Evaluating Lightweight Optical Hand Tracking for Virtual Reality Rehabilitation

Oana Ferche, Alin Moldoveanu, Florica Moldoveanu
University POLITEHNICA of Bucharest
313 Splaiul Independentei, Bucharest
E-mail: oana.ferche@cs.pub.ro, alin.moldoveanu@cs.pub.ro, florica.moldoveanu@cs.pub.ro

Abstract. As part of the TRAVEE (Virtual Therapist for Neuromotor Rehabilitation through Augmented Feedback) project we require a device for a precise tracking of the movements of the human hand. We concluded that the most appropriate for our purpose (due to its lightweight design and ease of use) would be a solution based on optical tracking. In this paper, we presented the way in which we evaluated two such optical tracking devices: Leap Motion and Intel RealSense. We used this method because we believe it fits the best our particular scenario.

Keywords: neuromotor rehabilitation, optical tracking, virtual reality

1. Introduction

The purpose of the TRAVEE project is to assist the rehabilitation of patients who have suffered a disability as a result of a stroke. Often, these patients are left with a disability in the upper limb. By providing them with augmented feedback presented in a dedicated virtual environment, TRAVEE aims at finding a way to support the rehabilitation process.

The TRAVEE system creates a virtual environment (VE) in which the patient can see an avatar for himself as well as an avatar named Virtual Therapist (VT), that exemplifies the movements that the patient has to try to execute as part of his session. The patient observes the movements of the VT and tries to reproduce the movements in the real world. His movements are also represented in the VE and – when improvement is detected, they are also amplified, in order to accentuate the progress made. The improvements may represent an increased amplitude of the movement, more precision in movement, or activity detected at the surface of the brain (through Electroencephalography – EEG) or in the muscles required to execute the movement (through Electromyography – EMG).

In order to represent, evaluate and enhance the real movements of the
patient in the VE, these have to be tracked by a specialized device. After studying the available technologies as presented in (Ferche & Moldoveanu A. & Moldoveanu F., 2015) we concluded that the most appropriate solution would be using an optical device for tracking the hand movements.

As the TRAVEE system wishes to be affordable and easy to use by people without technical background, we focused on lightweight, low cost solutions. We chose the Microsoft Kinect for body tracking, but for more detailed movements of the hand we found it did not provide enough detail (only the direction of the forearm, the palm and that of the thumb) – as described in (MSDN, Jointype Enumeration). For detailed tracking of each finger we found two solutions that would fit our system: Leap Motion and Intel RealSense.

In this paper, we will present the rehabilitation exercise movements that will be considered in the TRAVEE system and we will try to track each of these movements with a placement that we thought would be best for each of the two tracking devices and we will discuss the results.

2. Overview

2.1 Available technologies

As stated, the two hand tracking devices will be considered in our evaluation: the Leap Motion and an Intel RealSense Developer Kit with a F200 camera. All of the presented experiments were performed on an Asus G551J laptop (Intel Core i7 4750HQ, 8 GB RAM, Nvidia GeForce GTX 960M, 64 bit Windows 10 Education operating system). The RealSense device is the VF0800 model equipped with the F200 camera. We used Intel Realsense SDK version 7.0.23.8048. The Leap Motion device is the LM-101 model, with the Leap Motion SDK version 3.1.2.40841.

Leap Motion

Leap Motion as presented in (Leap Motion Home) is a small device that incorporates two infrared cameras and three infrared LEDs. According to the developer (Leap Motion Specs) the interaction area of the device is 2 feet above the controller by 2 feet wide on each side (150 degrees angle) by 2 feet deep on each side (120 degree angle).

For best results, (Leap Motion Coordinates) states that the user’s hands are guaranteed to stay within the field of view of the device as long as they
are within an interaction box. This box has 235x235x147 mm dimensions (width x height x depth) and is situated above the Leap Motion controller, at a height preset in the device’s control panel or that is automatically determined by the device if the automatic option is chosen in the control panel).

**Intel RealSense F200**

Intel RealSense (Intel RealSense Home) cameras combine infrared optical tracking technologies - an infrared laser projector and an infrared sensor, according to (Intel RealSense Hardware) - with a 1080p RGB camera (color sensor) into a small device that combines 3D depth information with 2D images, in order to create a high quality 3D depth close range video (Intel RealSense Comparison). As presented in the same source, the F200 camera was replaced by a newer model, the SR300 in currently available developer kits.

