Research Proposal

Immersive Physical Therapy Experience in Virtual Reality

Ritika Tejwani

Dr. Henry Fuchs, Husam Shaik

Computer Graphics and Virtual Reality Research

The University of North Carolina at Chapel Hill

2020-2021

Written in Fall of 2020

RATIONALE

Significance

Recently, many healthcare systems and medical organizations are looking to integrate augmented reality and virtual reality programs into doctor-patient interactions in hospitals and clinics, with the aim of increasing efficiency and reducing physician burnout. Furthermore, immersive experiences in virtual reality are becoming a focal point within the medical field as they can not only help physicians, but can also be used by medical students, trainees, and patients for learning purposes. For example, patients can learn exercises from physical therapists in virtual reality. With such a program, ideally, therapists would record themselves doing various exercises and include annotations in real-time. The recording would then be reconstructed in 3D space for a VR experience. Users would be able to see the exercises from different angles and walk around the environment to get a better understanding of how to complete exercises with proper form, thus reducing chances of injury.

An immersive physical therapy experience in virtual reality (VR) could effectively connect physical therapists with patients remotely, which is especially beneficial to physical therapy clinics during the COVID-19 pandemic. Several gyms and physical therapy clinics in various states across the world were forced to temporarily close during the pandemic as they are considered “superspreader” locations (Chang 2020). Virtual reality (VR) is a practical solution to allow training to continue effectively and efficiently without the risk of COVID-19.

Background Literature Review (Must have at least 6 references)

Immersive virtual reality is a simulated experience that looks and feels like reality (Psotka 1995). Advantages of such experiences include prompting “mindful engagement” and allowing students to train for various tasks without severe consequences of their actions. Muijzer, in his Master’s Thesis in 2014, recognizes not only the need for patients and physical therapists or personal trainers to communicate remotely over a medium more effective than video but also the need for “automated rehabilitation exercises detection and evaluation systems” that give feedback to patients on their performance and instructs them how to improve their exercise form to reduce chances of injury in the future. While the benefits to patients are obvious, physical therapy or exercise workout sessions that are held through a virtual reality program are beneficial to physical therapists and trainers as well, as they are able to work with more patients over a shorter span of time, thereby increasing their productivity (Muijzer 2014). Muijzer details exercise detection and comparison methods that can accurately provide feedback to patients in real-time, thereby reducing chances of injury for patients in the future.

The most prominent and novel advancements in virtual reality primarily involve advancements in the reconstruction pipeline. Previous research outlines a clear image reconstruction process that is effective in creating an immersive experience in virtual reality. As explained by Cha (2016), the reconstruction pipeline is separated into two primary components. First is the reconstruction of the static background environment and second is the reconstruction of the human model and its movements. Equipment includes depth cameras (ex. Kinect v2 or Intel RealSense), a high-performance computer, and a virtual reality headset, such as an Oculus Quest.

Cha (2016) emphasizes the novel advancements in reconstruction technology by detailing each step of the reconstruction process used in the project. The paper begins by explaining that the background environment is reconstructed through an initial video of the room. After the frames are extracted from the video, they are imported into a software known as Metashape, which performs photogrammetric processing of the images and generates 3D spatial data from them. After building 3D spatial data, the points are tied together: the point cloud is converted to a dense cloud and then to a mesh. It is also important to consider decimating the mesh and removing spatial points of low confidence in order to generate a more realistic environment.

Cha (2016) further explains that the human model is reconstructed through an initial static scan in a T-pose. The T-pose minimizes the deformations in the mesh and ensures that all points of the body are visible for the camera to generate spatial data. Another depth camera is used to track the skeletal movements of the human body. Skeletal tracking includes tracking of the head, neck, spine, limbs, hands, feet, and others. Recording such skeletal movements can provide an animation, which can later be applied to the static human mesh. Before applying the animation, a skeleton must be rigged to the static human model so that the skeleton of the human can move in accordance with the animation. When applying the animation to the rigged human model, it is important to remove deformations and inconsistencies. After both the static environment and the animated human model are reconstructed, they can be combined in Unity, a 3D development platform. In Unity, the VR experience is customized and a user interface is incorporated through Unity scripts. The program is then built and transferred to a virtual reality headset, where the user can walk around in the 3D environment, play/pause the movements, view real-time annotations in space, and use other user interface features.

