories with a rechargeable battery and 24 sensors. Unlike the HTC Vive headset, the controllers do not include a gyroscope or accelerometer. Their velocity is computed by analyzing the difference in their position at timestep one and their position at timestep two. The Steam VR tracking system can determine the positions with sub-millimeter precision.

Lastly, the HTC Vive base stations, also known as the Lighthouse tracking system, create a 360-degree virtual space with a radius of up to 15 square feet. The base stations synch up and communicate wirelessly. Each station emits infrared light pulses at 60 pulses per second to help determine each accessories' position.

The HTC Vive components are primarily used for their tracking capabilities [Nie+17] because the individual trackers do not generate large amounts of noise due to having separate components assisting with hardware tracking. Noise is a term used for unwanted data modification during signal processing [Tuz02].

3.1.2 HTC Vive Tracking Devices

The HTC Vive trackers are a wearable motion tracking accessory seen in figure 3.2 and essential for the realization of this thesis. The tracker's shell contains sensors that communicate with a connector, which in turn is plugged into the computer to help determine the tracker's position.

Other third-party accessories developed for the HTC Vive trackers are belts and straps to affix each tracking device to an object or limb. The user study requires seven tracking devices for optimal positional and rotational tracking of the limbs.

Figure 3.2: Two HTC Vive tracking devices.

3.2 Software

3.2.1 Unity

Unity is a game engine developed by Unity Technologies. It can be used to create two-dimensional, three-dimensional, virtual reality and augmented reality applications and simulations, therefore being a prominent choice for the implementation of the application used for the user study. The engine supports multiple platforms, but for the sake of this thesis, only running the program in the Unity editor is sufficient.

The engine's primary scripting API is C Sharp (C#), a general-purpose, object oriented-programming language. Visualization and physics simulations are primarily taken care of by the engine, decreasing the workload and allowing the focus to be shifted towards modeling, tracking and governing data.

3.2.2 Steam VR

Steam VR is an extension for Steam, a software distribution platform. It facilitates the use of virtual reality accessories in Steam games or other applications, such as Unity, serving as an interface between software and virtual reality hardware.

3.2.3 Final IK

Final IK is an inverse kinematic solution. By only using forward kinematics, which is achieved with the HTC Vive trackers, the individual body parts are tracked, but more subtle muscle movement such as that of the shoulder blades is not possible. Inverse kinematics solutions act as an additional stabilizer and calculate body movement based on the position and rotation of the limbs. This makes the movement look smoother and more akin to a human's.

3.2.4 Microsoft Visual Studio

The integrated development environment (IDE) of choice. This is used to edit and manage the C# scripts used in the application. Its code editor supports code refactoring and IntelliSense, a component that assists with automatic code completion and error detection.

3.2.5 Autodesk Maya

Autodesk Maya is a 3D computer graphics application, mainly used to model assets for 3D applications, architecture or animated films. Maya is used to create 3D assets such as the avatar for the user study. It also fully supports texturing and rigging, a delicate process for visualization and performance.

4 Concept

Now that all the necessary tools have been chosen, the thesis' practical portion has to be developed. To establish a better structure, this section is divided into two parts, the 3D application and the user study itself. By defining the requirements and the methods of fulfilling those requirements, a rough blueprint can be created, which functions as an outline for the application and the survey.

4.1 Application

4.1.1 Idea

Firstly, the application requires a way of representing a human body in virtual space. Since the program is a learning tool, the avatar should look humanoid and have proper proportions. Secondly, said avatar should have select bones that are mapped to the movements of the person performing the exercise. Once that is possible, the event of starting and stopping a recording must be manageable. Also, only relevant information should be saved in a recording, so an adequate file format must be chosen. That covers the requirements for data acquisition, which leaves those of data conversion and movement replication. Since the method of simulating an exercise depends heavily on which file format the data is converted to, the two tasks should not be designed arbitrarily, but in tandem. If possible, the file format should be flexible and not require too many calculations at run-time.

4.1.2 Design

The avatar is fully created in Autodesk Maya, including rigging and skinning. It is modeled with a standard height of 1.80m, which can be slightly adjusted if needed. Steam VR manages the trackers' location in virtual space [Pee+18], and the Final IK solution is functional with only three tracking devices, one attached to the head and each hand.

To reduce computational error and improve body and limb positions, more trackers are added to key points on the user's body. As seen in Figure 4.1, the avatar is mapped to ten tracked body parts - the head, elbows, hands, pelvis, knees, and feet. A user interface (UI) allows a user to start and stop recording, at which point the resulting file can be converted or discarded.

