HCI project: Does touch feedback improve the user's accuracy when composing VR paintings on a 2D surface?

Group member: Raphaël Demoulin, HAN Xinpei

Presentation of the scientific question addressed

In recent years, the development of VR has been remarkable, and its use is getting more prevalent with a lot of VR applications that are being designed. Therefore, it is necessary to explore the human interaction involved in using these VR applications, by defining the form and effectiveness of the interaction, and how to improve that.

In this project, our goal is to explore if adding stronger touch feedback will improve the user's capability, and more specifically, the user's accuracy when composing VR paintings on a 2D surface.

However, since VR drawing is quite a new technology, and so drawing in a virtual space by moving the user's body without any support, i.e. mid-air movement, might be less accurate and is performance-limited.

To answer this question, we designed an experiment to test in the case of using VR for 2D drawing to see if giving the touch feedback of a pen and a hard surface, close to real-world drawing, will improve the user's drawing efficiency and experience.

Rapid overview of the related work

We have read multiple papers related to this topic, but we will only talk about the two most important to our experiment. The first one is the bases of our experiment, and the second one we found after we finished our experiment design, confirmed it.

The first paper is “Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks” by Johnny Accot and Shumin Zhai. This paper introduces the idea of Steering law by using the Fitts' law index of difficulty (ID). The idea is to create more than one goal to pass and compute the Fitts'ID of this new task, before adding another goal. Then, by iterating this addition toward infinity, we have an infinite number of goals which create a tunnel that become the Steering task. This gives us the Steering law ID which is derived from Fitts' law ID, that is used the same way as doing Fitts' task. One key difference between the two task come from the fact that the Steering task doesn't require to have the drawn line to go back and forth, and just needs to go through the tunnel once.
The second paper we found is “Experimental Evaluation of Sketching on Surfaces in VR”,by Rahul Arora and al., which presents the experiment as we wanted it to be. The main difference between this paper and our experiment is the use of the Steering law as the basis of comparison.

Solution

In order to research if touch feedback can improve user accuracy, we designed an experiment to collect data and analyze the results to see if they indicate some conclusion.

Introduction of the ideal experiment:

The ideal experiment we designed in the first place is to have each of the participants draw multiple lines with the VR device inside shapes of different difficulty, with respect to the Steering law. This task has to be done in two conditions.

The first case is the control test. This means it doesn't have the additional feedback we want to test, only the one of the VR device and 2D drawing software we will use.
The second case is to do the same drawing with the same condition, but this time by holding a pen and having a hard surface to lay on while doing the task.
The implementation of those additional would have been done by using a pen shaped object that would have been tracked to act as the paintbrush inside the VR drawing software while using a wall to lay on for the drawing.
During the experiment, for each shape in both cases, we track the time of completion of the Steering task and save the drawing in order to analyze the accuracy.

The pre-requisite for setting up the experiment are, a VR headset, a VR 2D drawing application called "V-Art Painting Studio" on Steam, 10 Steering tasks with increasing difficulty indexes to provide scales for evaluation, a hard surface, and a pen for touch feedback.

For the experiment participants, we invited 14 users, most of whom were young computer science students who seldom draw and never received professional drawing education (we didn't count the middle school art course as professional education).

We had to make adjustments on this experiment due to devices and time limitations. We will give detailed instructions on that in the next section. Also, the biases that might have been introduced by the adjustments and the possible ways to solve them.

The adjustments of the experiment:

We had the choice between two VR devices. The first one is the HTC Vive Pro 2 kit (headset and controller), and the second one is a Meta Quest 3 kit (headset and controller).

To choose one, we decided to define how to add the pen feedback and use the one on which it will be faster to implement. Therefore, we want the user with a VR device to hold a pen, see it, and use it. We had the possibility to combine the controller with a pen, but this approach would change the position and the way of the user using the pen, due to the weight of the controller. The other possibility was to generate the pen's position digitally inside the VR world. For this, we needed to use the Optitracks cameras to capture a pen-shaped object compatible with the camera in order to generate its position and motion. However, due to the time limitation, we decide not to apply this method.