A primary application of image reconstruction techniques is creating an immersive surgical training experience in virtual reality for medical students. Being able to view extreme and rare cases in the operating room can teach medical students how to handle such cases and reduce surgical errors. Researchers at UNC's Computer Graphics and Virtual Reality Group did just that in 2016 when they used depth cameras to record and reconstruct a simple surgery (Cha et al., 2016). Virtual reality also provides a platform for medical students to practice surgeries via sensors, thus providing a real-time experience of rare or extreme situations. By creating a personalized and adaptable experience for each student, medical students are able to better prepare for difficult surgeries.

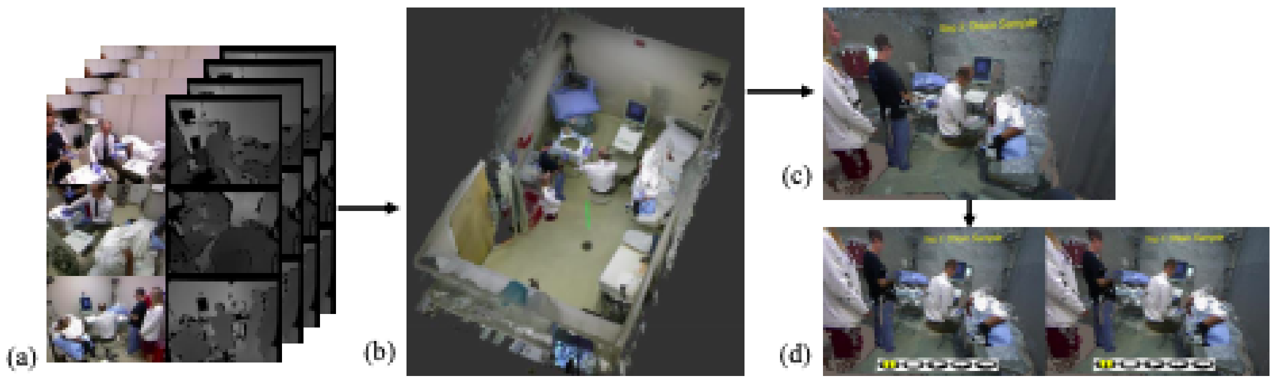


Figure 1 (Cha et al, 2016) demonstrates the 3D reconstruction pipeline, with the final product being an immersive experience with spatial annotations.

Real-time motion detection is an emerging advancement that has a wide range of applications, such as remote physical therapy sessions. Some very recent research describes how users can detect their own movements in real-time by using a single RGB camera and a mirror in front (Unberath et. al, 2020). Such advancements make it possible for physical therapists to conduct sessions with their patients remotely, effectively, and realistically. Real-time motion tracking can be used to compare user's movements with those of the trainers. Research shows that a cloud-based system can be used to create a system in which a therapist can record a movement in which the user can replay and the program will generate a score measuring how well the user completed the exercise (Wei, Lu, Rhoden, Dey, 2017). Similarly, a 2019 study conducted by Saraee and other researchers at Boston University experimented with a home-based physical therapy experience in virtual reality in which the user's completion of the exercise was compared to that of the tutorial and the results showed an improvement of performance through the use of the program.