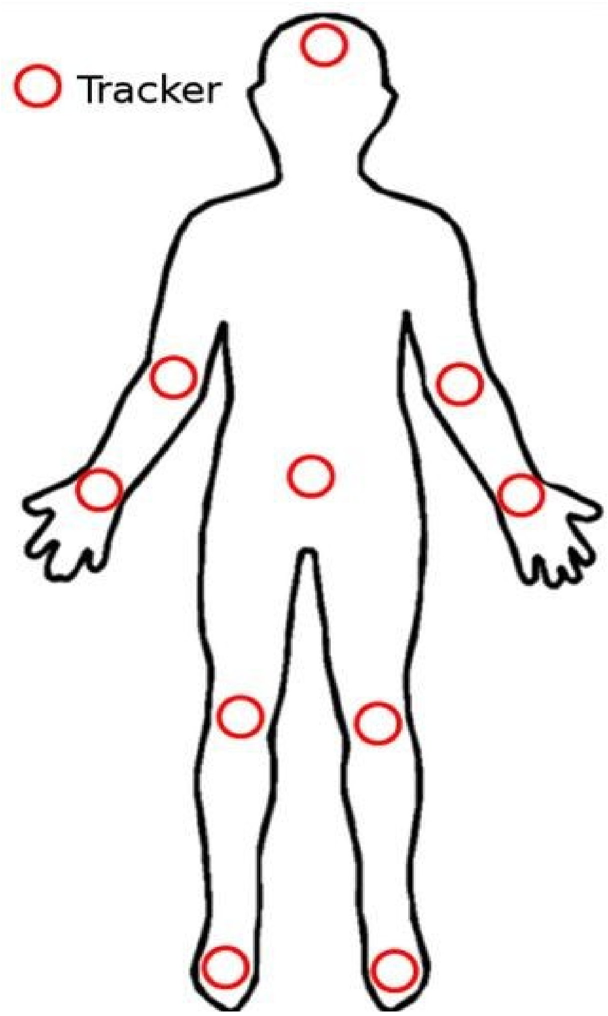

Figure 4.1: Adapted human outline showing approximate positions of tracking devices.

The file conversion could take place immediately after recording, but after careful deliberation, it was determined that it is better to be done in a separate step, as not every recorded movement may be satisfactory. Also, replaying the exercise should be delegated to a second avatar, because the first is associated with the trackers, which would override the recording otherwise. Since many trackers must be simulated at once at every frame, the most efficient way to convert the recorded data would be as an animation with 50 keyframes per second.

4.2 User Study

4.2.1 Idea

In order to acquire relevant information, the user study should include participants of varying ages and levels of sports experience in the best case. Furthermore, there should be an acceptable amount of participants to draw a conclusion. To ensure that focus and motivation are high, the preparation must not take long and the questions should be straightforward and easy to answer.

The study should be conducted in a safe environment [Bia12] and structured in a way that is simple and provides the most benefit towards drawing the conclusion on whether the application is useful or not.

In addition to the information gained from the initial questioning, the study should provide ample opportunities to document observational data derived from the participants' performance. To that end, there must be a scoring system to evaluate learning progress along with a reasonable time in which said learning progress is monitored.

4.2.2 Design

In consideration of this thesis' timeframe, a test group of around ten people is sufficient. As the primary concern is the evaluation of learning tools and not the recording of additional exercises, the preparation is kept simple. The study is conducted in a secluded area, free from other outside influences. The participants are chosen at random to avoid personal bias and after an initial questioning are given the first of three learning tools.

After a brief time of familiarization, the participants are then encouraged to perform the four increasingly complex exercises detailed in the learning medium. The chosen exercises are the one-arm dumbbell curl, the standing dumbbell press, a variation of the one-legged stance and the dumbbell row. They were chosen in a way that steadily increases the difficulty and complexity by recruiting more muscle fibers and using multiple limbs.

The interviewer, who is a certified personal trainer, then documents their performance and the participants answer the questions regarding their opinion of the learning tool. The cycle repeats with the second and finally the third tool, after which the participants are questioned a final time about their learning experience during the user study and whether or not they enjoyed working with the 3D learning tool.
They are then asked to give their opinion on the user study's execution in its entirety and, if possible, voice their criticism about each learning tool while comparing them.
After concluding the study, the interviewer can evaluate the survey data as a form of quantitative research in addition to the documented gradings.

5 Implementation

5.1 Avatar

Autodesk Maya was used to create and rig the avatar. As seen in figure 5.1, the avatar possesses individual limb, spine and finger bones to allow humanoid movement. Since real-world locomotion depends on ligaments and different types of joints [Bia12], it would seem obvious to try the approach of tracking their movement and applying it to the model. Since a rigged avatar lacks muscles and joints, however, all movement has to be applied to the bones. Hence, each worn tracker is assigned to a bone and moved accordingly.

