Abstract

In our project we studied the usability of a new multimodal virtual drawing technology and the potential of its artistic expression through the creation of sketching tool for artists. We examined if multimodality encourages more artistic expression from users, and its effect on the user experience. The project was inspired by Pablo Picasso’s projected light drawings from 1949. Picasso couldn’t observe his drawing in real time with the technology available at the time. We have developed a tool with which the drawer can see the strokes of their drawing in real time.

An usability test was conducted in a Visualization studio (VIC) at KTH University with real time live feed from a laptop computer back-projected through a projector. The users controlled a 8 LED flashlight to track the light. A piano sound was added as an additional modular, where the different coordinates of the light triggered different MIDI notes. We studied whether feedback of sound has a positive effect on artistic expression and motivates people to draw more.

The project concludes that we were able to make recommendations regarding the possible future studies of hand controlled virtual drawing. Our study found that there was no statistical significance between having sound and not having it. This study might give useful information in Virtual Reality Drawing and Augmented reality drawing in the future.

Keywords: Free-hand drawing, Multimodality, Usability, Artistic expression, Sketching.
1 Background

This project started when we saw a video from James Vincent’s article [1] about Glen Keane drawing in virtual reality. We wanted to experience this technology and thought of how it could be used. For example, as a virtual collaboration tool for students, a virtual sketching tool for designers or a tool for artist who want to create new kinds of art forms using sculptural drawing as a tool. Inspirations were also drawn from Picasso’s famous “Light Drawings” [2], where the images captured shows the artist drawing with a small electric light in a dark room. In his art, he was known for his excellent use of minimalistic lines to create highly expressive pieces. In these “Light Drawings”, his figure is shown in full concentration and it was almost as if the process of the creation of the art piece was reflected. We are not as talented as Picasso, drawing with light while imaging the painting in mind still seems magic for us. But why not realize the magic with our modern technologies? The biggest challenges for us to draw in Picasso’s way is that we can not imagine a painting in the air while seeing nothing-memory leaves following the light. So if we can figure out a way to presenting the painting, we will approach Picasso’s magic. To better understand the related field in our chosen project, we did some research on previous relevant work. Deering [3] introduced HoloSketch, a system which supports a variety of editing operations (i.e. movement, grouping, scaling and attributes). It was interesting to note that it mentioned most of the 2D and 3D drawing packages featured an array of screen buttons and menu selectors at one corner of the drawing space. McGill [4] also highlighted that interacting with reality is a challenge in VR Head-Mounted Displays, in that being aware of the presence of others in the real environment compromises with the virtual immersion the user experiences. In order to bridge this barrier to consumer adoption, the study suggests incorporating minimal amount of reality in, enabling users to be more aware of their real surroundings while still being immersed in virtual reality. Shin, Dunston and Wang [5] explored the two mechanism of view change in a 3D virtual environment, with it being observer movement, where the user explores the object by moving around the object, as compared to model rotation, where the user explores the object by rotating it. It concludes that a user’s perception of 3D objects in an augmented reality environment depends on the type of view change, and that its results reflected that performance was better for observer movement than that of model rotation. Keefe, Feliz, Moscovich, Laidlaw and Laviola [6] showed the

Figure 1: “An angel” Examples of art created during our study. The participants created several pieces of art using our sketching tool for designers.

Figure 2: “A Chinese Fish” Example of art created during our study. A Chinese fish was created during the experiments in our study.
system that provides artists the approach of using 3D analogy of 2D brush strokes to create 3D art works in an immersive cave environment. It is a big challenge to build a virtual reality system with which artists are able to interact naturally. The CavePainting system completed the task by showing that artists having minimal experience with computers had no trouble learning the gesture-based interface.

