CyberSketch

Multitouch Tablet-Based Rich Interaction as Input for Mixed Reality 3D Sketching

Course: CIU235 Individual Project in Interaction Design

Supervisor: Daniel Sjölie

Ajla Ćano & Daniel Roeven

Spring 2018

Abstract

We present CyberSketch, an exploration of the possibilities of accurate and rich multitouch tablet-based input for MR, corresponding to a synchronized reality setup. We have done so by designing a 3D sketching tool which demonstrates the possibility of modeless, natural, and gestural input. Contributions made are proving the feasibility of practical setup of tablet-based input for MR. We also demonstrate the possibility for rich tablet-based interaction through an approach to multitouch tablet-based 3D sketching. Finally, we identify a viable approach to two-handed multitouch interaction for MR and its possible use cases, and we describe the need to communicate the value of multitouch input for MR to users.

Contents

1 Introduction **1**

2 Related Work **2**
2.1 The Sketching Process 2
2.2 3D Sketching Tools 2
2.3 Design Rationale & Interaction Style 4
2.3.1 Interaction Styles 4
2.3.2 Current VR Input Techniques 5
2.3.3 Multitouch Input 6

3 Concept **8**

4 Process **9**
4.1 Design Log 9
4.1.1 Ideation Phase 9
4.1.2 High Fidelity Prototype 12
4.1.3 Video Prototype 13
4.2 Feature-level breakdown 14

5 Final Version of CyberSketch **18**

6 Evaluation **20**
6.1 Methods 20
6.2 Results 21
6.2.1 User Experience In General and Usability 22
6.2.2 Requirements, freedom, exploration and expressiveness . . 23

7 Discussion **24**
7.1 Learning outcomes 24
7.2 Limitations 25
7.3 Future work 25

8 Conclusion **26**

References **27**

1 Introduction

AR technologies are quickly becoming mainstream and promise large potential, but they will only take off universally when UX designers start considering how to incorporate AR into people's everyday life with the purpose of improving productivity or quality of experience (Interaction Design Foundation, 2017).

Similarly, VR technologies have quickly proliferated and are now available to the mainstream consumer (Jerald, 2016). Both AR and VR enable new forms of user interfaces: 3D UIs. Such UIs have impact in many different application domains ranging from art and design to education or entertainment (Interaction Design Foundation, 2017) (Joseph J., Kruijff, McMahan, Bowman, & Poupyrev, 2017).

Current AR and VR input devices such as handheld controllers are limited in terms of rich interaction possibilities. Problems include reliable and precise tracking, user fatigue (the gorilla-arm effect), and practical setup. Simultaneously, state-of-the-art computer vision is too limited, computationally expensive, or requires an elaborate setup to support physical, real-time and real-world interaction. As such, we identify a need for rich interaction input techniques for AR and VR. We propose the touchscreen as an input device for AR and VR: it reliably and accurately captures input, commercially available technologies allow for rich interaction, and it does not require an elaborate setup.

We have developed a functional prototype in a synchronized reality setup, which sidesteps common MR challenges such as advanced tracking (Sjölie, 2015). The prototype is a 3D sketching tool, allowing users to quickly model 3D artefacts in a sketchy and exploratory fashion. Envisioned users are 3D designers, for example architects, industrial designers, and CG modellers.

2 Related Work

We discuss three avenues of research and related work. First, we examine the designerly sketching process itself, and from this distill requirements and suggestions to guide the design of our product. Then, we examine existing 3D sketching tools, which provides some open problems and design opportunities. We also examine current VR input techniques in general. Based on the open problems and opportunities, we explore a variety of interaction design styles, which then serves as inspiration for our design rationale.

2.1 The Sketching Process

Sketching is a fundamental part of the design process (Eissen & Steur, 2011), (Buxton, 2010). Sketches are distinct from finished drawings or rendering, in that sketches are rapidly created "suggestions of meaning" and "externalizations of mental imagery", rather than a "polished recording of an idea" (Ware, 2012). The sketching process is a loop of externalizing concepts and analyzing the result (Ware, 2012). Based on that analysis, the designer then imagines additions, modifications, or subtractions to the sketch, and externalizes those, thereby creating a new loop. Sketching is advantageous because of its speed, interactive imagery and expressive wealth (Hummels, 2000). A sketching tool should thus allow for a both quick and expressive interactions. Finally, the design sketching tool should offer tentative interactions to support creativity (Ware, 2012). In summary, the tool should support:

- Recreation of mental imagery (*sketching an envisioned car*).
- Analysis of the sketch (*a wheel is misplaced*).
- Modification the sketch (*correctly placing the misplaced wheel*).
- Rapidly creating suggestive meaning (*sketching a car is quicker than describing or building it*).
- Expressiveness (*not all cars look and feel the same*).
- Exploration (*what would a three-wheeled car look like?*).

2.2 3D Sketching Tools

An ongoing avenue of research is investigating how to assist the user in the sketch-to-3D translation process (Olsen, Samavati, Sousa, & Jorge, 2009), with the goal of getting computers to understand and interpret sketches in three dimensions. A survey of sketch-based modelling tools poses the following open problems (Olsen et al., 2009):

- Providing a natural interaction method.

- Designing 3D sketching interfaces with a worthwhile increase in utility over traditional 3D modelling interfaces.
- Ensuring discoverability of functionality.
- Precision of manipulation, as compared to traditional control-point manipulation and finding "a better paradigm that fits with human experiences and conventions".

A brief survey of some existing academical and industrial examples of sketch-based modelling tools by the authors confirms some of these problems, and we identify additional areas for improvement.

ErgoDesk is an early example of a sketch-based modelling tool (Forsberg, LaViola, & Zeleznik, 1998). Users draw gestural lines on a 2D surface with a stylus, and these lines are interpreted to create three-dimensional objects. It is an advanced tool, especially considering its time. However, it requires elaborate setup and very specialized equipment.

EverybodyLovesSketch is a similar, sketch-based modelling tool based on stylus input on a 2D surface (Bae, Balakrishnan, & Singh, 2009). It is a powerful tool that allows for sophisticated 3D curve sketching, but its interaction model appears rather complex and hard to discover[1]. Additionally, since sketch construction is highly dependent on input mode and sketch state, it requires a specific order of steps.

Tilt Brush[2] is a commercially available 3D sketching tool for VR. The two handheld VR controllers represent a paintbrush and tool palette. The paintbrush allows the user to paint in 3D space, and the tool palette allows the user to change colors, paint types, and modify the sketch. Tilt Brush offers great expressive wealth, as evidenced by spectacular creations made with it[3].