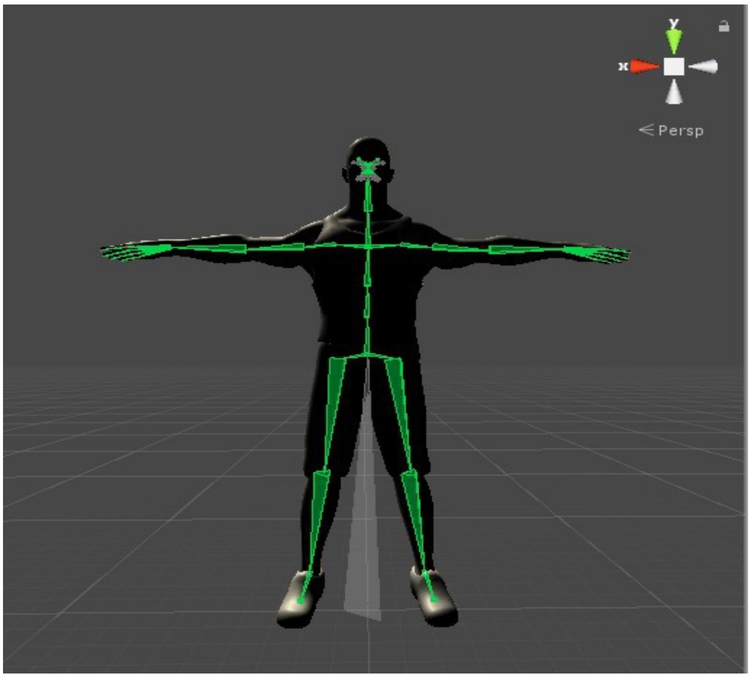

Figure 5.1: Rigged 3D model of a human body.

5.2 UI

To fulfill the requirements, the user interface should allow the user to start and stop recordings. Also, the user should have the option to pause an animation to study the movement at a specific point in time. Figure 5.2 shows the user interface with buttons required to enable these actions.

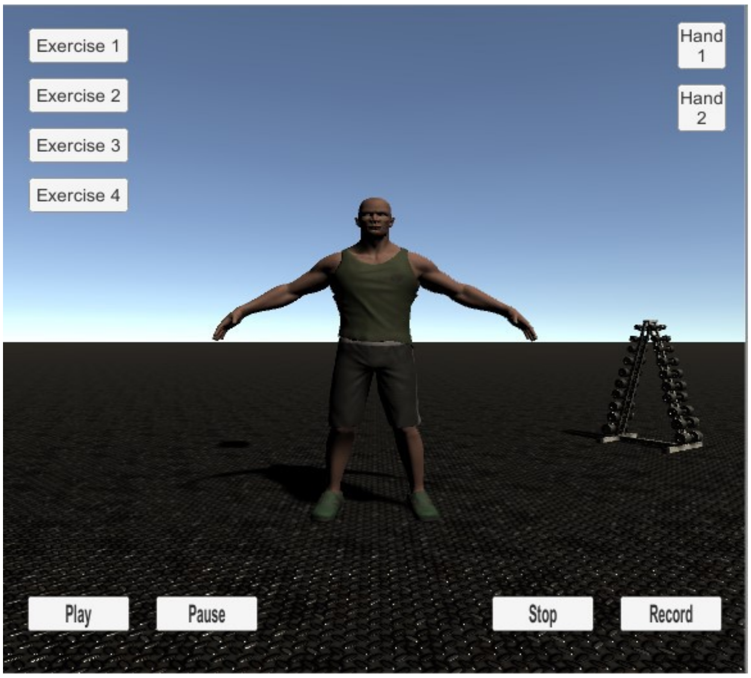

Figure 5.2: User interface for the 3D application.

In addition to the essential buttons, more quality of life enhancements have been integrated into the UI. The exercise buttons allow a user to dynamically swap exercise animations and the hand buttons add or remove dumbbells from the avatar's hands.

5.3 Tracking

The tracking for most body parts is taken care of by the SteamVR Tracked Object [Uni19] script which is part of the Steam VR plugin. Only the hands are exempt, as they are associated with the controllers, which use the Steam VR Behaviour Pose [Uni19] script. The avatar's extremities move in accordance with the position and rotation of their correlated trackers [Cas+18].

Since the Final IK solution is added as a stabilizing and smoothing component, the scene's hierarchy has to be changed. The VR IK [Roo19] script calculates body movement based on the location of targets, which are the head, hands, pelvis and feet, and so-called bend goals, which are the elbows and knees. Properly implemented bend goals rectify the position of joints, where they would be deduced otherwise. (Figure 5.3) shows the way a properly configured tracker setup for the upper body should look and the resulting avatar movement.

Figure 5.3: Comparison of real-world and virtual upper-body movement

It's of great importance to not choose the tracked objects themselves as targets and bend goals. Otherwise, the avatar may become warped depending on how the trackers are worn.

Naturally, the tracker's virtual representation should follow its rotation, but the problem arises as soon as the virtual object's fixed up-vector differs from its real-world counterpart's. To resolve this issue, the object follows the position and rotation of its tracker but has a child object which in turn is the target for the Final IK solution. The child's orientation can be aligned with the tracker's orientation, correcting the limbs visually while still allowing the parent object to rotate accordingly.