Instead of doing the steps mentioned above, we simplified our experiment by taking advantage of the passthrough functionality of the VR device. VR passthrough is functionality that allowe the during a virtual reality experience by using a set of camera on the front of the helmet.

We printed out the shapes of the experiment, making sure that we respected the scaling to keep the same difficulty indexes for the Steering tasks between the digital and real ones. By using the passthrough to stay close to a virtual environment, and a real pen, they were able to draw with the pen feedback.
In ideal conditions, the participants should wear the same VR device for the 2 cases, but we had to use both of them to take advantage of the Meta Quest passthrough and HTC Vive steam support.

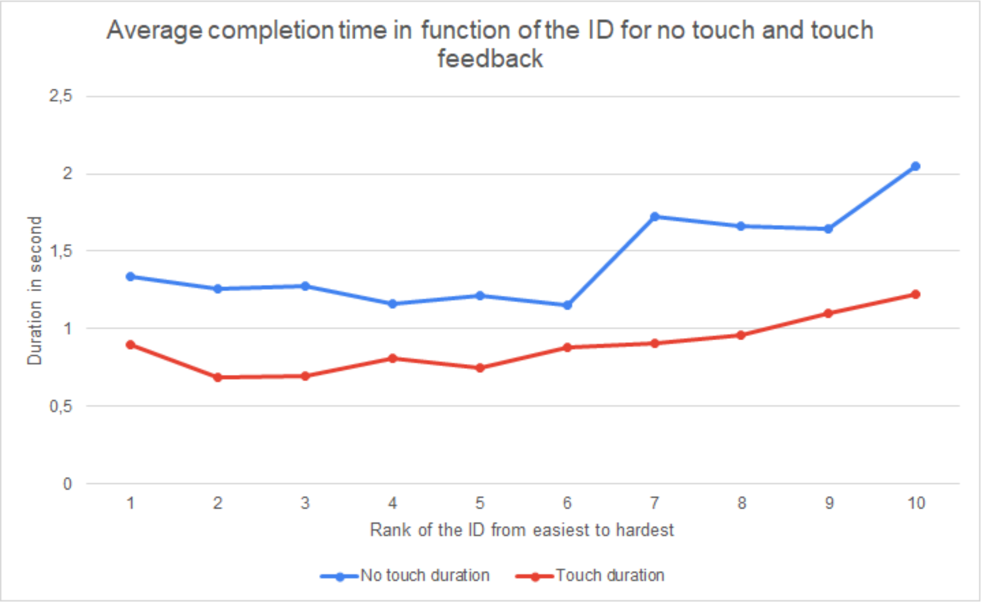
For the hard surface feedback, we just used the table on which we put the printed Steering tasks.

Evaluation

For each Steering task, we ask the participant to draw three lines from the wider opening to the smaller one, without crossing the lines of the figure. Using the stopwatch, we take the time of completion of each lines which we associate to the corresponding student and the index of difficulty.
The goal is to compare the average completion time of all participants on a given index of difficulty in each case. We then compare the time between the 2 cases to see if, the user is faster or not with touch feedback. We then compare the drawing of both cases to see if there are less mistakes made when adding the touch feedback. The positive answer to our research question can be indicated in the case that all the participants have better average results when drawing with touch feedback.

Result

Of the 14 participants, only 6 datasets were trustworthy due to the inconsistencies in the collection of time. Since the data was collected using the stopwatch function of the phone, with the time stamp used to define the duration for passing the goals being the lap button. (Further explanation in the next part).
Using the 6 datasets, we can make the following graph :



The graph is made by computing the average completion time of each participant for each index of difficulty (ID). Note that, the graph represents the rank of difficulty of the ID and not the value of it, i.e. the ID of rank 1 and 2 have for value 1,92 and 2,51 (see annexe for all difficulty).
We compute separately the tasks with and without touch, which generate 2 curves.
The blue curve represents the experiment without pen and table feedback. The flat part at the beginning of the curve is due to the narrowing tunnel of the first 5 rank are wide enough for the participant to go through them without much effort. The 6 and 7 one, participant nearly did or did 1 mistake in average, which slowed them down.