We note a few opportunities for improvement with regards to its interaction model: it is heavily reliant on modes and mode switching to unlock functionality, it is prone to user fatigue known as the "gorilla arm effect" (Hincapié-Ramos, Guo, Moghadasian, & Irani, 2014), and finally, there exists an inherent conflict in creating 3D shapes with 2D lines in 3D space. In a 2D medium, it is possible to create the illusion of depth and volume by indicating contour lines. In a 3D medium, however, the contour lines are dependant on angle of view. Consider a sphere: it can be represented in a 2D medium by a circle, but no singular contour line exists in a 3D medium. Thus, to create convincing 3D shapes in Tilt Brush, the user needs to fill them in entirely. This induces a tension between the amount of effort required and the desired suggestive qualities of a sketch.

Blocks[4] is another tool to rapidly create 3D models in VR. Unlike Tilt Brush, it is based on creating and modifying geometrical primitives. This resolves the tension between effort and suggestive quality as described above. Fur-

[1] `https://www.youtube.com/watch?v=hGmE362ldZY`
[2] `https://www.tiltbrush.com`
[3] `https://www.youtube.com/watch?v=P_Azf1sSnRg`
[4] `https://vr.google.com/blocks/`

thermore, taking primitives as starting point aligns with product design sketching methods, which start with basic shapes such as cubes, cylinders, cones, spheres, and planes (Eissen & Steur, 2011). Nevertheless, like Tilt Brush, it is reliant on modality and faces similar issues with arm fatigue. Additionally, it makes heavy use of the control-point paradigm as described by Olsen et al. (2009).

In summary, we identify the following design opportunities for sketch-based modelling systems:

- Creating an interaction model that:
  - uses natural interaction techniques,
  - is less dependent on modes and order of action,
  - is suited for prolonged use (i.e. avoids "gorilla arms"),
  - is rapid and suggestive, and
  - captures expressiveness.
- Differentiating 3D sketching from 3D modelling by:
  - providing a worthwhile increase of utility over traditional 3D modelling, and
  - finding a "better paradigm that fits with human experiences and conventions".

2.3 Design Rationale & Interaction Style

To satisfy the requirements and address the open problems described above, we employ design strategies from a variety of interaction styles. We then discuss current VR input techniques and multitouch interaction with regards to these styles.

2.3.1 Interaction Styles

Embodied interaction is "the creation, manipulation, and sharing of meaning through engaged interaction with artifacts" (Dourish, 2004). This definition captures the designerly sketching process with clarity and precision. Pen and paper are used to create, manipulate, and share meaningful representations of ideas and mental imagery. This interaction style emphasizes the importance of the aforementioned suggestive qualities in sketching.

Freedom of interaction expands on the notion of embodied interaction by stating that interaction should: "take advantage of a person's perceptual motor skills, offer myriad ways to achieve a products functionality, allow the person to act at multiple points at once, and allow for easily reversible actions" (Wensveen, Djajadiningrat, & Overbeeke, 2004). This interaction style emphasizes the importance of modeless interaction in our product.

Rich interaction is a related approach to product design that takes advantage of a person's perceptual-motor skills, cognitive skills, and emotional skills (Frens, 2006). Perceptual-motor skills are emphasized because humans *do*, in addition to *think*. This is in accordance with the act of sketching, through which designers are able to create something "greater" than they can mentally imagine, by offloading mental images to marks on paper (Ware, 2012). Furthermore, emotional skills are emphasized because sketches embody an expressive quality: they can be quick and dirty, unsure and ambiguous, rough and suggestive, bold and emphasised, or beautiful and polished.

Natural interaction is an interaction style that, similarly to rich interaction, takes full advantage of it's user's capabilities, such that their behaviour and experience feels natural (Wigdor & Wixon, 2011). An approach to achieve this is to make the coupling between action and function direct or to employ augmented feedforward to indicate the coupling (Wensveen et al., 2004). A direct coupling means that when a user acts on a product, the action is met with the products reaction that is equal to the input action's time, location, direction, modality, dynamic, and expression. A coupling through augmented feedforward is employed when a direct coupling cannot be established (as an example, consider a gradient indicating a GUI button's unpressed state despite a mouse or finger not actually being able to depress the button).

Gestural interaction promises opportunities for capturing strongly expressive qualities as input. This has been explored specifically with regards to design tools in Hummels's (2000) PhD thesis. In it, she points out that most gestural interfaces are based on "pre-defined symbolic commands" due to technological constraints. However, determining the meaning of a gesture that is not pre-defined within a logical ruleset is proves a formidable challenge. The resolution of this conflict, Hummels proposes, lies in using gestural interaction to "invite and seduce the user to explore".

We see this proposal as an extension of the aesthetics of interaction that aim for a **enjoyable, delightful, and magic experience**, through use of the product. This design approach has been employed in VR and 3D UIs that use magical interaction to simultaneously create an enjoyable experience and overcome human and technological limitations (Jerald, 2016)(Joseph J. et al., 2017). We can use a similar approach as described above when designing for natural interaction. Instead of creating a direct coupling between action and function, however, we deliberately reverse the coupling along some aspects. As (Wensveen et al., 2004) notes, this leads to:

> *Electronic products instill moments of magic and surprise that seem to surpass the laws of nature and physical causation.*

2.3.2 Current VR Input Techniques

Current handheld VR controllers offer interaction through buttons and positional tracking. Buttons offer few possibilities for rich interaction: different functions are unlocked through the same action. There is thus no room for perceptual or motorical differentiation (Djajadiningrat, Overbeeke, & Wensveen, 2000). Furthermore, buttons offer few possibilities for expressiveness: a discrete on-off state captures much less information about its use than continuous input controls.

The positional tracking system captures much more richness and expressiveness: three degrees of freedom (3-DOF) for the controller's position, another 3-DOF for its rotation, its movement speed and acceleration, and the interplay between two different controllers in different hands. A notable drawback to handheld controllers, however, is that they require the user to hold the controllers, thereby preventing the user from interacting with other devices and the world around them. In effect, the controllers become the users hand, at the cost of replacing their hands' capabilities, rather than extending them.

There are alternatives addressing this issue: instead of using controllers, employing computer vision to detect and model hands. The Leap Motion[5] is a commercially available example. The issue with this strategy, however, is one of precision and stability. Satisfactory results are attainable, but they require precise calibration and lighting conditions. Another issue with in-air gestural interaction is dubbed the "live mic" problem—the system must differentiate between actions that are intended as input and those that are not (Wigdor & Wixon, 2011). Handheld controllers often use button to "record" or confirm an action, but bare hands have no clear "on" switch.

Tracked controllers or hands can thus provide 6-DOF controls (3-DOF position and 3-DOF rotation) controls, but an issue persistent across 6-DOF interactions in general, is the "gorilla-arm effect": fatigued arms resulting from raised or outstretched arms for mid-air interactions (Hincapié-Ramos et al., 2014). This makes 6-DOF control unsuited for prolonged use.

Finally, we note the drawback of having many degrees of freedom: coordinated control of 6-DOF demands much more effort from the user than 2-DOF manipulations. As such, reducing degrees of control is crucial when high-precision interaction is needed, such as when creating 3D models (Joseph J. et al., 2017). In addition to being hard to perform, they are harder to design correctly—as unbounded action possibilities allow for expressiveness, but require the correct reaction from the product (Wensveen et al., 2004). This poses a significant challenge to the interaction model, as it would allow a wealth of action possibilities, but simultaneously needs to communicate all possibilities and state of the system clearly, if the user is not to feel overwhelmed.