5.4 Movement Recording

After setting the foundation, a method of recording the exercises had to be implemented. The Animation Serializer script creates serializable vectors for positions, and serializable quaternions for rotations in order for them to be saved as XML nodes which in turn form an XML file.

The format was chosen because it's simple and more flexible than some alternatives like HTML, but still well structured compared to a basic text file.

To begin with, the script needs a reference to the transform of every object being tracked. Also, in preparation for saving the data as an XML file, the classes Workout Animation and Workout Animation Node are constructed beforehand. The former forms the root of the XML file while the latter creates nodes that give it structure. Each node contains the position as a vector and the rotation as a quaternion of every one of the tracked objects. A bool determines if the recording is taking place or not. If the Record button is pressed, the bool's value is set to true and a Workout Animation Node is written and added to a list of nodes with every execution of the Fixed Update function, which runs at 50 frames per second (Code Listing 5.1). Once the recording is stopped, the serialized data is converted from a list to an array of nodes and saved as an XML.

```
public void FixedUpdate()
{
      if (!isRecording)
             return;
      WriteNode();
}

/// <summary>
/// Add every relevant Transform to the nodelist
/// </summary>
public void WriteNode()
{
      WorkoutAnimationNode node = new WorkoutAnimationNode()
      {
             headposition = new SerializableVector3(head.localPosition.x,
             head.localPosition.y, head.localPosition.z),
             headrotation = new SerializableQuaternion(head.localRotation.x,
             head.localRotation.y, head.localRotation.z,
             head.localRotation.w),
             .
             .
       };

       nodeList.Add(node);
}
```

Code Listing 5.1: Node with all Positions and Rotations being added every Execution of Fixed Update (shortened).

5.5 Data Conversion

As the exercises are, in the scope of this thesis, only recorded in the Unity Editor, a custom Editor Window is an excellent way of managing the data conversion. The window has an object field for an XML file and a text field for naming the resulting animation.

The conversion process is split into two steps, each implemented in a separate function. The first function extracts the data from an XML file to make it available for processing by loading in the XML text file and filtering the data by its associated tag (Code Listing 5.2). Each of the twenty resulting node lists contains data about the position or rotation of one of the ten tracked body parts.

To access that information, each node's child nodes have to be passed through so the float values for the axes can be subsequently combined into vectors and quaternions.

```
/// <summary>
/// Extract data from XML file and prepare for conversion
/// </summary>
void ExtractXMLData()
{
    XmlDocument xmlDoc = new XmlDocument();
    xmlDoc.LoadXml(textAsset.text);
    XmlNodeList xmlHeadPosition =
    xmlDoc.GetElementsByTagName("Headposition");
    XmlNodeList xmlHeadRotation =
    xmlDoc.GetElementsByTagName("Headrotation");
    .
    .
}
```

Code Listing 5.2: XML nodes are sorted into lists based
on their associated tags (shortened).

To ensure that the extracted values are not affected by any regional differences such as the interpretation of the decimal point, which in some regions is omitted in favor of the decimal comma, an IFormatProvider is passed as the third parameter while parsing, as seen in Code Listing 5.3.

```
foreach (XmlNode positionInfo in xmlHeadPosition)
{
    Vector3 position = new Vector3();
    XmlNodeList positionDetails = positionInfo.ChildNodes;

    foreach(XmlNode detail in positionDetails)
    {
        if (detail.Name == "X")
        {
            float.TryParse(detail.InnerText, NumberStyles.Float,
            CultureInfo.InvariantCulture, out position.x);
        }

        if (detail.Name == "Y")
        {
            float.TryParse(detail.InnerText, NumberStyles.Float,
            CultureInfo.InvariantCulture, out position.y);
        }

        if (detail.Name == "Z")
        {
            float.TryParse(detail.InnerText, NumberStyles.Float,
            CultureInfo.InvariantCulture, out position.z);
        }
    }
    headPosition.Add(position);
}
```

Code Listing 5.3: Parsing values using IFormatProviders to ensure correct decimal interpretation.

After all the positional and rotational information has been stored in the corresponding lists, the data is ready to be converted into an animation file.

The created animation clip has an associated animation curve along with a timeline. By adding keyframes to the curve for every frame of the animation timeline, the previously recorded movement can be recreated.