Intel RealSense provides multiple functionalities, including 3D scanning, face detection, hand tracking, object tracking, etc. (Intel RealSense Comparison) (Intel RealSense Ranges). Out of these possibilities we are currently interested in the hand tracking features of the F200 camera.

As stated in (Intel RealSense Comparison), the recommended ranges for the hand tracking feature of the F200 are a distance between 20-60 cm away from the camera and a speed of maximum 0.75m/s.

### 2.2 Related work

From our research, related works assessed the precision of the tracking devices by using a robotic arm that can be precisely placed and moved from one point in space towards another through comparing the variations in the detected positions or movements compared to the established ones, by using a robotic arm with a pointing device (that the Leap Motion device has the ability to detect) placed at the end (Accuracy Robustness LM) or by placing a plastic arm model holding a pointing device at the end of the robotic arm, such as in (Analysis Precision LM). Both experiments evaluated the measurement precision in both static and dynamic scenarios. (Accuracy Robustness LM) discovered sub-millimeter precision in both static and dynamic scenarios, while (Analysis Precision LM) determined a sub-millimeter precision for the static experiments but stated that an inconsistent performance was obtained for the dynamic scenario.

We did not find extensive research regarding the evaluation of the
performances of the Intel RealSense. In (Approach to Physical Rehab) the authors developed a simple Virtual Reality system for medical rehabilitation of the shoulder using Kinect and Intel Realsense for movement tracking. In the preliminary evaluation, four experts working in the field of physiotherapy tested the system containing the two devices and reported that they found all the animations of the movement in the virtual world to be fluid.

In (Augmented World RS) 35 participants with various degrees of experience with a professional image editing tool were presented with a special graphics editor that was controlled by either classical user input (mouse and keyboard) or by using special gestures recognized by the RealSense. The participants completed a survey regarding the usability of the two user interfaces. The authors discovered that users that were more accustomed to the professional editing tool were more inclined to prefer the classical input methods, while novice users preferred the non-conventional input methods. All groups found the second one to be more engaging.

Another experiment assessing a possible use of the Intel RealSense is presented in (Biometric Hand Recognition). Here, the authors created 3D biometric hand recognition system. The results were stated to be comparable to some of those based on more expensive sensors.

All the aforementioned experiments show promising results regarding the precision of the two devices.

2.3 Considered movements

The TRAVEE system already implements four movements (finger flexion-extension, thumb flexion-extension, shoulder raise, forearm adduction-abduction) that are analyzed and for which a feedback is provided as a grade representing the degree to which the movement was executed.

In its final version, TRAVEE will include 10 rehabilitation movements, out of which only four involve movement of the fingers and are thus relevant to the subject of this article:

- Finger flexion-extension
- Thumb flexion-extension
- Fist flexion-extension
- Thumb touches

In the following paragraphs we will describe the correct execution of each of the four selected movements.
**Finger flexion-extension**

This exercise involves flexing and extending four fingers (index, middle, ring, pinky) while the thumb remains still. The correct execution of the flexion movement – as presented in (GetMedEdu Movements) - starts with the four fingers extended and consists of bending the fingers so that the angle between them and the palm is decreased. Oppositely, the extension part of this movement involves increasing the angle between the fingers and the palm.

![Figure 1. Finger flexion (left) and extension (right)](image)

We will also consider the execution of the flexion-extension movement for each finger, separately.

**Thumb flexion-extension**

This movement involves moving the thumb across the palm in the coronal plane towards the palm (flexion) and away from the palm (extension) (GetMedEdu Movements).

![Figure 2. Thumb flexion (left) and extension (right)](image)
Wrist flexion-extension
This movement requires that – while holding the fingers flexed – the wrist is moved towards the forearm (flexion) and away from the forearm (extension).

![Figure 3. Fist straight (left), flexed (center) and extended (right)](image)

Thumb touches
This movement (WebMd Movements) requires touching each of the fingertips with the thumb while holding the wrist straight.

![Figure 4. Thumb touches: index (1), middle (2), ring (3), pinky (4)](image)

3. Methods
As a disclaimer, the experiment does not precisely assess qualitatively or quantitatively the tracking abilities of the two devices. The subjects performed the exercises in front of the cameras of the devices and decided for each repetition, whether they felt immersed in the VE – thus they could have been “fooled” that the virtual hand was actually theirs. This method allows a great degree of subjectivism. Also, as the experiment was
performed by humans, and not by a robotic hand that could have maintained the same posture and executed the movements in an identical manner for each repetition – we are aware and insist that there is an important degree of human error involved. One aspect that must be mentioned is that – although they were instructed to keep their hand within the range of the device, it is obvious by the provided feedback (some loss of precision for certain movements, but in singular cases) that sometimes the tested users were not able to maintain their hand within that exact area. This will be taken into consideration when placing the devices for the installation of the TRAVEE system in a hospital environment.