2.3.3 Multitouch Input

We address the issues described above by looking to the multitouch tablet as input device. As Buxton (2016) notes, multitouch expands our gestural vocabulary, can capture more of the richness of the real world, and exploit the skills that we have acquired of living in it.

MRTouch (Xiao, Schwarz, Throm, Wilson, & Benko, 2018) is a recent is

[5] www.leapmotion.com

an academical example showing the viability of touch input for head-mounted MR. It's input tracking is based on computer vision however, and precision is therefore limited. Our approach is based on the multitouch tablet as input device, thereby creating multitouch input with reliable, high precision.

In summary, we identify the multitouch tablet as a valuable input device for MR, due to its high accuracy, the possibility for rich, gestural, and modeless interaction, while simultaneously being suited for long-term use, decreasing required effort due to the 2-DOF interaction surface, and solving the "live-mic" problem by providing tangible feedback while the system is sensing input: touching the screen.

3 Concept

Using a multitouch tablet as input device for MR, we create a 3D sketching tool. We use the multitouch tablet as a reliable and high precision input device. It is capable of offering interaction styles which match the sketching process naturally.

The MR setup consists of a VR headset connected to a computer running a VR application. Also connected to the computer is the multitouch tablet, through which the primary interaction will take place. The position of the tablet in the real world is synchronized with a virtual representation of the tablet in the virtual world. The setup resembles a synchronized reality setup (Sjölie, 2015), by selectively bringing physical artifacts into the virtual world (in this case, the tablet), which then serve as input to the virtual world but also as tangible grounding.

This project is thus simultaneously an exploration of using a multitouch tablet as input for MR, and creating a sketch-based modelling tool in MR with multitouch input.

4 Process

This project is framed as a *research through design* project (Stappers & Giaccardi, 2013), where we gain general knowledge about the possibilities for rich input for MR through a tablet, by designing a specific tool as demonstrator of this newly gained knowledge. As is common in research through design projects, design direction was determined both by literature studies and by prototyping, where literature and prototypes heavily influenced each other. In section 2 and section 8 you can read about the different insights we gained. In this chapter we explain how we carried out the process when designing our prototype as one possible solution.

As the community is still establishing a formal process for research through design, there are no universally agreed upon guidelines of how to delineate the design process into concrete phases. This project started the process with an ideation phase, continued by creating and evaluating a high-fidelity prototype, and ended the process with making the video prototype. We consider the the literature studies and this report, the high-fidelity prototype, and the video demonstrator all to be carriers of the newly generated knowledge.

4.1 Design Log

Here we chronologically detail the design process and methods used.

4.1.1 Ideation Phase

Brainstorming sessions were performed in order to explore different kind of use cases both in terms of scenarios but also in terms of tools. An early idea was to sketch with a stylus in 2D on the tablet and then translate that into 3D through an AR interface, as seen in figure 1. Another idea was to use a tablet and stylus to sketch directly in VR on top of the tablet, as seen in figure 2.

Figure 1: Sketching in 3D with a stylus pen on a tablet and viewing this through an AR interface.

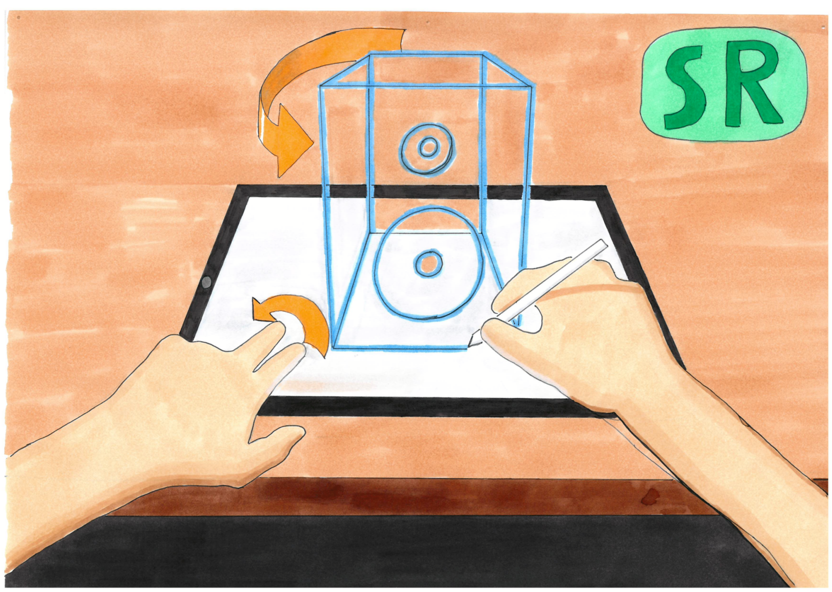

Figure 2: Sketching in 3D in a synchronized reality setup with a tablet and stylus.

When we found that we did not have access to a stylus due to limited resources, we explored finger-based multitouch interactions instead. The ideation phase continued with a day long *bodystorming* session. Bodystorming is an ideation method well suited for MR design processes (Vitazko & Zeller, 2018). This session gave new insights into "sketchily" working in 3D. With modelling clay that represented primitives and a piece of paper that represented the tablet, we could explore this type of sketching and gained insights into the different steps and features that are needed in order to support the intended goal of 3D sketching. A car was the typical model we tried to re-create with different primitives, seen in figure 3. After this bodystorming session we could decide upon potential features and gestures. This concept was again captured in a sketch, see figure 4.

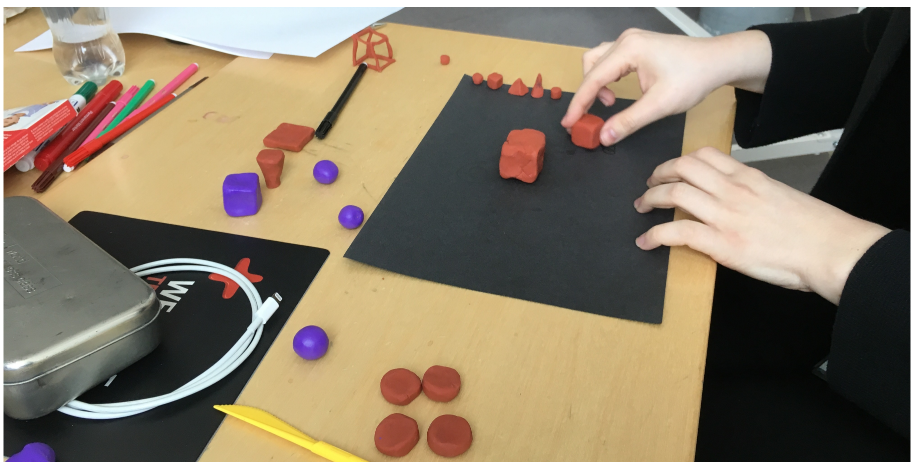

Figure 3: Bodystorming session with modelling clay and piece of paper.