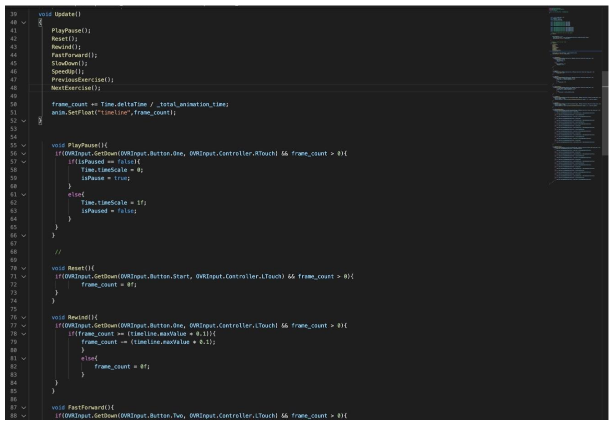
As explained by Barnart (2020), a relevant factor for consideration is how the population of interest will adapt to the new technology. For example, deploying virtual reality headsets with such immersive surgical training experiences is not only an expensive venture but also noticeably different than conventional methods of learning. Criticism, user backlash, and transition periods are expected with new technology. It is also important to re-evaluate the program frequently and measure its progress to continually improve the effectiveness of virtual reality in the education setting.

RESEARCH DESIGN

Specific Aim 1: To develop an immersive physical therapy virtual reality experience that allows the user to control the experience from the controllers.

Motivation: Just like a patient can view their physical therapists perform exercises from all angles in real-life, allowing the patient to use the controllers to move around in the virtual reality environment is both practical and necessary for a virtual reality physical therapy session to be effective and efficient. Because viewing the exercises remotely can pose several challenges, I also aim to add several features, such as play, pause, speed up, slow down, rewind, fast forward, and reset to allow the patient to better understand how to properly perform each exercise, thus reducing chances of injury for the patient. I aim to program the controllers for these features instead of displaying a canvas on the screen to allow for more efficient use of the user interface and to not disrupt the viewing of the exercises on screen.

Supporting Preliminary Data: Thus far in the project, I have been able to program three buttons on the virtual reality controllers to play, pause, and reset the exercise animation. When the first button on the right controller is pressed, it pauses the animation and the second button will play the animation again. This ease of use allows the user to quickly play or pause the animation while still being able to walk around the environment and view the exercise from different angles.



```
void Update()
{
    PlayPause();
    Reset();
    Rewind();
    FastForward();
    SlowDown();
    SpeedUp();
    PreviousExercise();
    NextExercise();

    frame_count += Time.deltaTime / _total_animation_time;
    anim.SetFloat("timeline",frame_count);
}


    void PlayPause(){
     if(OVRInput.GetDown(OVRInput.Button.One, OVRInput.Controller.RTouch) && frame_count > 0){
         if(isPaused == false){
             Time.timeScale = 0;
             isPause = true;
         }
         else{
             Time.timeScale = 1f;
             isPaused = false;
         }
     }
    }

     //

    void Reset(){
     if(OVRInput.GetDown(OVRInput.Button.Start, OVRInput.Controller.LTouch) && frame_count > 0){
             frame_count = 0f;
     }
    }

    void Rewind(){
     if(OVRInput.GetDown(OVRInput.Button.One, OVRInput.Controller.LTouch) && frame_count > 0){
         if(frame_count >= (timeline.maxValue * 0.1)){
             frame_count -= (timeline.maxValue * 0.1);
             }
             else{
                 frame_count = 0f;
             }
     }
    }

    void FastForward(){
     if(OVRInput.GetDown(OVRInput.Button.Two, OVRInput.Controller.LTouch) && frame_count > 0){
```

Figure 2 displays a C# script that enables the patient to use buttons on the controllers to play and pause the exercise animation.

Methods: To create a functioning user interface through the virtual reality controllers, I will program in C# in Unity. Through creating scripts and coding methods, each button on the controller can be programmed to perform a different function. In the update function for the script, I will call methods for each button using a conditional statement. For example, in the update function of the script, I will include methods such as slow down, so that if a certain button is pushed on the controller, the exercise animation will decrease to 0.75 speed of the original clip. Such functions will allow the user to view the exercises multiple times and at a pace that is right for them.

