11-17-2021

Hand-tracking Object Interaction System

Zihao Li
zl9422@rit.edu

RIT

Hand-tracking Object Interaction System

by

Zihao Li

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Fine Art in Visual Communication Design

School/Department of Design
College of Art and Design

Rochester Institute of Technology
Rochester, NY
November 17th, 2021

RIT | College of Art and Design

Thesis Approval

Hand-tracking Object Interaction System

Thesis Title

Zihao Li

Thesis Author

Submitted in partial fulfillment of the requirements for the degree of Master of Fine Arts

The School of Design | Visual Communication Design

Rochester Institute of Technology | Rochester, New York

Adam Smith

Name

Chief Thesis Advisor

Title

Electronic Signature: Use Adobe Acrobat

Mike Strobert

Name

Associate Thesis Advisor

Title

Electronic Signature: Use Adobe Acrobat

Name (Optional for some programs)

--- select a Title ---

Title (Optional for some programs)

Electronic Signature: Use Adobe Acrobat (Optional for some programs)

Name (Optional for some programs)

--- select a Title ---

Title (Optional for some programs)

Electronic Signature: Use Adobe Acrobat (Optional for some programs)

Abstract

At this point, VR, AR and MR technologies are thriving in different industries and business. They are increasing substantial productivity for large enterprises like Boeing and Ford at this point. Therefore, it’s well recognised that in future, VR, AR and MR will dive deeper into people's life, especially in to-business area. With VR, AR and MR being applied deeperly, more use cases and scenarios will show up.

However, from interaction design and user experience perspective, current interactions for VR, AR and MR is still very primitive and not user-centered enough along with several major problems. Problems include limited accessibility, being harmful to muscles, violations to social norms, lack of measurable precsions and so on. This might keep VR/MR from further influencing the industries and the world.

This project aims to offer an emerging gesture interaction system in VR that can help users manipulate virtual objects more confidently, more efficiently and more fluently with natural gestures and hand-tracking ability. More importantly, this project serves as an interaction template for different industries, that different industries can create their own interaction systems in terms of vertical needs based off this project. Final design aims to help define the next level of VR and MR user experience.

Keywords

Virtual reality, Mixed reality, Interaction design, UX design, Gesture interaction

With having analyzed plentiful of VR and MR demos online, current interactions for VR, MR meet with several major user-centered problems. The first problem is they rely on heavy body movements, that users have to reach out their arms and hands to manipulate virtual objects[1], which easily could exhaust users, and is physiologically harmful to muscles for long-term usage. The second problem is that, in many cases users have to repeat the same manipulations and movements for many times, which brings low productivity. The third problem is, it violates current social norms, as it's embarrassing and not safe to perform such interactions in public space[2]. Imaging you are in the metro and waving your arms, not only the others will feel embarrassed, but also it is easy to lead to accidents. The forth problem is, current interactions have limited accessibility. In many cases, users rotate virtual objects by twisting their wrist, which is physiologically limited within 85 degrees in vertical direction and within 50 degrees in horizontal direction[3]. The fifth problem is, current interactions are not ready for the complexity in real world either. For example, what if the virtual object is very huge or is overlapping with a real person. How can users interact reasonably with such objects? The last problem is, current interactions are hard to measure and fine-tune, that users can only feel and guess the outcomes. While fine-tuning, users have to spend even more muscle tensions and anxiety due to smaller scale and smaller velocity[4]. Therefore, the solution statement of this project is, to design a hand-tracking interaction system for VR and MR that offers fluent and handy experience with rich cues, optimal gesture ergonomics and measurable precisions.

The intention of design is to provide an interaction template for different industries and look into what future VR/MR interactions can be with current major interaction problems solved. Based on this template, industries can detailedly create their own interaction systems in terms of their vertical business and user portraits. For example, to teachers or educators who uses MR technology for teaching and demonstrate virtual objects in a class, in regular cases they may not need strong precisions or may not want any gesture or UI to frequently distract students, which means the interaction system here can remove some interactions for precision and visual effects to help better achieve the specific scenario goals.