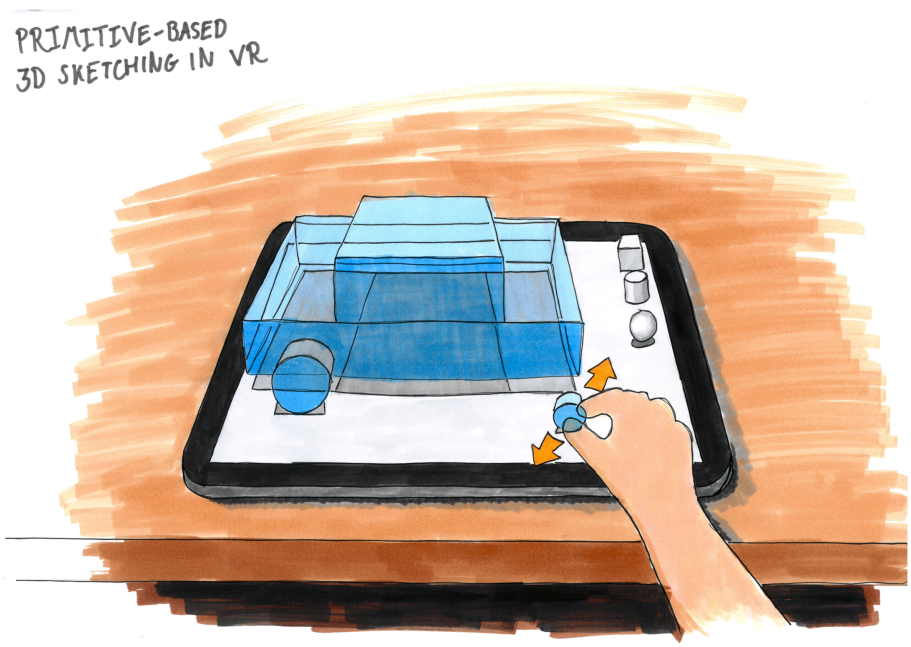

Figure 4: Sketching in 3D based on primitives.

4.1.2 High Fidelity Prototype

From the list of possible features (full list in subsection 4.2) we decided upon a few features that we wanted to include in our minimum viable product. Those features were; creating and deleting primitives, moving/panning in the horizontal plane, rotate around the vertical axis, and scale the primitives through pinch. Roughly half of the design process was spent on hi-fi prototype implementation.

The setup for the high fidelity prototype consists of two main parts: the touchscreen tablet input and the VR head mounted display output. This division is reflected in the implementation: we implemented the gesture recognition by building an iOS application in Xcode 9, which communicates with an Unreal Engine 4 application running the VR simulation and powering the HMD. In Unreal we translated those positions into the VR environment and applied the transforms as calculated by the iOS app.

In a later iteration we managed to include the roll feature and the vertical translation. We wrapped up the hi-fi prototype by including instructions to our five gestures into the VR-environment so the user have something to fall back on in case they forget.

Finally, we conducted a UX evaluation of this prototype, which you can read more about in section 6.

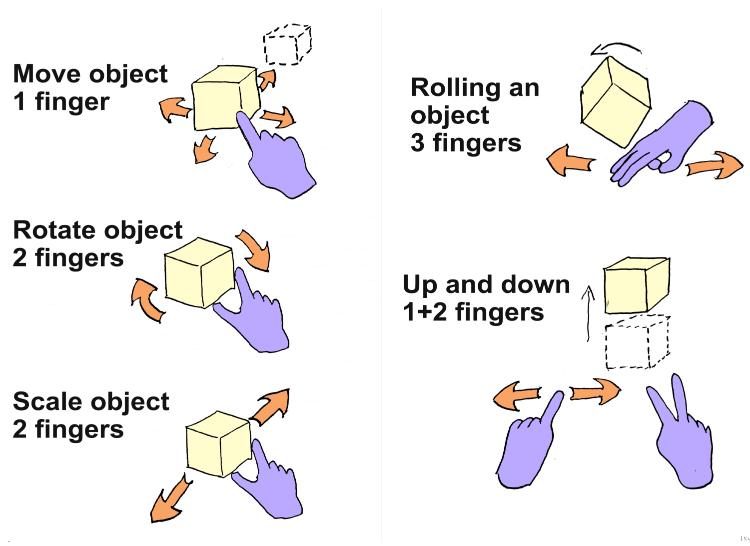

Figure 5: The five possible gestures in the Hi-Fi prototype of CyberSketch, together with instructions.

4.1.3 Video Prototype

Because of limitations in the hardware, we weren't able to implement all desired features. Also, through implementing and prototyping, we gained new insights. Finally, as we consider the video to be a carrier of knowledge as well as this report, we present the envisioned concept and reasoning behind it through a video prototype. The points that we wanted to include were: why use a tablet and not hand-controllers, what 3D sketching really means, how our prototype works right now, what we envision in the *near* future (improvements on our prototype), and what we envision in the *distant* future (the whole concept of 3D sketching in the future). A one day video shoot session was performed where we included shoots of both the real life interaction as well as the 3D translation in the VR-environment. For the future concepts we used Adobe After Effects and Cinema 4D to animate what we are envisioning.

Figure 6: A view of the video prototype showing interaction with the tablet.

4.2 Feature-level breakdown

This is a list of features which together describes the full envisioned 3D sketching tool CyberSketch. This list thus covers features we implemented, features we failed to implement due to scope of project, and features which we envision in future versions.

Implementation software and hardware Unreal Engine 4.19 was used to create a VR application. An iPad Air 2 was used to capture input. An Oculus Rift headset was used as VR output. An app for iOS was developed in Xcode, making use of both standard and custom-built gesture recognizers. It interfaces with Unreal Engine by sending messages over WiFi using the OSC protocol. The iPad recognizes the gestures and calculates transform information. It then sends the position of action (XY screen coordinates) and transform information to the Unreal Engine application. The Unreal Engine application then finds a primitive with corresponding primitives if it exists, and applies the transform to it.

Gestural interaction model We created a gestural interaction model, based on natural interaction where possible, and "magic" interaction where we needed to break the physical constraints. Specifically, we created the following gestural interactions:

- **Translate horizontally**: one or two finger drag.

  Primitives are moved across the interaction surface, or "canvas", by dragging them with one or two fingers. We support one-finger drag as that is the most common approach (Hinrichs & Carpendale, 2011). We also support two-finger drag, such that it can be performed simultaneously (modelessly) with scaling or rotating yaw, thereby increasing the freedom of interaction. The action and function are coupled in time, location, direction, and dynamics, which means that this interaction approaches natural interaction (Wensveen et al., 2004).

- **Scale**: two finger pinch.