Data Analysis: I will use trial and error as a method to evaluate the user interface system. For each modification in the C# programming scripts, I will build the project and determine whether the modification enabled a user interface function or not. By analyzing what functions work and which don't as a result of changes in coding, I will be able to analyze the user interface system. In addition, I will also compare the ease of the use of the VR controllers user interface to the canvas user interface made during the summer. I will distribute one survey to five people and each person that completes the survey will try the project made over the summer and the project made during the school year (the order that they try the VR experiences will be determined randomly with the flip of a coin). The survey will have ten questions and will be divided into two sections: five questions about the first user interface and five questions about the second user interface type. All the questions will ask for a rating between 1-10 (with 10 being the best) regarding five qualitative measurements. For each survey completed, I will sum up the five ratings given for each interface. Whichever user interface type has the greatest sum will be considered to have "won" that survey. Out of the five surveys completed, whichever user interface type has the majority of wins will be considered to be the most effective type.

Expected Outcomes and Limitations: I expect the user interface controlled by the Oculus Quest controllers will receive higher ratings than the user interface controlled by canvas buttons. It is expected that the controllers will have several functions, including play, pause, speed up, slow down, rewind, fast forward, and reset, each initiated through pressing a specific button on the controllers. It is also expected that pressing a specific button will bring up 3D spatial exercises for the specific exercise animation being played, instructing the user on the proper form for each movement. All interactions with the VR experience will be through the controllers, including moving around (the joysticks) and the user interface features (buttons). Limitations in this system arise if the patients wish to practice the exercises while watching the physical therapist in virtual reality. In such a situation, the patient would have to set the controllers down and would not be able to use the functions in the user interface while trying the exercises on their own.

Specific Aim 2: To engineer a way to record complex physical exercises with high accuracy using the Kinect v2 depth camera and apply the animations to a human model with minimal deformations.

Motivation: In creating an immersive VR experience for physical therapists to connect with patients while social distancing, the most important component of this project is recording exercises. The physical therapist's movements must be reconstructed and displayed accurately and reliably in order to ensure that the project is functional for patients to learn the proper form for exercises in this VR experience.

Supporting Preliminary Data: Thus far, I have recorded exercises I performed while standing in front of the Kinect v2 through Unity software. I was able to do this through coding C# scripts where I tracked JSON data of the positions and rotations of all 22 joints, relayed the joint movements on the human model in Unity, and recorded the exercise reconstruction so that I could play it back. In doing this, I was able to reconstruct several exercises, including squats, jumping jacks, high-knees, and more.

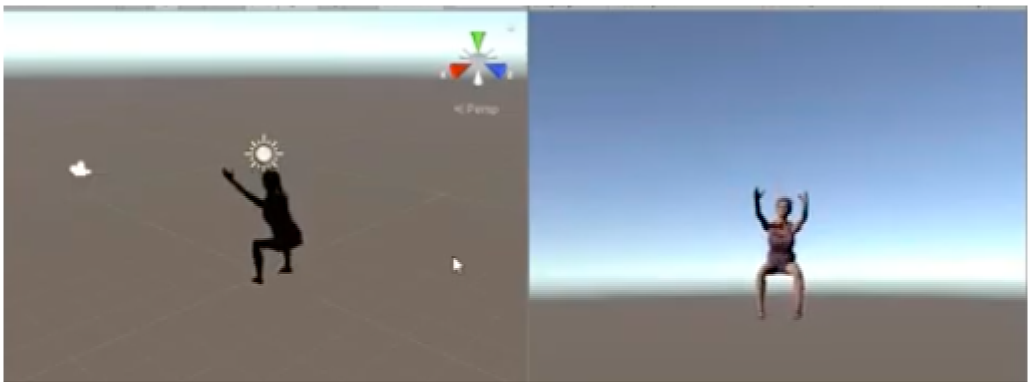


Figure 3 displays a 3D reconstruction of squats in Unity through two different perspectives. The image on the right is the VR experience view.