The components of every vector and quaternion must correspond to an individual keyframe, which signifies three keyframes per vector and four keyframes per quaternion. Each keyframe is then assigned to an object that should be modified. In this case, the objects in question are the targets and bend goals of the Final IK solution. Code Listing 5.4 shows the process for the head target's position.

```
 for (int i = 0; i < key_headPosition_X.Length; i++)
 {
        key_headPosition_X[i] = new Keyframe(timeline, headPosition[i].x);
        key_headPosition_Y[i] = new Keyframe(timeline, headPosition[i].y);
        key_headPosition_Z[i] = new Keyframe(timeline, headPosition[i].z);
        .
        .
 }

 curve = new AnimationCurve(key_headPosition_X);
 clip.SetCurve("Head Target Parent", typeof(Transform), "localPosition.x",
curve);

 curve = new AnimationCurve(key_headPosition_Y);
 clip.SetCurve("Head Target Parent", typeof(Transform), "localPosition.y",
curve);

 curve = new AnimationCurve(key_headPosition_Z);
 clip.SetCurve("Head Target Parent", typeof(Transform), "localPosition.z",
curve);
```

Figure 5.4: Positions and rotations for every tracked body part is added to the timeline with keyframes (shortened).

To conclude the data conversion, all lists used during data conversion are cleared to allow the conversion tool to be used in succession, and the newly created animation asset is saved in a separate folder with the other animations.

5.6 Movement Replication

With the creation of the animation asset, the last step is moving an avatar in a way that replicates the initially recorded movement. As stated before, the avatar used to record the movements is not a good option due to being associated with the trackers. Otherwise, the association would have to be severed every time an exercise is played. For that reason, a second avatar is created with the sole purpose of simulating the movements. This avatar also has the Final IK solution component VR IK but is not influenced by the Steam VR plugin. Instead, the targets are simulated by creating objects with similar behavior. Once again, the child object is the actual target for the solution whereas the parent object is moved according to the played animation. Since the IK solution always behaves the same with identical input, the resulting movement is indistinguishable from the original.

6 User Study

The following section details the user study's procedure. It is split into three sections, which describe the preparations for the study, the method of execution and the evaluation of the results, respectively. For the sake of clarity, evaluations done by the participants are defined as ratings and those done by the instructor as scores.

6.1 Preparation

The user study's preparatory phase began with ensuring a secure examination environment and gathering the necessary equipment. The room where the study was held had a cleared out space of roughly two square meters where the exercises were to be performed. In addition, a desk was provided for the periodic learning phases where the participant dealt with the supplied learning utensils. In consideration of safety concerns regarding the participant's health and the existing electronic equipment, exercises requiring weights such as dumbbells were performed with a softer alternative.

The examination room was supplied with two sets of participant sheets: one for user-evaluations and one for gradings to be done by the instructor. It it important to note that the data collected from user-evaluations gave insight into the participants' perception of the tool's learning effect whereas the instructor's score reflected the actual effectiveness through analyzing the learning progress made by using the learning medium.

The instructor prepared the learning tools used during the user study, which consisted of a set of two pictures depicting an exercise in its initial and final state, a video showcasing the full exercise in motion and a recorded 3D simulation of the exercise. The order of learning tool distribution was chosen deliberately, as the results should reflect whether the third learning tool is an alternative or upgrade from the others or not. For this study, the instructor did not perform any correctional measures unless absolutely necessary and explanatory texts were omitted from the learning tools.

Since the learning effect of the individual tools was the primary focus and said texts could have been edited into either of the learning tools, they were disregarded.

6.2 Execution

The user study was conducted with one participant at a time and took around 60 minutes. Each participant did a warm up routine to prevent injury, as performing even short bouts of exercise should be done in a responsible and healthy manner [Bia12]. Once limber, the participants were handed the first learning tool that they should familiarize themselves with. Each participant was given ten minutes after which the four depicted exercises were completed in succession while the instructor took notes.

Once all the movements had been performed, the participants were given the opportunity to rate each exercise based on how well it was explained by the learning tool. The points ranged from one to five, with one point being the lowest rating meaning the exercise was not defined well at all, and five points being the highest rating meaning the exercise was explained perfectly.

Meanwhile, the instructor evaluated the previous notes and formulated a score based on certain criteria that had to be met for each exercise. For example, the instructor deducted points for an improper hip angle during dumbbell rows or too much momentum during the standing dumbbell press. The participants were then handed the second learning tool, video footage showing the exercises. Once again, the participants were given time to process the information and asked to perform the movements. By this time, the participants had the experience to make a direct comparison between the first and second learning medium which could have influenced their second rating. After calculating the next score, the instructor provided the participants with the third and final tool, the 3D program. The participants finished the exercises and wrote their assessed rating down. Finally, the user results and instructor results were compiled and evaluated.

6.3 Results

The user research was conducted with twelve participants, whose names and genders have been anonymized as they have no impact on the evaluation result. Each user's experience level was taken at face value as attested by themselves since there is no way to quantify experience in weight training or physical therapy. Table 6.1 shows the compiled list of user evaluations for each learning tool. The participants are labeled with letters A through L for referencing purposes.

The median age was 25 years old with the youngest being participant D who was 19 years old and the oldest being participant F who was 44 years old. Of the 12 users, only five had prior experience, with the other seven having done some other form of sport or none at all.