  Primitives are scaled uniformly by a two finger pinch. This is another common surface-based interaction technique (Wobbrock, Morris, & Wilson, 2009). The main difference with the common multitouch approach is that the pinch here scales the object in three dimensions, rather than two. Action and function are coupled in time, location, direction, and dynamics.

- **Rotate (yaw axis)**: two finger rotate.

  Primitives are rotated along the vertical (yaw) axis by touching the screen with two fingers and rotating them with respect to the center between the two contact points. As opposed to a single touch on a corner to rotate, this approach allows for simultaneous rotation and translation (Hancock, Vernier, Wigdor, Carpendale, & Shen, 2006). The action and function are coupled in time, location, direction, and dynamics.

- **Rotate (roll axis)**: three finger drag.

  Primitives are rotated along their roll axis by a three finger drag. By grabbing the primitive with three fingers, the primitive tips over in the direction which it is being pushed or pulled. This gesture falls under magic rather than natural interaction, as it is only coupled in time and dynamics. It is not coupled in location (moving fingers does not move primitive), or direction (horizontal action to achieve rotational function).

- **Translate vertically**: two handed gesture.

  Primitives are moved along their vertical axis by touching it with two fingers, and moving a third finger (from the other hand) towards or away from the object. This interaction is inspired by the "balloon selection" technique (Benko & Feiner, 2007), which is a metaphor of moving a helium balloon up or down by anchoring its string in a pulley, and pulling on the string away from the pulley. As such, a horizontal action effects a vertical displacement. This interaction is also considered magical rather than natural, since it is not coupled in location or direction.

Creating and deleting items On the right side of the canvas, we permanently display the primitives that are possible to create. Primitives are created by dragging them onto the canvas. This creates an editable duplicate of the primitive, while the original stays in place.

Items are deleted by dragging them off the edge of the canvas. Primitives will turn red when they are near the edge, indicating that they will be deleted on release. If deletion is not desired, the user can drag them back onto the canvas.

Modeless Interaction As outlined in section 2.3, modeless interaction is an important feature of this tool. We support modeless interaction in several ways: firstly, moving, scaling, and rotating can be done simultaneously. Secondly, the interactions that cannot be performed simultaneously require only a very subtle mode switch: touching with an additional finger. Because this is embedded in the gesture itself, we consider it almost modeless. Thirdly, selecting a primitive to act upon is done simply by acting upon the primitive. There is no specific selection mechanism or "active mode". Finally, changing the angle of view is done by moving the head. Since this doesn't interfere with the actions the hands are doing, view changes are also modeless.

In-game help: gesture lookup Because this sketching tool introduces some new gestures, we included an explanatory graphic on the virtual table. We placed it such that if upon looking towards the tablet it is in peripheral vision. That way it is doesn't distract or interfere, but can easily be used to look up gestures.

Hands in VR A planned feature that we were not able to implement is bringing hands into the virtual world. We looked into employing computer vision to do this (precise tracking would still be left to the tablet). Experiments with a head-mounter Leap Motion failed due to reliability. We were able to bring hands into VR, but issues with positional reliability caused it to be unusable. We synchronized the virtual tablet's position based on detected real world hand position, but if a hand left the Leap Motion's view was and was rediscovered, its position would be slightly offset compared to the previous position, and thus the virtual tablet. This rendered the virtual hands unusable very quickly. This problem was compounded by the fact that the glare and reflection from the tablet's screen interfered with the Leap Motion's hand detection.

Nevertheless, we consider it vital to be able to see hands in a MR 3D sketching tool. It not only helps increase presence, but is also essential in guiding the users hands and fingers, and by extension the action, to the right primitive. As a workaround, we show a circular orange highlight on the virtual tablet at the location where the user touches the real tablet. An issue with this approach is that the highlight is obscured by primitives on top of it.

Two-handed interaction Bringing hands into VR through computer vision would greatly help with selection, as described above. However, we also identify the opportunity to augment the multitouch input data with hand recognition data. Even relatively imprecise hand position data would be enough to determine which touch on the screen corresponds to which finger.

This opens up the possibility to create true two-handed interaction, where different hands are able to perform unique actions. Envisioned interactions are copying, stretching, bending, and moving the canvas. Copying would work by holding a primitive with one finger on one hand, while dragging away to "pop off" a copy with another finger the other hand. To a hand-blind multi-touch system, this is the same gesture as scaling, but augmenting touch points with corresponding hands allows this gesture to work. Similarly, stretching and bending are interactions with two finger on each hand (four total). Finally, we envision manipulating the sketch with the dominant hand, while moving and rotating to orient the canvas with the non-dominant hand, in line with Guiard's (1987) two-handed interaction theories.

5 Final Version of CyberSketch

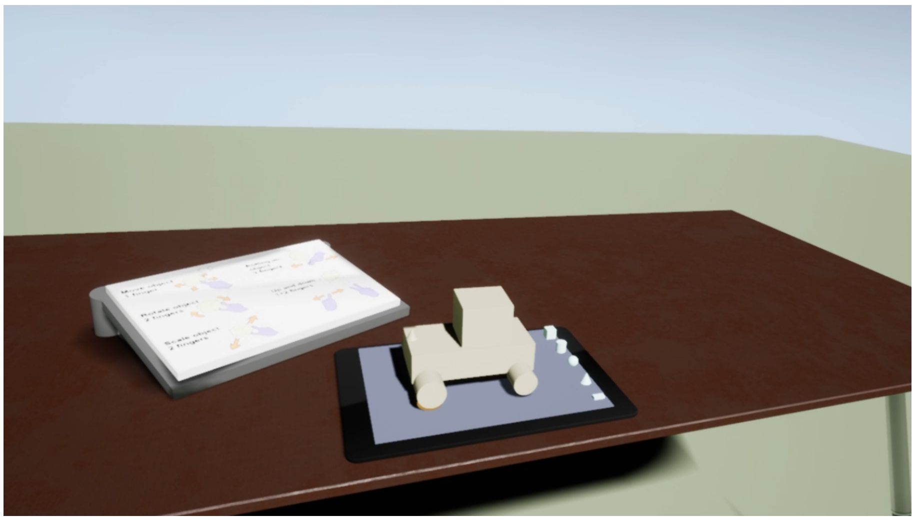

Figure 7: A screenshot of a car being built in the high fidelity prototype.

We consider the “final version” of CyberSketch to be both the high-fidelity prototype and the video prototype. Together they sum up the designed artefact in terms of implementation and vision.

To summarize, the implemented prototype as of today consists of a VR application and an iOS application[6]. It contains features such as:

- **Creating primitives** by dragging them from the side onto the canvas
- **Moving on the horizontal plane** by dragging.
- **Scaling uniformly** by pinching with two fingers.
- **Rotating along the yaw (vertical) axis** with two fingers.
- **Rotating along the roll axis** by dragging with three fingers.
- **Moving vertically** with a two handed, three finger gesture.
- **Deleting a primitive** by dragging the primitive off the edge of the tablet.
- **In-game instructions** to look up gestures.
- **View manipulation** by moving and rotating the head freely.