In order to determine the most appropriate configuration for our system, we applied two evaluation methods that we believed are sufficient for a decision in our particular situation.

For the first method we established five classes of movements – five approximate angles to which we bent four fingers (index, middle, ring, pinky) and the thumb respectively – in order to observe the detection of movements of smaller or greater amplitude with each of the selected devices (30 repetitions of each movement).

In the second method, we executed each movement of the four movements of the hand included in the TRAVEE system and observed the accuracy with which the tracking follows the movement for a long period of time (50 repetitions). One execution was considered correct if the subject felt that the movement of the virtual hand represented his own to an immersive degree.

Performance evaluation

Each of the three subjects involved was requested to perform the specified number of repetitions for each exercise. For each exercise, he was then asked to provide a percentage representing the degree to which he believed to be immersed in the VE. A percentage of 100% represents a perfect rendering of the real movement onto the virtual avatar. A smaller percentage is a subjective measure of the correctness of the movement, as seen by the user.

For example, a user did not observe a difference between the rendition of the fully closed fist and a movement in which the fingers are bent but do not touch the palm. In this case, he considered that the fully closed fist was not correctly represented and approximated that the fist was closed only to a degree of 9/10 for each repetition, therefore he evaluated the precision of the rendition with a 90% degree of success.
Similarly, if the user detected a small additional movement in the VE for example, but that he thought did not disrupt the immersion sensation, he could choose to decrease the success rate by only a small percentage, to reflect that observation.

4. Device placement

We selected the positions of the devices so that the movement is performed in the distance range recommended for the devices, as we mentioned in their description. As for the actual location of the device, we chose it so that all the important elements of the movement would be visible to the IR cameras, for the device to observe the differences in depth – as we thought would be most appropriate.

Figure 5. Placement of the devices on a flat surface with less (top) or more (bottom) indirect lighting
More precisely, for the Leap Motion we performed the movements with the device placed on a flat surface, as can be seen in the image below. We respected the official recommendation that the hands should be perpendicular to the field of view (Leap Motion Best practices) for best tracking reliability.

Also, we wanted to see whether more or less (indirect) light will have any effect on the tracking reliability.

We performed the movements with the RealSense while it was placed on a flat surface, but with the hands perpendicular to its surface as well as a second set with the hands parallel to its surface.

Both devices were placed on the same flat surface, in an environment with sufficient lighting, but away from direct sunlight – in order to avoid interference in the infrared light spectrum.

We also tested the configurations separately, in order to avoid the interference with one another – while one of them was working, the other was disconnected from the computer.

We placed the hand performing the movements straight in front of the device, as we believed it is most suitable and we made 30 repetitions of each movement.

5. Tracking for movements with various amplitudes

As we are interested in the detection of the movement by the tracking devices, we considered that an approximate approach would suffice for determining whether various degrees of movement are detected by the devices.

Movement classes

As we mentioned, we established distinct sets of exercises for the group of fingers consisting of index, middle, ring and pinky fingers and another set of movements for the thumb.

Therefore, for the group of five fingers we established five classes of movement amplitude for the flexion of the fingers:
- Very small (XS) – the fingers are kept in a straight position and are bent with approximately 30 degrees.
- Small (S) – the finger is slightly bent (the angles between the distal and intermediate phalanges and the one between the intermediate and the proximal phalanges are of about 150-160 degrees) and the angle between the proximal phalanx and the metacarpal bone is of approximately 45 degrees.
- Medium (M) – the finger is bent in the shape of the letter “C” (between each two of the phalanges there is an angle of approximately 120 degrees)
- Large (L) – the finger is bent event further, the angle between the intermediate and the proximal phalanges is decreased to approximately 90 degrees, but the fingertip does not touch the palm.
- Full (F) – the fingertip touches the palm.

Each of these classes is represented in the figures below.