Methods: To reconstruct a physical therapist's movements while they guide the user through the proper form of exercises, two components are needed. First is a static model of the physical therapist, created using an Intel RealSense depth camera and Itseez3D program. The static human model will be reconstructed in a T-pose to allow for the movements to be more effectively applied to the human model. Then the exercises are recorded using the Kinect v2 depth camera and they are recorded. The joint transformations are automatically applied to the human model created previously so the reconstruction is complete. The exercises must be performed slowly and directly in front of the camera in order to minimize deformations or inconsistencies in the reconstruction.

Data Analysis: The reconstruction of the exercises will be analyzed through a trial and error method. Using an initial reconstruction, I will analyze the number of inconsistencies in the exercise reconstruction by looking at the skeletal display and counting the number of times that a joint moves in a direction it is not supposed to (indicating that the joint was not detected properly). For each modification I make in the reconstruction process, I will repeat the count process to determine whether the reconstruction adjustments worked or not. The trends in the number of inconsistencies will be analyzed by creating a line graph, with the modifications (trials) on the x-axis and the number of inconsistencies in the display on the y-axis.

Expected Outcomes and Limitations: It is also expected that several realistic physical therapy exercises will be reconstructed and recorded for the VR experience, each with very few inconsistencies and deformations. I expect the line graph for each exercise will be down-sloping, indicating that the modifications to the reconstruction process worked in decreasing the number of inconsistencies in the exercise reconstruction. The exercise animations should have high clarity and be effective enough for patients to learn how to complete the exercises with proper form in the VR experience. Challenges in achieving this goal are related to the efficiency of the program and how accurate the recordings are with complex exercises. Oftentimes, when one limb goes in front of the other, the Kinect v2 skeleton tracking gets confused/disjointed and is unable to correctly display the exercises. This is incredibly important for physical therapy exercises, which often require stretching, and many exercises are performed while laying on the ground.

Specific Aim 3: To explore real-time motion tracking concepts by creating a physical therapy tutorial that prompts the user to do the exercise and tracks the number of repetitions the user completes.

Motivation: This goal addresses the interaction between the physical therapist and the patient in an in-person physical therapy session. To make the virtual reality experience as realistic as possible, I will include an interactive experience for the patient to ensure that they are performing the exercises with the proper form, thus reducing chances of injury.

Supporting Preliminary Data: Thus far, using the JSON data and joint position and rotation data I collected in recording movements in front of the Kinect v2, I have read several papers that explain how to achieve this goal and I have started coding in C# scripts do this. I have outlined the transformations for specific joints involved in each specific exercise and have begun calculating the numerical position and rotation changes. These pre-set criteria will be used in methods in the future to determine which exercise the user is performing.

Methods: To identify which exercise the patient is performing in front of the camera, there must be pre-set criteria for each exercise. These criteria should consist of major joint transformations and rotations for each exercise. Each exercise will also be separated into phases: starting phase, phase 1, phase 2, etc. The positions and rotations of each joint will be recorded for each phase and the transitions between the phases in order to develop the pre-set criteria for each exercise. Then, the Kinect v2 will track the joint positions and rotations of the patient's skeleton as he/she completes an exercise. By comparing these joint transformations to the pre-set criteria and choosing the exercise whose criteria best match the patient's movement, the exercise can be identified. Once the exercise is known, the Kinect v2 will continue tracking the patient's movements and will count the number of times the patient runs through all the outlined phases of the exercise in order to count how many repetitions the user is completing.