[6]https://github.com/DanielRoeven/vr-3d-sketching

The envisioned future version is more elaborate, and can be seen in our video prototype[7]. In addition to the mentioned features above, a future version of CyberSketch would also contain features such as:

- **Showing user's hands**, which will increase presence and help with targeting shapes.
- **Copying modified primitives**, thanks to the differentiation of hands, this interaction speeds up the workflow by not needing to recreate shapes manually.
- **Stretching and bending primitives** which allows for more expressive shapes.
- **Moving the canvas** with the non-dominant hand, which together with moving the head increases freedom of view manipulation.

The current prototype is based in VR, but we see much value to be gained by moving more towards MR and AR. Such a version would require less context switching, make it more mobile and enable collaboration.

[7] `https://vimeo.com/272779701`

6 Evaluation

Since this project is an exploratory research of using the tablet as rich input device for 3D sketching, there is a wide variety of elements that we want to evaluate. The *user experience in general* is the most important aspect that we evaluated, followed by *requirements* and *usability*. We are not specifically evaluating the learnability of the tool, but instead focus on evaluating the possibility of creating rich and natural interaction with tablet-based MR input for sketching. We are also not evaluating the stability of the application since it acts as a prototype for the concept. However, we will note when it interferes with a good user experience.

6.1 Methods

Various methods for evaluation were performed in order to test the main elements (user experience, requirements, usability) in detail. The sample size of this evaluation is seven (7), and each evaluation session had a duration of 15 minutes.

- **Moderated Task Test**
  Participants were introduced to the concept of 3D sketching and the setup of the tools, as well as an introduction to the different components and gestures the tool provides. After signing a consent form for permission to record the whole study, the participants could start interacting with the tool. The interaction was divided into three parts; getting familiar with the tool for one minute, one moderated task (sketching a car) for four minutes, and one free form task (sketching something of their choice) for four minutes.

- **Questionnaire - Microsoft Product Reaction Cards**
  In order to capture the user feedback in concrete words, a suitable way for doing that is using the Microsoft Product Reaction Cards (Benedek & Miner, 2002), which is a list containing 118 different and contradictory words. Turner (2017) explains that this method is good for capturing experiences and feelings towards a product while at the same time identifying areas for improvement. The participant were provided with a questionnaire containing the shorter version of the Reaction Cards, meaning 64 words instead of 118, which still captures a wide variety of responses while shortening the needed time (Turner, 2017). They were then asked to pick seven words that describe the experience the best and rank them.

- **Follow-up Interview**
  The user study was wrapped up with a semi-structured interview in order to collect first hand accounts of opinions and perceptions of the experience (Hanington & Martin, 2012). The starting point of the interview was questions regarding the participants reaction based on the reactions cards. Depending on the time left, the participants were also asked about their

opinions of 3D sketching as a addition to a designers work flow, as well as the whole setup of tablet input for VR.

The following table explains the detailed aspects we are testing and what methods cover them:

Aspect of study	**What are we evaluating?**	**How are we evaluating?**
User Experience		
The emergent UX as a whole	Every product always evokes a UX through its use: what is the emerging UX of this prototype?	Questionnaire, Interview
Enjoyment of the experience	As covered in 2.3, we focus on enjoyment of the experience over ease of use: does the unlocking of the functionality contribute to the overall experience?	Questionnaire, Interview
Freedom of interaction	Do participants make use of modeless interaction, and perform steps simultaneously? Does the prototype allow users to accomplish the same goal in different manners or order of steps?	Task Test
Exploration	Does the prototype support exploratory behaviour during sketching? Does the prototype stimulate exploration of interaction possibilities?	Task Test
Expressiveness	Do all cars look the same? Does the prototype allow the users to, through different ways of interaction, to express themselves?	Task Test
Requirements		
Recreation of mental imagery	Is the end product recognizable? Does the prototype allow the user recreate an imagined object?	Task Test
Analysis of the sketch	Do users analyze the state of their sketch?	Task Test
Modification of the sketch	Does the prototype allow users to modify the sketch based on the analysis step?	Task Test
Usability		
Gestural interaction	Are the action-function coupling strong enough for the user to understand the interaction possibilities? Are the gestures themselves usable?	Task Test; Questionnaire
2D input for 3D manipulation	Does the 2D interaction surface allow for 3D interaction, specifically on the Y axis? Can users achieve all 3D manipulations they want?	Task Test

6.2 Results

The sample size of this evaluation was seven participants, and each participant being asked to choose seven reaction cards. As such, the total amount of gathered reactions cards was 49. The division between the total number of positive and total number of negative emotions was 59.2% and 40.8% respectively. Below is a graph visualizing the top 11 experiences that occurred more than once:

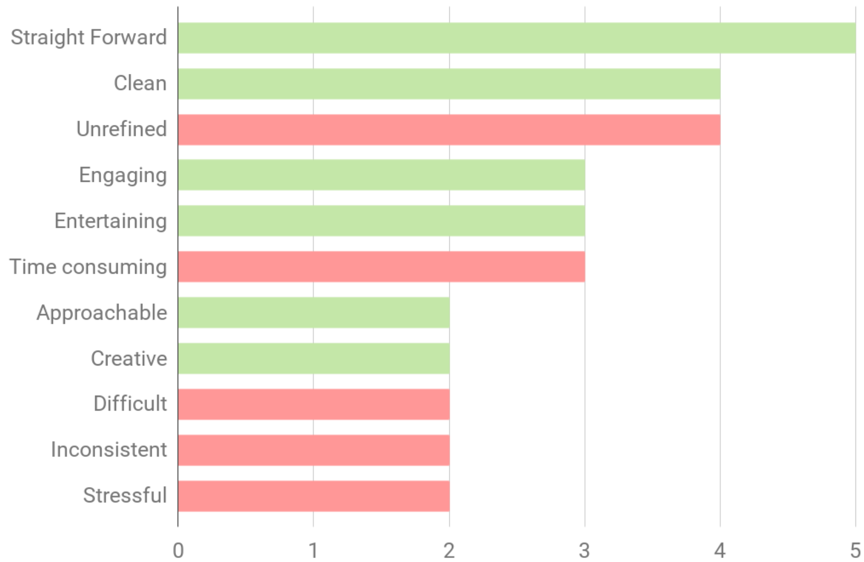

Figure 8: Top 11 experiences (describing 65.3% of the whole experience) and the number of times they each appeared. Other positive experiences that occurred once were: *collaborative, consistent, desirable, easy to use, familiar, flexible, friendly, innovative, inviting, simplistic*. Other negative experiences that occurred once were: *annoying, confusing, frustrating, impersonal, ineffective, overwhelming, rigid*.