Table 6.2 shows the individual participant scores for each exercise as calculated by the instructor. To evaluate the data gained through the user study as well as possible, the scores represented in both tables were analyzed together to show the juxtaposition of estimated and actual effectiveness.

The first tool consisting of the two images received a low rating, with a mean average of approx. 1.6 out of 5 points. The tool obtained a better rating from participants who did not have any prior sports experience, such as participants A and F. The average rating of inexperienced users (IU) was approx. 1.9 whereas the experienced users (EU) rated the tool with 1.2 points. It could be deduced that the EU see the tool in a worse light because they have practical knowledge and have already interacted with other learning tools, drawing a direct comparison.

This was undermined further by examining the instructor scores. The average score for all participants was 2.5 points, although EU fared much better with 3.7 points compared to IU with approx. 1.6 points. Seeing that the user-evaluation rating for IU was much higher in contrast to the respective instructor score, it was reasoned that the first learning medium was qualitatively poor.

Evaluating the results for the second learning tool confirmed this assumption. Firstly, the tool earned a higher overall user-evaluation rating of 3.7 points, a 2 point increase from the first tool. Also, the ratings for IU and EU were more uniform in their distribution with mean averages of 3.7 points for both. Secondly, after having some time to study the video material, users A and F came to the conclusion that they had been too benevolent with their rating of the first tool. Seeing that both participants did not have any experience and that the overall inexperienced user rating for the set of images was higher than the experienced user rating, it was concluded that users without experience have difficulty assessing what qualities to look for in a learning tool without a comparison.

This is relevant for real-world scenarios in which a person is interested in weight training and uses a learning tool that describes movements incorrectly, which can potentially lead to injury. Lastly, the speculation of the second learning medium being the superior learning tool was supported by the instructor score. With an IU score of approx. 3.3 and an EU score of 4.3 points, this tool only has a score difference of 1 point compared to the 2.1 point difference of the first leaning tool.

The third learning tool's results were the focal point of this study, determining whether the 3D program is an upgrade, a worthy alternative or an inadvisable learning medium compared with the others. The overall user rating was the highest of the three tools with a mean average of approx. 4.4 points, 0.6 points higher than the video material. The IU rated the application with approx. 4.4 points, the same as the EU. A fourth of all participants gave the tool a perfect score, with the main reason being the option of changing the perspective while pausing the animation to analyze the movement. All three were participants with prior experience.

Participant I deviated from the norm with a rating of 3 points, the same rating given to learning tool two. This is reflected by participant I's instructor score of 3 points, reaching the lowest level of progress out of all IU. The participant voiced sound criticism about every learning tool, the main point being missing text elements or markers to show a user what body parts to focus on during the exercise. This shows that not every user responds well to a purely visual learning method.

Other participants have also noted that tracking-based anomalies led to some confusion but were ultimately not a great concern. A point speaking for the user study's success is the evaluation of the instructor scores. Scores for IU were approx. 3.8 points on average, a 0.4 point increase from the previous learning tool. This is impressive considering the users did not have any knowledge about exercise movements before. Participant E not only gave each exercise the same rating as the instructor score but also showed the most growth with a difference of 3 points. Contrarily, the user with the least growth was participant G, who reached the maximum score of 5 points for each learning tool. EU, in general, did well across all learning mediums but were still able to get an average score of 4.7, a 0.4 point increase. The 1 point difference between EU and IU remains.

In addition to the improved scores, participants reported that they saw the learning tool's merit and wished to see it improved upon and made accessible in the future. Users B, C, F, H, and L also disclosed that they thought the application was enjoyable and users C and G revealed that their rating was slightly influenced by the tool's flexibility and extensibility. Thus, the 3D program was well received by participants of varying ages, making the user study a success for this thesis.

Participant	**Age**	**Experience**	**Tool 1**	**Tool 2**	**Tool 3**
A	24	No	3/5	5/5	5/5
B	24	Yes	1/5	4/5	4.5/5
C	30	Yes	2/5	4/5	5/5
D	19	No	2/5	4/5	4/5
E	21	No	1/5	3/5	4/5
F	44	No	3/5	4/5	5/5
G	32	Yes	1/5	3/5	4/5
H	26	Yes	1/5	4/5	4.5/5
I	23	No	1/5	3/5	3/5
J	28	Yes	1/5	3.5/5	4/5
K	20	No	2/5	3.5/5	5/5
L	39	No	1.5/5	3.5/5	4.5/5

Table 6.1: Results of the User Evaluation

Participant	**Age**	**Experience**	**Score 1**	**Score 2**	**Score 3**
A	24	No	1.5/5	3/5	3.5/5
B	24	Yes	3.5/5	4/5	4.5/5
C	30	Yes	4/5	4.5/5	5/5
D	19	No	1/5	3/5	3/5
E	21	No	1/5	3/5	4/5
F	44	No	2.5/5	4/5	4.5/5
G	32	Yes	5/5	5/5	5/5
H	26	Yes	4/5	5/5	5/5
I	23	No	2/5	3/5	3.5/5
J	28	Yes	2/5	3/5	4/5
K	20	No	2/5	4/5	4/5
L	39	No	1.5/5	3/5	3.5/5

Table 6.2: Results of the Instructor Evaluation

7 Future Developments

As the 3D application used for the user study was only designed to fulfill its purpose as a learning tool, it still has untapped potential. The following are different scenarios in which the program, given enough time and resources, could be used in the future. In these cases, the field of application can either be the same as this thesis' or something beyond the scope of sports.