The project is developed with several design strategies from different aspects. The first strategy is to divide the interactions by dimensions, by which users will only need to manipulate one dimension at a time. For example, users only need to focus on moving the virtual objects forward while not affecting other dimensions. This can reduce a lot cognitive load and pressure for users especially when MR interactions are new to them. Also, the interactions are divided and allocated for both hands, that one hand decides the interaction modes, the other performs the interaction. This ensures plentiful degrees of freedom for MR interactions while keeping them feasible. The second strategy is to design the ergonomic gestures for VR/MR. Considering nowadays users have got used to using mobile phones and seeing others using them, it's the best that the ergonomics utilizes the muscle memories trained by mobile phones and social norms formed by mobile phones. Therefore, the ergonomics mainly refers to mobile's ergonomics, and optimal interaction zone for UI in VR/MR is defined[5]. The UI in VR/MR will locates at where users would

---

[1] "Mixed Reality UX Elements - Mixed Reality". 2021. *Docs.Microsoft.Com*. https://docs.microsoft.com/en-us/windows/mixed-reality/design/app-patterns-landingpage.

[2] Shengzhi, Wu. 2020. "为什么我不看好AR眼镜？". *Zhihu.com*. https://zhuanlan.zhihu.com/p/138203496.

[3] "STRETCHING AND FLEXIBILITY - Normal Ranges Of Joint Motion". 1996. *MIT*. https://web.mit.edu/tkd/stretch/stretching_8.html.

[4] "Control Of Muscle Tension | Boundless Anatomy And Physiology". 2021. *Lumenlearning*. Accessed May 8. https://courses.lumenlearning.com/boundless-ap/chapter/control-of-muscle-tension/.

[5] Namwongsa, Suwalee, Rungthip Puntumetakul, Manida Swangnetr Neubert, Sunisa Chaiklieng, and Rose Boucaut. 2018. "Ergonomic Risk Assessment Of Smartphone Users Using The Rapid Upper Limb Assessment (RULA) Tool". PLOS ONE 13 (8): e0203394. doi:10.1371/journal.pone.0203394.

normally hold their mobile phones for daily usage. Moreover, a secondary interaction zone is defined for rapid responses, referring to how users hold their mobile phones for taking photos. The third strategy is to define visual and sound feedbacks for different interaction states and interaction actions. Geek blue is used for normal state, as it is the most common color in sci-fi movies and high-tech scenes. Orange is used as active state, and purple and green are used as special states, as they both feel unstable. Then, when it comes to feedback UI, it should be consistent with the task complexity, and refer to several industry products including Google Map AR[6], Tencent Map AR[7] and Apple AR guideline[8], which all presents less than three elements in their feedback UI. Therefore, the priority of different elements is ranked and the top three elements are chosen for feedback UI, which is number, direction and interaction mode. The forth and last strategy is to guide users' eye flow. Guiding user's eye flow is considered very important as users no longer stare at any physical screen in VR and MR environment. In fact, users stare at their daily life, which means they could look everywhere if their eye flows are not guided. The goal here is to minimize user's focal point movement, while also providing the intuitiveness of how the elements move. A technique called Rubber Band Animation is applied here. It can present the movement trends while reducing focal point movement.

Based on the design, an executable demo is prototyped in Unity3D, to measure how the interaction system generally work. Overall, the feasibility of the system is measured by technical index like FPS and the variation of FPS. The usability of the system is measured with assessments and feedbacks from interviewees.