6.2.1 User Experience In General and Usability

Five out of seven participants felt that the tool was very *straightforward* and four out of seven felt that it was *clean*. From the interview we found out that the reason to that was that they got a feeling of “you get what you see” and you have limited amount of actions you can perform making it easy to understand. They also mentioned that the primitives felt accessible and thanks to the straightforward instructions it was easy to learn. Having primitives for the 3D sketching made some of the participants feel *engaged* and since it was a new way of sketching they thought it was fun and *entertaining*.

However, since the prototype isn’t stable enough to function perfectly, a lot of the negative experiences stemmed from the instable prototype itself. The prototype can be perceived as jumpy and *unrefined* making it harder to perform the actions the participants desire. All of this makes it *time consuming* and *difficult* to handle.

Contradictory Perceptions Most of the participants shared the overall experience of the tool. Nonetheless, it happened that they sometimes perceived opposite experiences around the same feature. One of the features that they had split opinions about was sketching with pre-defined primitives. Some participants claimed that this limitation stimulated their creativity positively in a way that they had to create their own meaning out of the primitives. Others argued that this limitation rather challenged their creativity in a way that they couldn't fully express themselves which gave them a feeling that they had to do it the way the tool wanted it.

Another feature where the participants had opposite experiences was the different gestures. The participants were familiar with a majority of the gestures (pan, scale, rotate), however the *new* gestures (roll and raise/lower) received mixed reactions. Some participants thought that even though the gestures are new to them, once they got it explained it was easy to learn. Other claimed that the new gestures were not intuitive enough for them to remember.

Concept When it comes to the opinions of the general concept of 3D sketching on a tablet being a part of a design process in the real world, the participants had a range of different thoughts. A majority of the participants had difficulties understanding what the purpose of the tablet is since they see it as a physical constraint. They claimed that being in a VR environment made them want to interact with primitives in and around the environment as well, rather than have them on a tablet. Some of them also claimed that the hand-controllers that comes in the VR-set felt more intuitive to them. Most of them also argued that they would have preferred to see their hands in the environment together with the tablet. Although, a few could see different scenarios where the limitation of a tablet would be desirable, such as small meeting rooms where there is no way for bigger movements.

The general concept of quick 3D sketching, especially when prototyping models such as cars and buildings which is a challenge to sketch on a paper, is welcomed by most of the participants.

6.2.2 Requirements, freedom, exploration and expressiveness

Requirements In six out of seven cases in which the participants were asked to create cars, the end result was clearly recognizable as a car. Users often analyzed their sketches and made adjustments based on this analysis. For example, users put effort into ensuring that the wheels of the car were the same size, by adjusting them while looking at other wheels for reference. We can thus say that CyberSketch meets the requirements of the sketching process: recreation of mental imagery and analysis and modification of the sketch.

Freedom of interaction The task tests showed that users build their cars in different ways: some start by placing wheels while others start by placing the body. Additionally, we noted that users make use of modelessness by simultaneously moving and scaling or rotating.

Exploration In the free-form task test, users created varying "sketchy" structures, like football fields or castles. Some did this by immediately building what they had in mind, while others played with the available shapes for a while to decide how and what to build. In the task of creating a car, one user started building one style of car, but upon finding another available shape, changed course and built a different style of car. One user even created a three-wheeled car. We thus find that the prototype allows for exploration.

Expressiveness Related to the previous point, the cars that users created do not look the same at all. Even with the small variety of available primitives, cars with very different "feelings" were created. See figure 9. Thus, the prototype allows for expressiveness.

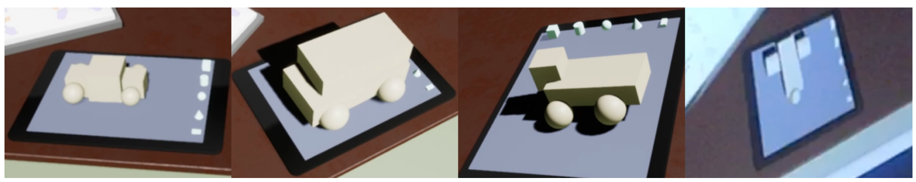

Figure 9: Some of the cars that were created.

7 Discussion

7.1 Learning outcomes

Since this project was carried out for the course CIU235 Individual Project in Interaction Design, we list some educational insights and learning outcomes here.

Performing research through design Literature studies and prototype design influenced each other heavily while providing guidance for the design direction. Since this was our first (large) research through design project, problems we encountered were striking the correct balance between literature and design, scoping the literature studies correctly, and spending a lot of time on implementation. Through trial and error (spending too much time on implementation, then needing to catch up in other areas), and reflecting on our the process, we learned to carry out a research through design project. We also learned how the contribution of this project is not the designed prototype itself, but rather generalized insights that came through creating this prototype. Finally, we have been thinking about how to communicate the generated knowledge, and have adopted a three-pronged approach: publicly available prototype, report, and video. We consider each of these to carry the generated knowledge, and use them to convey the insights in different manners.

Designing and developing for MR Neither of the authors had extensive experience developing for AR or VR. This was a learning experience in terms of VR-related design concepts (presence, degrees of freedom, different types of tracking), in terms of hardware involved, and in terms of implementation and development. Furthermore, MR offers a new design space as compared to physical products or GUIs, which we learned to navigate through literature and design methods.

Disciplines and tools Because MR has such a different design space than what the authors had worked with previously, we developed an appreciation of, and affinity for, new tools, such as Unreal Engine, 3D modelling and texturing tools, Xcode, WiFi networking with the OSC protocol, and After Effects and Cinema4D to create a video prototype.

Evaluation In the evaluation we made use of new methods, such as the Microsoft Product Reaction cards. This UX evaluation method was considered useful and pleasant to use by the authors, and will likely be used again going forward.

Communication Any successful project requires good communication. We gained experience with regards to internal communication (between the authors to set up a design process, communicate design vision, and giving feedback to each other), with the supervisor (communicating design decisions, describing development problems), and with the world (through report and video). These insights will be carried forward, and be of use in further design projects, such as the Master's thesis.

7.2 Limitations

The main limitations of the prototype are not showing hands due to hardware issues with the Leap Motion, imperfect multitouch gesture recognition leading to frustration and functions that are hard to use, a small canvas, since primitives can only be placed on the tablet, and finally, not supporting copying, stretching, or bending primitives.

The main limitations of the evaluation are sample size (7) and educational background of users (all users study interaction design). Furthermore, all users are acquaintances of the authors, introducing significant bias to the evaluation. However, this bias should be limited to the questionnaire and interview, as the other aspects were evaluated through the task test, specifically in aspects that the users were unaware of. For example, users were not asked to to manipulate position and scale simultaneously (modeless interaction), or create a car with a "unique feeling" (expressiveness), but did so anyway.

7.3 Future work

Another development iteration addressing bugs in the current prototype and another evaluation with more participants from diverse educational backgrounds would be valuable, specifically to ensure that any experienced frustration was indeed due to bugs in the prototype.

Apart from that, we envision additional features for CyberSketch, as covered in section4.2, section 5, and the video prototype.