Data Analysis: This data will be analyzed by mathematically computing the accuracy of identifying the exercise being performed and how many repetitions are being completed. I will ask three other people to help me in this process. I will ask each person to complete five different exercises in an order of their choice and with the number of repetitions for each exercise of their choice as well. For each exercise that is completed, I will record whether the program detected it accurately or not and whether the program counted the number of repetitions correctly or not. At the end of the 15 exercises completed, I will determine the percentage of accuracy of the program using the following formulas: accuracy of detection = (# of exercises detected accurately)/15 * 100 and accuracy of repetitions = (# of exercises where the repetitions were counted accurately)/15 * 100. Using these percentages, I can describe the effectiveness of the interactive component of the project.

Expected Outcomes and Limitations:
It is expected that the program will be able to identify which exercise the patient is completing and to accurately count the number of repetitions a patient performs of an exercise in front of the Kinect v2 with at least a 70% accuracy rate. If there is a lot of variability in these results and the program is not able to identify the exercise or not accurately count the number of repetitions, then the program is not successful. If enough time remains, I can also develop a method to reliably compare these movements to that of the physical therapist to determine if the patient is performing the exercises with proper form, For example, if a physical therapist records themselves doing the exercise, the patients can perform the exercise in front of the Kinect and the program will calculate how well the patient completed the exercises in comparison to the physical therapist. I can do this by comparing the position and rotation values I collected. These values can also be used to count how many repetitions of a certain exercise the patient has completed. Exercises with minimal joint transformation or exercises that are complex in which several joint transformations may limit the performance of this program, as it will be harder to track and thus will reduce the accuracy of the program.

Risk and Safety

There is a chance of risk when completing exercises to be reconstructed. But the exercises being reconstructed are all simple exercises that are traditionally performed by physical therapists. Before recording exercises, the proper form for each exercise will be researched and implemented and all exercises will be performed in an open space.

There is also a risk involved for the patients while in the VR environment as they are using a headset and are not able to see their real-life environment. It is important for the user to initially outline their boundaries on the VR headset and to use an open space in order to reduce the chances of injury.

BIBLIOGRAPHY

Barnard, M. (2020, June 22). Best Practices for Deploying Immersive Learning Programs. Retrieved October 06, 2020, from https://virtualspeech.com/blog/deploying-immersive-learning-programs

Cha Y.-W., Dou M., Chabra R., Menozzi F., State A., Wallen E., MD, Fuchs H.(2016). Immersive Learning Experiences for Surgical Procedures, Medicine Meets Virtual Reality 22: NextMed / MMVR22, 220, 55-62, doi: 10.3233/978-1-61499-625-5-55

Chang, S., Pierson, E., Koh, P.W. et al. Mobility network models of COVID-19 explain inequities and inform reopening. Nature (2020). https://doi.org/10.1038/s41586-020-2923-3

Muijzer, F. (2014). Development of an Automated Exercise Detection and Evaluation System Using the Kinect Depth Camera. Biomedical Engineering MSc (66226).

Psotka J. (1995). Immersive training systems: Virtual reality and education and training. U. S. Army Research Institute.Kluwer Academic Publishers. Instructional Science 23, 405-431

Saraee, E., Singh, S., Hendron, K., Zheng, M., Joshi, A., Ellis, T., Betke, M. (2017). ExerciseCheck: Remote Monitoring and Evaluation Platform for Home Based Physical Therapy. 10th Annual International Conference on Pervasive Technologies Related to Assistive Environments (PETRA'17), Rhodes, Greece, June 2017. DOI: http://dx.doi.org/10.1145/3056540.3064958

Unberath M., Yu K., Barmaki R., Johnson A., and Navab N. (2020). Augment Yourself: Mixed Reality Self-Augmentation Using Optical See-through Head-mounted Displays and Physical Mirrors. IEEE VR TVCG.

Wei W., Lu Y., Rhoden E., Dey S. (2019). User performance evaluation and real-time guidance in cloud-based physical therapy monitoring and guidance system. Multimedia Tools and Applications, 78, 051–9081. DOI: https://doi.org/10.1007/s11042-017-5278-5