Figure 7. The established classes of movements: None (top left), Very small (top center), Small (top right), Medium (bottom left), Large (bottom center), Full (bottom right)
For the thumb, we established four exercises:

- Extending the thumb away from the other four fingers (in the Tables represented as “Close”) – bringing the thumb towards them to a position of the palm where all fingers are kept straight and pointing in the same direction;

![Figure 8. Extending and flexing the thumb toward the other fingers](image)

- Flexing the thumb towards the palm (the pinky) with three degrees of flexion:
  - Small (S): the thumb is brought from an extension (where it is placed in the same plane as the other extended fingers) to a position where it is straight and just above the palm, approximately perpendicular to the plane of the extended fingers;
  - Medium (M): the thumb is bent towards the pinky so that its distal phalanx is approximately parallel to the plane determined by the other four fingers;
  - Full (F): the fingertip of the thumb touches the palm.

![Figure 9. Extending and flexing the thumb with various degrees toward the pinky finger (1-extension, 2-small, 3-medium, 4-full)](image)

**Observations**

We experimented with each device in the following manner:

- For the movements of the four fingers and those of the thumb we tested both settings (more or less lighting);
- 30 repetitions of each class of movements, as presented in figures 7, 8 and 9;
- Counted the repetitions that were represented accurately in the VE (accurately was defined subjectively in this experiment – we considered that a movement was accurate if the user felt immersed in the VE while
performing it – he felt that the hand in the VE could have been his, no additional movement was apparent and the movement of the virtual hand was the one that the subject was testing at the time);  
- To observe the tracked movements in a VE we used one of the applications provided by the producer in order to eliminate possible third party errors in mapping movement onto avatar:
  - For the Leap Motion we used the AllHandModels scene from the Leap Motion Core Assets (Leap Core Asset).
  - For the Intel RealSense we used the Hands Animation (Unity) from the Intel RealSense SDK Sample Browser (RealSense Support).

**Leap Motion**

![Figure 8. The evaluated position for the Leap Motion, in both more and less lighting setup](image)

<table>
<thead>
<tr>
<th></th>
<th>Less lighting</th>
<th>More lighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>XS</td>
<td>S</td>
</tr>
<tr>
<td>User 1</td>
<td>100% *</td>
<td>100% *</td>
</tr>
<tr>
<td>User 2</td>
<td>100% *</td>
<td>100% *</td>
</tr>
<tr>
<td>User 3</td>
<td>100% *</td>
<td>100% *</td>
</tr>
</tbody>
</table>

*The measurements seemed to be very precise as long as the fingertips were straight above the controller – if the fingers were stuck together. If there was space in between the fingers then they could be held close to the edges of the field of view and the tracking was not affected.

**The user reported that he did not observe a clear visual difference in the representation of the fully closed movement (F) and the one where the large flexion (L).
Table 2: Visually observed performance of Leap Motion for the four types of movement of the thumb

<table>
<thead>
<tr>
<th></th>
<th>Less lighting</th>
<th>More lighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Close</td>
<td>S</td>
</tr>
<tr>
<td>User 1</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>User 2</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>User 3</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

*The user stated that the difference in the visual representation of the M and F movements of the thumb was not easily observable. The difference in lighting did not produce any differences that could be observed, therefore for the rest of the exercises we chose to test only in the setup with more lighting.

**RealSense**

![Image of RealSense setup](image_url)

Figure 8. The evaluated position for the RealSense, in both more and less lighting setup

Table 4: Visually observed performance of RealSense for the five classes of movements

<table>
<thead>
<tr>
<th></th>
<th>Less lighting</th>
<th>More lighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>XS</td>
<td>M</td>
</tr>
<tr>
<td>User 1</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>User 2</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>User 3</td>
<td>99.5%</td>
<td>100%</td>
</tr>
</tbody>
</table>

*For the movements with more amplitude, when the fingertips become less visible because they approach the palm, there seems to be a slightly decrease in the accuracy of the movement of the virtual hand.

**The user stated that the difference in the visual representation of the L and F movements of the fingers was not easily observable. The presence of more indirect light does not seem to affect in any way the tracking process.
Table 5: Visually observed performance of RealSense for the four types of movement of the thumb

<table>
<thead>
<tr>
<th>Less lighting</th>
<th>More lighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>Close</td>
</tr>
<tr>
<td>User 1</td>
<td>100% S 100%</td>
</tr>
<tr>
<td>User 2</td>
<td>100% S 100%</td>
</tr>
<tr>
<td>User 3</td>
<td>100% S 100%</td>
</tr>
</tbody>
</table>

Also, with the RealSense we observed that while holding the palm perpendicular to the surface of the device the tracking was not very precise, as we could see the device detected additional displacement of the fingers and became unstable during movements. The users perceived that a much more coherent movement seemed to be represented in the VE when the hand was placed parallel to the surface of the device, similarly to the Leap Motion.