8 Conclusion

We have explored the possibilities of accurate and rich multitouch tablet-based input for MR, corresponding to a synchronized reality setup (Sjölie, 2015). We have done so by designing and developing a 3D sketching tool called CyberSketch, which demonstrates the possibility of modeless, natural, and gestural input.

Specific contributions made by this project are:

Practical setup We have shown the feasibility of using a multitouch tablet as input to VR. We have shown this by creating an app running on an iPad Air 2 with iOS 11. It interfaces over WiFi by sending OSC messages to Unreal Engine 4.19, which runs the VR application, displayed on a Oculus Rift headset. Reliability, accuracy, and latency of interaction in general were all satisfactory.

Rich interaction We have shown the feasibility of modeless, natural, gestural, and expressive interaction with the tablet through creating a prototype making use of these interaction styles. Evaluations showed that users can make use of the gestural, modeless and natural interaction to create expressive cars.

Sketching in 3D We have shown the feasibility of creating a 3D sketching tool based on interaction on a two-dimensional surface, which addresses issues with unconstrained 6DOF interaction such as "gorilla arms" and the "live-mic problem".

Two-handed Interaction We have identified a viable strategy for creating two-handed multitouch interaction, by augmenting the precise multitouch input with rudimentary computer vision. We have also identified valuable applications of this technique: manipulating the canvas with the non-dominant hand, speeding up the workflow by copying with gestures, and increasing expressiveness with gestures to bend and stretch.

Communicate value of constraining DOF This project is an exploration of harnessing the value of the multitouch tablet as input to MR. In the evaluation, however, we found that some participants reacted negatively to the idea of constraining the degrees of freedom. The complaint is that because anything is possible in VR, everything should be possible. The authors hope to convey the value of tablet-bound interaction through this report, but we have found that it is important to also convey this value to the user of the tool, lest it negatively affect the user experience.

References

Bae, S.-H., Balakrishnan, R., & Singh, K. (2009). Everybodylovessketch: 3d sketching for a broader audience. In *Proceedings of the 22nd annual acm symposium on user interface software and technology* (pp. 59–68).

Benedek, J., & Miner, T. (2002). Measuring desirability: New methods for evaluating desirability in a usability lab setting. *Proceedings of Usability Professionals Association*, *2003*(8-12), 57.

Benko, H., & Feiner, S. (2007). Balloon selection: A multi-finger technique for accurate low-fatigue 3d selection. In *3d user interfaces, 2007. 3dui'07. ieee symposium on.*

Buxton, B. (2010). *Sketching user experiences: Getting the design right and the right design.* Morgan Kaufmann.

Buxton, B. (2016). *Multi-touch systems that i have known and loved.* Microsoft Research. Retrieved from `http://www.billbuxton.com/multitouchOverview.html`

Djajadiningrat, J., Overbeeke, C., & Wensveen, S. A. (2000). Augmenting fun and beauty: a pamphlet. In *Proceedings of dare 2000 on designing augmented reality environments* (pp. 131–134).

Dourish, P. (2004). *Where the action is: the foundations of embodied interaction.* MIT press.

Eissen, K., & Steur, R. (2011). *Sketching: The basics.* BIS Publishers.

Forsberg, A. S., LaViola, J., & Zeleznik, R. (1998). Ergodesk: a framework for two-and three-dimensional interaction at the activedesk. In *Proceedings of the second international immersive projection technology workshop* (pp. 11–12).

Frens, J. W. (2006). *Designing for rich interaction: Integrating form, interaction, and function* (Unpublished doctoral dissertation). Eindhoven University of Technology.

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of motor behavior*, *19*(4), 486–517.

Hancock, M. S., Vernier, F. D., Wigdor, D., Carpendale, S., & Shen, C. (2006). Rotation and translation mechanisms for tabletop interaction. In *Horizontal interactive human-computer systems, 2006. tabletop 2006. first ieee international workshop on* (pp. 8–pp).

Hanington, B., & Martin, B. (2012). *Universal methods of design : 100 ways to research complex problems, develop innovative ideas, and design effective solutions.* Rockport Publishers, Osceola. Retrieved from `https://ebookcentral.proquest.com`

Hincapié-Ramos, J. D., Guo, X., Moghadasian, P., & Irani, P. (2014). Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the 32nd annual acm conference on human factors in computing systems* (pp. 1063–1072).

Hinrichs, U., & Carpendale, S. (2011). Gestures in the wild: studying multi-touch gesture sequences on interactive tabletop exhibits. In *Proceedings of* *the sigchi conference on human factors in computing systems* (pp. 3023–3032).

Hummels, C. C. M. (2000). *Gestural design tools: Prototypes, experiments and scenarios* (Unpublished doctoral dissertation). Delft University of Technology.

Interaction Design Foundation. (2017). Augmented reality – the past,the present and the future. *Interaction Design Foundation*.

Jerald, J. (2016). *The vr book: Human-centered design for virtual reality.* New York, NY, USA: Association for Computing Machinery and Morgan & Claypool.

Joseph J., J. L., Kruijff, E., McMahan, R. P., Bowman, D., & Poupyrev, I. P. (2017). *3d user interfaces: Theory and practice (usability).* Addison-Wesley Professional.

Olsen, L., Samavati, F. F., Sousa, M. C., & Jorge, J. A. (2009). Sketch-based modeling: A survey. *Computers & Graphics*, *33*(1), 85 - 103. Retrieved from http://www.sciencedirect.com/science/article/pii/S0097849308001295 doi: https://doi.org/10.1016/j.cag.2008.09.013

Sjölie, D. (2015, 12). *Approaching the matrix through synchronized reality.*

Stappers, P., & Giaccardi, E. (2013). Research through design. In *The encyclopedia of human-computer interaction, 2nd ed.* (chap. 43). Interaction Design Foundation.

Turner, N. (2017, 11). *Capturing user feedback with microsoft's product reaction cards.* Retrieved 2018-05-11, from http://www.uxforthemasses.com/product-reaction-cards/

Vitazko, M., & Zeller, M. (2018). *Case study - expanding the design process for mixed reality.* Retrieved from https://docs.microsoft.com/en-us/windows/mixed-reality/case-study-expanding-the-design-process-for-mixed-reality

Ware, C. (2012). *Information visualization, third edition: Perception for design.* Morgan Kaufmann.

Wensveen, S. A., Djajadiningrat, J. P., & Overbeeke, C. (2004). Interaction frogger: a design framework to couple action and function through feedback and feedforward. In *Proceedings of the 5th conference on designing interactive systems: processes, practices, methods, and techniques* (pp. 177–184).

Wigdor, D., & Wixon, D. (2011). *Brave nui world: designing natural user interfaces for touch and gesture.* Elsevier.

Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1083–1092).

Xiao, R., Schwarz, J., Throm, N., Wilson, A. D., & Benko, H. (2018). Mrtouch: Adding touch input to head-mounted mixed reality. *IEEE transactions on visualization and computer graphics*, *24*(4), 1653–1660.