The red curve represents the experiment with pen and table feedback, which was done right after the task without feedback. We can see that the curve have a steady increase in time as the difficulty increase. This trend is what we expected to see, as it is the trend shown in the Steering law paper. We can also notice that the participant aren't surprised by the tunnels since they have experience with the previous task.

Overall, we can see that the 2 curves clearly show that the pen and table feedback have an impact on the efficiency of the drawing. But due to the bias of the following part, it is hard to state if it is really the addition of feedback that improved the efficient or some other parameters.

For the evaluation on the improvement of accuracy, there isn't much that is noticeable. Indeed, over the 14 participants, most of them made a mistake on the ID of rank 7, but there is a non-negligible chance it was due to the surprised in increase of difficulty than a real accuracy mistake. This idea is supported by the fact that there was often no mistake after this one. The one that happened are more linked to the trouble of using VR in general (loose of balance, overshoot when moving with the pad, etc)

Biases and what to improve:

1. Group of participants: most of them are computer science young students, who know better about VR, and learn how to interact with machines faster and easier than general people. Also, this group is too biased to be considered as a representing group. We would need more different backgrounds, and ages with a good ratio between males and females.

2. Use two different VR devices for experiments. Indeed, as previously mentioned, we used the Vive Pro 2 since it is stable for using the drawing application, but its passthrough has a very bad resolution. Meta Quest 3 has a good quality passthrough and should be able to use the Steam application, but the link worked on the local network, which led to inconsistency in the use of the drawing software. In the end, we decided to take advantage of both devices. We used Vive Pro 2 for VR drawing and Meta Quest 3 for pass-through. Without limitation of the resources, the best is to use only one VR device for the whole process to avoid the extra variable introduced in the experiment.

3. The VR controller gives vibration feedback. This created a condition to the VR case of the experiment that we couldn't reproduce in the second case. It would have been better to have a similar condition, meaning both cases shouldn't have any vibration feedback.

4. Manually measuring the time and make mistakes. For each of the participants, we count the total time from the moment they begin to draw until the passing of the left goal. We have 10 drawing tasks and for each of them the participant should draw 3 times to have an average rate, so in total we should have 60 records for each participant under each experimental condition. Since we counted the time manually, there are miss-clicks that leads to a number of data different from 60. We invited 14 participants, but in the end, only 6 participants had valid data. We tried to restore the data, by comparing the time of completion with the one that are supposed to be on the same index of difficulty (ID) to find out the inconsistencies to ignore the faulty ID. But this wouldn't make the data trustworthy, so we decided to abandon the data of those that don't have 60 measures in both cases. This leads us to use 6 valid datasets out of 14. It will be more scientific and have more valid data sets if we can improve the way to count the time. The best way would have been to record the experiment in both cases and put time stamps on the resulting video to obtain consistent data.

5. The canvas in the VR environment is vertical, and the table in real life is horizontal. We didn't think it would have much effect, but nearly all participants made a remark on that. Also, by re-reading the paper of Rahul Arora, we can see that the results are different depending on the inclination of the drawing surface. The virtual canvas was also too high for some participants (One had to be on the foot tip to draw), so this also can have influence on the result. We should make sure the virtual canvas and the hard-surface are in the same position and with accessible size.

6. The drawing task we designed only has one shape but with different difficulty indexes, it will be more relevant if we can have more different shapes with different difficulty indexes.

Difficulties

The experiment has to be done in a limited time amount of time that clashes with time required in other courses. Also, different groups from our course and from other courses had to share the VR headset, so it really constrained the testing time necessary to remark mistakes and biases.
We only have two VR devices as choices, and both of them have pros and cons. Sometimes, they were buggy, so the experiment couldn't be done in ideal conditions.
Doing the ideal experiment required technical skills, which we didn't have time to develop, as they were too difficult to resolve.

Summary

This project was interesting as it was the first time we realized an experiment from scratch. There were a lot of unexpected problems that came up as we progressed the project, and the methodology to do an experiment with people is surprisingly complex. Noticeably, a very surprising thing comes from the fact, an explanation that work for someone can poorly work for someone else. It showed us too late the necessity to do a training experiment with the participant to ensure he understands it and also avoid the bias of surprise. This of course only works in the case that doesn't involve discoverability.