5. Tracking the TRAVEE movements

Since the variation in lighting did not have an observable influence on the tracking ability of the two devices, we tested each of the four movements only in the setup with more light.

Finger flexion-extension

Table 6: Visually observed performance for the finger flexion-extension movement

<table>
<thead>
<tr>
<th>User</th>
<th>RealSense</th>
<th>Leap Motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>80%</td>
<td>100%</td>
</tr>
<tr>
<td>User 2</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>User 3</td>
<td>100%*</td>
<td>100%*</td>
</tr>
</tbody>
</table>

* The user stated that the fingers did not appear to be touching the palm although in reality they were fully flexed.

Thumb flexion-extension

Table 7: Visually observed performance for the thumb flexion-extension movement

<table>
<thead>
<tr>
<th>User</th>
<th>RealSense</th>
<th>Leap Motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>100%*</td>
<td>100%</td>
</tr>
<tr>
<td>User 2</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>User 3</td>
<td>90%**</td>
<td>90%**</td>
</tr>
</tbody>
</table>

*Perfect performance if the fingers are spread. If the four fingers (index, middle, ring, pinky) are stuck together, very often for a thumb flex it detects another finger flexing as well.
** The user stated that the thumb did not appear to be touching the palm although in reality it was fully flexed.

Wrist flexion-extension

Table 8: Visually observed performance for the wrist flexion-extension movement

<table>
<thead>
<tr>
<th>User</th>
<th>RealSense</th>
<th>Leap Motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>100%*</td>
<td>100%</td>
</tr>
<tr>
<td>User 2</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>User 3</td>
<td>98%*</td>
<td>98%*</td>
</tr>
</tbody>
</table>

*The detection of this movement seems to be of 100% as long as the movement is executed with the palm facing the device and the fingers spread. If the fingers are touching or if they are held into a fist the accuracy is diminished.

Thumb touches

Table 9: Visually observed performance for the thumb touches movement

<table>
<thead>
<tr>
<th>User</th>
<th>RealSense</th>
<th>Leap Motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>100%*</td>
<td>90%**</td>
</tr>
<tr>
<td>User 2</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>User 3</td>
<td>99%*</td>
<td>95%**</td>
</tr>
</tbody>
</table>

* The RealSense has perfect accuracy for this movement, as long as the hand is positioned as discussed (parallel to the field of view of the device).
** For the Leap Motion all the misses were observed for the thumb touch of the middle finger.

6. Observations

From observing the accuracy of the tracking process (visually – comparing the movements in the real world with those indicated by the visualization software provided by the producers) and by the feedback we received from our subjects, we determined that for the movements that must be tracked by our system, the best placement for the Leap Motion device is the classical one, with the hand directly above the controller, especially if the movement involves keeping some of the fingers together. For the RealSense we think the best placement would be a similar one, where the controller is placed directly under the hand.

We also observed that the two degrees of lighting that we tested did not interfere in any way (that could be observed in the visual representation of the movements of the avatar hands) with the precision with which our specific exercises were tracked.

For the gradual exercises that we classified into five degrees, we noticed that both the devices were able to detect each movement as well as to track
the movement with a satisfactory precision. By holding the palm approximately parallel to the RealSense it would seem that hiding the fingertips, by performing some of the exercise movements may slightly interfere with the tracking process.

The TRAVEE movements were represented quite well in the VE in the configuration that we tested. The RealSense seemed to add some additional visual movement for the finger flex movement and the Leap Motion did not always detect the thumb-middle finger touch.

Details were provided alongside the tables dedicated to the movements for which the user felt that the degree of immersion was less than 100%.

7. Conclusion

By performing these experiments we found a placement for each of the two devices that provides good tracking capabilities. For this purpose, we performed some experiments that were based exclusively on visual observations. Given the fact that the TRAVEE system must be highly immersive we believe that an accurate representation of the real movements in the VE is critical and that the purpose of such accuracy is to have a credible aspect – therefore this was our main expectation when choosing the optimal configurations for our situation.

As we did not mean to obtain a qualitative measurement or comparison between the two, we believe our goal of finding what works best for each of the movements in our system was accomplished. Further studies are necessary for an accurate evaluation of these tracking devices.

We also found out more information regarding the way the two devices track our desired movements and the opinion of a healthy subject that tries to perform them while being immersed in a VE.

References


Analysis Precision LM. An Analysis of the Precision and Reliability of the Leap Motion Sensor and Its Suitability for Static and Dynamic Tracking. J. Guna, G. Jakus, M.


May 2016.