Mäkelä, Reunanen, and Takala [7] introduced a 3D painting system in its early stages, with the main goal of exploring the program’s usability and artistic potential. The usability testing was assessed through both given tasks and free artistic sketching, and these were recorded through observations and interviews. They also found that common technical limitations of virtual environments were reflected in the results, with factors such as latency and tracking inaccuracy caused “inaccurate cursor placement, shaky strokes and overall unresponsiveness of the system”, whereas the clumsiness of hardware devices caused some restrictions to the user’s fine motor movement, hence hindering his handicraft work.

All the studies explored the possibilities and potentials of drawing or other manipulations in virtual reality from different directions. The results showed not only that abundant possibilities hold in drawing in virtual reality, but also that by approaching in different ways, unexpected achievements may appear. With all these in mind, we camp up the idea that probing the use of light in virtual reality drawing and also adding other modalities like sound to examine any significance could be made while drawing due to the new feature.

2 Method

2.1 The experiment

Our goal was to study if multimodality (the feedback of sound) in virtual drawing motivates people to draw more and whether multimodality has a positive effect on artistic expression.

2.2 Experimental setup

The experiment was conducted in the Visualization studio (VIC) at KTH. The real time image from laptop computer was back-projected through one projector (JVC D-ILA 4K) to a big 4K screen. The computer used its integrated web camera which was situated in front of the 4K screen. The sounds used Yamaha DM1000 sound mixer and were played from four loudspeakers using mono sound with an output sound intensity of around 70 dBA. The user handled a one hand 8 LED flashlight (NEBO Larry™ 8 LED Pocket Work Light) that we called a “lightsaber”. The flashlight had a bright and long shaped light source. We chose to use this, because it allows for the user to regulate the amount of light. It also had a button to switch the light on and off.

2.3 Design and procedure

We set up a stationary environment in a semi dark room, where users still managed to see their surroundings. Here, users were able to draw and explore the canvas with the use of a torchlight. The drawing occurred when the torch was switched on and a light was detected against the dark background, leaving its mark as the user moves the torch around in the canvas.

For this, we did the programming in “Max 7”, a visual programming software developed by Cycling '74. The program itself was ran on a MAC OS X. In the Max patch, what happens is that live video input is received from the webcam through the object “jit.grab”, where it digitize videos and decompresses the signals into a Jitter matrix. From the video, we then decide to track the color white as it provides the best contrast against a dark background. The light coordinates is then tracked from “jit.findbounds”, where it calculates the bounding dimension for a range of values indicated. In this object, it scans a matrix of values in the range, sending the minimum point value as a list of numbers out from the left outlet, and the maximum point value as a list out the second outlet. From there, we randomised the colors from the coordinates detected, combining the drawing image together with the live
video using “jit.alphablend”, where the alpha channel of the input matrix (live input video from webcam) is used as a per-cell crossfade value, and is blended between the input matrices in the left and right inlets (the drawing).

To add an additional modular to our program, we included sound, where the different coordinates of the light would then trigger different MIDI notes. Here it mimics a basic midi controller with the tracking of the light, and various thresholds were implemented to keep the notes in a range, where the x-coordinates manipulate the pitch and the y-coordinates manipulate the frequency. This is done with the object “noteout”, where the coordinates of the light are packed and transmitted as a value to a MIDI device.

2.4 User tests

After setting up the program, we designed an usability testing with a general aim to understand if our users can use the product well, to assess the overall effectiveness of our prototype and to identify the obstacles and challenges faced while using the prototype. But more specifically, we are interested if the multimodality, from the addition of sound, add anything to the user experience of our prototype.

The 10 test users were students from KTH, that belonged to different majors. Half of them were exposed to the program with the additional modality of sound, and the other half, with only the visual modality. Once in the dark room, they were brief on the study and given a consent form to sign (refer to Appendix A). They were then given brief instructions on how the light would enable them to draw. They were then given a minute to explore freely on the canvas. Following that, the users were given 4 different tasks (drawing a straight line, basic geometric shapes, connecting the dots on the screen and drawing an animal of their choice), one after another once they completed the given task on hand. The users were also briefed to voice out their thought process as they work their way through the tasks. After the experiment, the users were asked a question: “The test is now finished. Would you like to continue drawing a little bit more?” To further understand how the users thought of the program, they were given a questionnaire at the end of the testing (refer to Appendix B).