7.1 Layered Movement Comparison

By implementing a mode in which the avatar replicating the exercise movements is replaced with a translucent substitute, a user equipped with tracking devices can perform the same action while standing inside it. While viewing through the HTC Vive headset, the user can try to make the movements overlap. Optionally, multiple movements can be played in succession to create a complete workout. The time in which the movements overlapped can be tracked and eventually scored based on how close the movements were followed.

An important point to consider is the problem of height difference, which would either require and size adjustable avatar or predefined avatars with specific sizes. In that case, the virtual target positions must be adaptable as well, since the discrepancy between the position of worn trackers and virtual trackers would result in wrongly interpreted movement otherwise.

7.2 High-Performance Sports Analysis

Another scenario in which the program could be applied is in high-performance sports. Athletes with particularly respectable abilities could record movements for their respective sport, which others could use to study and train with.

As minuscule movements can decide the outcome of a competition, athletes may choose to emulate the motion to spot differences between the recorded movement

and their own. For example, football or soccer players could observe the kicking or throwing motions, respectively.
As of today, this approach is not yet feasible using the aforementioned tracking hardware. The trackers are far too large to not have an impact on the athlete's performance and the tracking area may be too small for every type of sports movement. Also, the tracking does not extend to the individual finger joints, which is imperative for sports with precisely timed grabbing or releasing motions. This problem may diminish in the future with steadily evolving tracking technology.

7.3 Motion Capture

The last featured adaptation of this 3D application could be used for motion capture in games or movies. Once the tracking hardware has been set up, the user can record whatever animations are required. One advantage of using this method of motion capture is efficiency since the process of recording itself is quick, easy and cheap. Animating a humanoid movement by hand can take a much longer time and mistakes have a greater impact.

Another advantage is the fluidity and naturalness of the motion. To achieve the same effect any other way, the movement must be interpolated as little as possible, meaning it has to be framed by hand, which is a strenuous task. If implemented in games, the user can also add colliders to parts of the avatar's body, allowing it to physically interact with objects, for example pushing an obstacle out of the way or throwing a damaging punch.

8 Conclusion

This empirical research aimed to determine the measure in which a 3D program simulating a person's movements was an adequate alternative to well known learning media in sports and physical therapy. Based on quantitative analysis in form of a score-based user study and qualitative analysis from participants' feedback, it can be concluded that the application is not only a possible substitute, but an enhancement. The results show that this approach can convey more information, especially towards inexperienced users. Experienced participants praised the application as well, further demonstrating that it fulfills the necessary qualities for a learning tool.

The participants' criticism, however, also showed that the program has it's flaws. It was concluded that the visual learning method is not appropriate for everyone, limiting the tool's assumed universality. Also, the application is held back by some previously mentioned tracking-based anomalies. Even so, it fulfilled its purpose and confirmed the established hypotheses.

List of Figures

Figure 2.1	Picture by Hay Kranen/CC BY-SA 3.0. *“The Utah Teapot, rendered in 3DSMAX 5.0 (Mar. 2007)”,* URL:https://upload.wikimedia.org/wikipedia/commons/4/47/Utah_teapot_3dsmax.png, (last visited on 11/13/2019).
Figure 3.1	Picture by Daniel Jazz Young. *“HTC Vive components (Oct. 2019)”.*
Figure 3.2	Picture by Daniel Jazz Young. *“Two HTC Vive Tracking Devices (Oct. 2019)”.*
Figure 4.1	Picture by unknown artist, released into the public domain. Adapted by Daniel Jazz Young. *“black outline of a human male body (Nov. 2007)”,* URL:https://upload.wikimedia.org/wikipedia/en/0/0e/Outline-body.png, (last visited on 11/13/2019).
Figure 5.1	Picture by Daniel Jazz Young. *“3D model of human with bones (Oct. 2019)”.*
Figure 5.2	Picture by Daniel Jazz Young. *“3D application UI (Oct. 2019)”.*
Figure 5.3	Picture by Daniel Jazz Young. “*Real-World and Virtual Movement Comparison (Nov. 2019)”.*

List of Tables

List of Code Listings

References

[Ban04] A. Bandura. “Observational Learning”. In: *Learning and Memory (2nd Edition) (2004)*, pp 482-484.