2.5 Result methodology

Participants were divided in two sets, with five participants in each set. Each participant was to indi-
individually do a series of tasks, and give their subjective opinion on their experience with using the tool. The first set of participants did the tasks with sound, and the other set without, to see if introducing an additional modality (sound) would have an impact on the subjects response to the tool. The participants were asked to rate their experience with six choice questions (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree or Strongly Agree, which for the purposes of analysis will be assigned the numerical values 1 to 5) along with four free-text questions.

We used the Rank-sum test [8] (also known as “Mann-Whitney U test”, “Wilcoxon-Mann-Whitney test”, or “Wilcoxon Rank-sum test”, to interpret the data for the first six questions, since this was likert-scale type data and was assumed to be ordinal and might not have followed a normal-distribution. The null hypothesis was that there was no difference with sound or without. We chose a 5% ($\alpha = 0.05$) level of statistical significance for a two-tailed test.

### 3 Results

<table>
<thead>
<tr>
<th>Question</th>
<th>With Sound</th>
<th>Without Sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The interface was easy to navigate.</td>
<td>(3, 3, 3, 4, 5)</td>
<td>(2, 3, 4, 4, 5)</td>
</tr>
<tr>
<td>2. I think most people would figure out how to use interface this very quickly.</td>
<td>(4, 4, 4, 5, 5)</td>
<td>(4, 4, 5, 5, 5)</td>
</tr>
<tr>
<td>3. Figuring out how to draw with this interface was difficult.</td>
<td>(2, 2, 2, 3, 4)</td>
<td>(1, 2, 2, 2, 3)</td>
</tr>
<tr>
<td>4. I think I would need to ask questions to know how to use this interface.</td>
<td>(2, 2, 4, 4, 5)</td>
<td>(1, 1, 2, 2, 3)</td>
</tr>
<tr>
<td>5. This has integrated the basic notion of drawing.</td>
<td>(2, 4, 4, 5)</td>
<td>(4, 4, 4, 5)</td>
</tr>
<tr>
<td>6. I would be interested to use this interface again.</td>
<td>(3, 4, 4, 4, 5)</td>
<td>(3, 4, 4, 4, 5)</td>
</tr>
</tbody>
</table>

#### 3.1 Analysis of quantitative results

We used the Rank-Sum test, which meant that we first sorted the responses according to score, and then ranked them in order, if responses shared score, their rank was the average rank of all responses with that particular score. We then got the total rank-score with:

$$U_V = \sum_{n=1}^{\mid V \mid} V_n - \frac{1}{2} \times \mid V \mid \times (\mid V \mid + 1) \quad (1)$$

The lowest U score was then compared to the critical value of the critical z-value table [9].

**3.1.1 Question 1, The interface was easy to navigate with**

$S_{rank} = \sum_{n=1}^{5} S_n = 1.5 + 6 + 6 + 6 + 6 = 25.5$

$U_S = \sum_{n=1}^{5} S_n - \frac{1}{2} \times 5 \times 6 = 10.5$

$T_{rank} = \sum_{n=1}^{5} T_n = 1.5 + 6 + 6 + 6 + 10 = 29.5$

$U_T = \sum_{n=1}^{5} T_n - \frac{1}{2} \times 5 \times 6 = 14.5$

The smallest value, 10.5 was selected for Critical value test. The critical value in this case ($\mid S \mid = 5$ and $\mid N \mid = 5$) was 2. Since 10.5 was not less or equal to the critical value, we had to accept the null hypothesis in this instance.

**3.1.2 Question 2, I think most people would figure out how to use interface this very quickly**

$S_{rank} = \sum_{n=1}^{5} S_n = 2.5 + 2.5 + 2.5 + 7.5 + 7.5 = 22.5$

$U_S = \sum_{n=1}^{5} S_n - \frac{1}{2} \times 5 \times 6 = 7.5$

$T_{rank} = \sum_{n=1}^{5} T_n = 2.5 + 7.5 + 7.5 + 7.5 + 7.5 = 32.5$
\[ U_T = \sum_{n=1}^{5} T_n - \frac{1}{2} \times 5 \times 6 = 17.5 \]
The smallest value, 7.5 was selected for Critical value test. The critical value in this case (|S| = 5 and |N| = 5) was 2. Since 7.5 was not less or equal to the critical value, we had to accept the null hypothesis in this instance.

3.1.3 Question 3, Figuring out how to draw with this interface was difficult

\[ S_{rank} = \sum_{n=1}^{5} S_n = 4.5 + 4.5 + 4.5 + 8.5 + 10 = 32 \]
\[ U_S = \sum_{n=1}^{5} S_n - \frac{1}{2} \times 5 \times 6 = 17 \]
\[ T_{rank} = \sum_{n=1}^{5} T_n = 1 + 4.5 + 4.5 + 4.5 + 8.5 = 23 \]
\[ U_T = \sum_{n=1}^{5} T_n - \frac{1}{2} \times 5 \times 6 = 8 \]
The smallest value, 8 was selected for Critical value test. The critical value in this case (|S| = 5 and |N| = 5) is 2. Since 8 was not less than the critical value, we had to accept the null hypothesis in this instance.

3.1.4 Question 4, I think I would need to ask questions to know how to use this interface

\[ S_{rank} = \sum_{n=1}^{5} S_n = 4.5 + 4.5 + 9 + 9 + 9 = 36 \]
\[ U_S = \sum_{n=1}^{5} S_n - \frac{1}{2} \times 5 \times 6 = 21 \]
\[ T_{rank} = \sum_{n=1}^{5} T_n = 1.5 + 1.5 + 4.5 + 4.5 + 7 = 19 \]
\[ U_T = \sum_{n=1}^{5} T_n - \frac{1}{2} \times 5 \times 6 = 4 \]
The smallest value, 4 was selected for Critical value test. The critical value in this case (|S| = 5 and |N| = 5) is 2. Since 4 was not less or equal to the critical value, we had to accept the null hypothesis in this instance.

3.1.5 Question 5, This has integrated the basic notion of drawing

\[ S_{rank} = \sum_{n=1}^{4} S_n = 1 + 4 + 4 + 8 = 17 \]
\[ U_S = \sum_{n=1}^{4} S_n - \frac{1}{2} \times 4 \times 5 = 7 \]
\[ T_{rank} = \sum_{n=1}^{5} T_n = 4 + 4 + 4 + 8 + 8 = 28 \]
\[ U_T = \sum_{n=1}^{5} T_n - \frac{1}{2} \times 5 \times 6 = 13 \]
The smallest value, 7 was selected for Critical value test. The critical value in this case (|S| = 4 and |N| = 5) was 1. Since 7 is not less or equal to the critical value, we had to accept the null hypothesis in this instance.

3.1.6 Question 6, I would be interested to use this interface again

\[ S_{rank} = \sum_{n=1}^{5} S_n = 1.5 + 6 + 6 + 6 + 6 = 25.5 \]
\[ U_S = \sum_{n=1}^{5} S_n - \frac{1}{2} \times 5 \times 6 = 10.5 \]
\[ T_{rank} = \sum_{n=1}^{5} T_n = 1.5 + 6 + 6 + 6 + 10 = 29.5 \]
\[ U_T = \sum_{n=1}^{5} T_n - \frac{1}{2} \times 5 \times 6 = 14.5 \]
The smallest value, 10.5 was selected for Critical value test. The critical value in this case (|S| = 5 and |N| = 5) is 2. Since 10.5 was not less or equal to the critical value, we had to accept the null hypothesis in this instance.