[Bha+16] K. K. Bhagat, W. Liou and C. Chang. “A cost-effective interactive 3D virtual reality system applied to military live firing training”. In: *Virtual Reality (Jun. 2016)*, Volume 20, Issue 2, pp 127-140, URL: https://link.springer.com/article/10.1007/s10055-016-0284-x, (last visited on 11/10/2019).

[Bia12] B. D. Biagioli. “*Advanced Concepts of Personal Training (2012)*”.

[Bor+18] M. Borges, A. Symington, B. Coltin, T. Smith and R. Ventura. “HTC Vive: Analysis and Accuracy Improvement”. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (Oct. 2018)*, ISBN: 978-1-5386-8094-0.

[Cas+18] P. Caserman, A. Garcia-Agundez, R. Konrad, S. Göbel and R. Steinmetz. “Real-time body tracking in virtual reality using a Vive tracker”. In: *Virtual Reality (Jun. 2019)*, Volume 23, Issue 2, pp 155-168.

[Che+09] B. Chen, F. Huang and H. Lin. “Using virtual world technology to construct immersive 3D virtual university”. In: *17th International Conference on Geoinformatics (2009),* ISBN: 978-1-4244-4562-2.

[Chr+88] R. W. Christina and D. M. Corcos. *“Coaches Guide to Teaching Sport Skills (1988)”.*

[Cla+02] M. J. Clayton, R. B. Warden and T. W. Parker. “Virtual construction of architecture using 3D CAD and simulation”. In: *Automation and Construction (Feb. 2002)*, Volume 11, Issue 2, pp 227-235.

[Ham94] B. P. Hamill. “Relative Safety of Weightlifting and Weight Training”. In: *Journal of Strength and Conditioning Research (1994)*, Volume 8, pp 53-57.

[Htc19] HTC. “*Vive”*. URL: https://www.vive.com/us/support/vive/, (last visited on 11/13/2019).

[Jod+14] T. Joda and T. O. Gallucci. “The virtual patient in dental medicine”. In: *Clinical Oral Implants Research (2015)*, URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/clr.12379, (last visited on 11/13/2019).

[Kat+06] L. Katz, J. Parker, H. Tyreman, G. Kopp, R. Levy and E. Chang. “Virtual Reality in Sport and Wellness: Promise and Reality”. In: *International Journal of Computer Science in Sport (2006)*, Volume 4, pp 4-16.

[Luc18] E. Luckett. *“A Quantitative Evaluation of the HTC Vive for Virtual Reality Research” (May 2018).*

[Nie+17] D. C. Niehorster, L. Li and M. Lappe. “The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research”. In: *i-Perception (Jun. 2017).*

[Pee+18] A. Peer, P. Ullich and K. Ponto. “Vive Tracking Alignment and Correction Made Easy”. In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (Mar. 2018)*, ISBN: 978-1-5386-3365-6.

[Roo19] Root Motion. *“Final IK Documentation”*. URL: http://www.root-motion.com/finalikdox/html/index.html, (last visited on 11/13/2019).

[Sal+95] K. Salisbury, D. Brock, T. Massie, N. Swarup and C. Zilles. “Haptic rendering: programming touch interaction with virtual objects”. In: *1995 symposium on Interactive 3D graphics (1995)*, pp 123-130, ISBN: 0-89791-736-7.

[Shi+14] A. Shingade and A. Ghotkar. “Animation of 3D Human Model Using Markerless Motion Capture Applied To Sports”. In: *International Journal of Computer Graphics & Animation (Jan. 2014)*, Volume 4.

[Suz+98] N. Suzuki, A. Hattori, T. Ezumi, A. Uchiyama, T. Kumano, A. Ikemoto, Y. Adachi and A. Takatsu. “Simulator for virtual surgery using deformable organ models and force feedback system”. In: *Medicine Meets Virtual Reality (1998)*, pp 227-232.

[Tor06] A. Torrence. “Martin Newell's original teapot”. In: *SIGGRAPH '06 (2006).*

[Tuz02] V. P. Tuzlukov. “Signal Processing Noise”. In: *The Electrical Engineering and Applied Signal Processing Series (Apr. 2002).*

[Uni19] Unity. *“Unity User Manual”*. URL: https://docs.unity3d.com/Manual, (last visited on 11/13/2019).

[Usa19] Usability. *“User Research Basics”*. URL: https://www.usability.gov/what-and-why/user-research.html, (last visited on 11/13/2019).

[Yoo+13] H. Yoon and J. Park. “Avatar Animation Using Skeleton Tracking With Kinect Sensor”. In: *International Journal of Advancements in Computing Technology (Aug. 2013)*, Volume 5.

[Zen+12] M. Zeng, Z. Liu, Q. Meng, Z. Bai and H. Jia. “Motion capture and reconstruction based on depth information using Kinect”. In: *Image and Signal Processing (Oct. 2012)*, pp 1381-1385.