3.2 Analysis of qualitative results

In addition to the quantitative questions, the participants were also asked some more qualitative questions, what follows is a short summary:

“Which task did you find it hardest to complete and why?”

- The hardest task was to draw an animal since it needs some sharp details to look good. It is hard to draw something sharp with the stick.
The circle, it was hard to end up where you started!

Connecting the dots, because they were near the edges of the screen and a bit hard to reach.

To draw a straight line.

The results suggest that different people thought different tasks were the most difficult, which would be encouraging, since it would imply that there is no ‘single’ point of failure with the process.

“What problems did you face while going through the tasks, and how do you think these can be improved?”

• Drawing thinner or thicker lines was not immediately understood.

• Depending on the distance to the camera it seemed like the view got distorted.

• Some light reflections on the floor.

• I can not draw a very straight line. Maybe some kind of automatic correctness can be utilized.

“What do you think can be improved in your experience with the prototype?”

• Easier on-off switch to make interruptions during the drawing.

• Give it a better connection with the screen.

• Maybe some "anti-shake" control which makes it easier to draw straight lines.

• Maybe an undo button instead of clearing the entire screen.

• Better on/off button on the light sable.

“If you could make a significant change with the interface, what would you do and why?”

• Make uniform the thickness of the line according to the position.

• Making it easier to clear the screen, some kind of eraser function perhaps.

• Maybe showing a visible pointer where you are aiming.

In summation of the points above, most users wanted a way to enhance the precision of the tool, to make it easier to accomplish the given tasks and to make it easier and more intuitive to use.

4 Discussion

Mäkelä, Reunanen, and Takala [7] study found technical limitations of virtual environments, with factors such as latency and tracking inaccuracy caused “inaccurate cursor placement, shaky strokes and overall unresponsiveness of the system”, whereas the clumsiness of hardware devices caused some restrictions to the user’s fine motor movement, hence hindering his handicraft work. Our results reflect this common concern, with most users wanting a way to enhance the precision of the tool, in order to make it easier to accomplish the given tasks and to make more intuitive to use. Although our study didn’t use virtual reality helmets, many users experienced similarities with Mäkelä, Reunanen, and Takala [7] study such as clumsiness and low latency.

Our study didn’t find significant results if the feedback of sound in virtual drawing motivates people to draw more or whether multimodality has a positive effect on artistic expression. This lack of statistical significance could be due to the result of a small sample size, and therefore lack of information gathered to contribute to a statistical significance. Future studies could look into this with a larger sample size as some recent studies [10] suggest that the activity of visual cortex can be modulated by sound. Some experiments [11] [12] [13] also prove that the perception of objects is a cross-modal process, and sound can enhance the detection of visual imagery and sometimes alter the perception of a visual result. In our study, four people out of ten participants, where two people with sounds and two people without sounds wanted continue drawing after we asked if they are eager to explore the interface more. There was no significant evidence if the sound made them to continue drawing or not. Further studies could also be made with a
questionnaire directly related to the experience with sounds.

The two modalities studied in our project were sound and vision. Movement of a hand added and added impression of depth while drawing was also incorporated. All the users internalized the use of the new media relatively quickly after a debriefing study session. Our experiment also suggested that the participants enjoyed the drawing tasks. The sounds played while the participant was drawing seemed not to have big difference in their artistic expression. The participants seemed to be fully immersed in using the new tool with and without sounds.

Future studies may consider changing the drawing sound from a piano key to a more relevant sound, for example, to the sound of a swiping pencil or paintbrush as it might give a different result due to the perhaps fuller experience. One limitation in our project was that the room had to be dark enough in order for the program to work well. The darkness might have interrupted the user’s interaction with the tool while seeing themselves from the screen at the same time. A lighter room might help the user to see their silhouettes more clearly and perhaps contribute to a more pleasant drawing experience with them being more aware of their surroundings.

References