The project contributes to current UX designs for VR and MR, and help define the next level of VR and MR experience. Xindi Liu, UX designer from Nreal, positively comments this project, "This system can work. It actually inspires me about the feedback UI design. One concern is that the limited tracking angle of motion cameras may lead to wrong results." Overall, the system works well in translating, scaling and rotating the virtual object with measurements. Users can intuitively manipulate the virtual object by simple and neutral gesture interactions, and handily see how the manipulation is performed in real time with clear feedback UI. Moreover, the system provides continuous manipulation options that free users from repeating the same behavior, which boosts up productivities. However, the system is limited by hand tracking technologies, and lacks more detailed interaction feedbacks like sounds and motions. In future, I intend to improve the system by digging into delicate micro-motions and sound feedbacks. While the outcomes are not prefect, the main strategies of this project is promising, which is to serve as an interaction template for industries to develop their own versions based off specific scenarios and user needs.

---

[6] "AR Home". 2021. Google AR & VR. https://arvr.google.com/ar/.

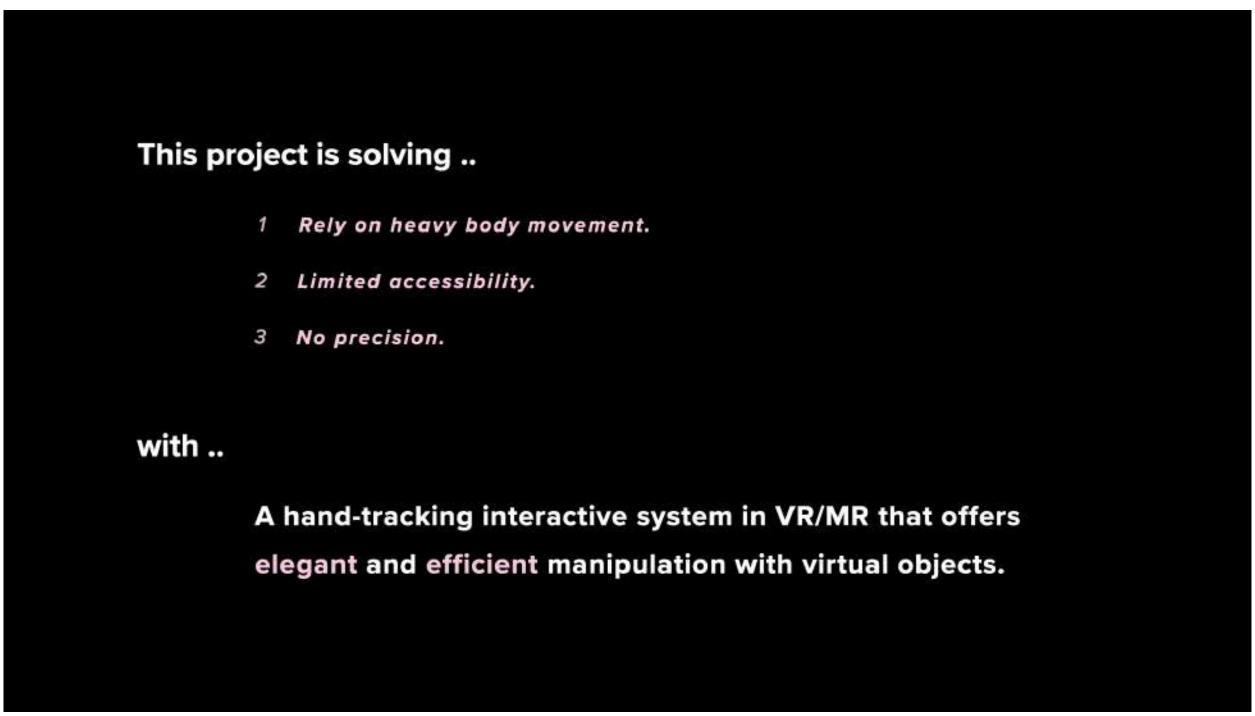
[7] "腾讯地图AR步行导航设计总结". 2021. 微信公众平台. https://mp.weixin.qq.com/s/fdjI9-GsahXVtHrPFuLGAA.

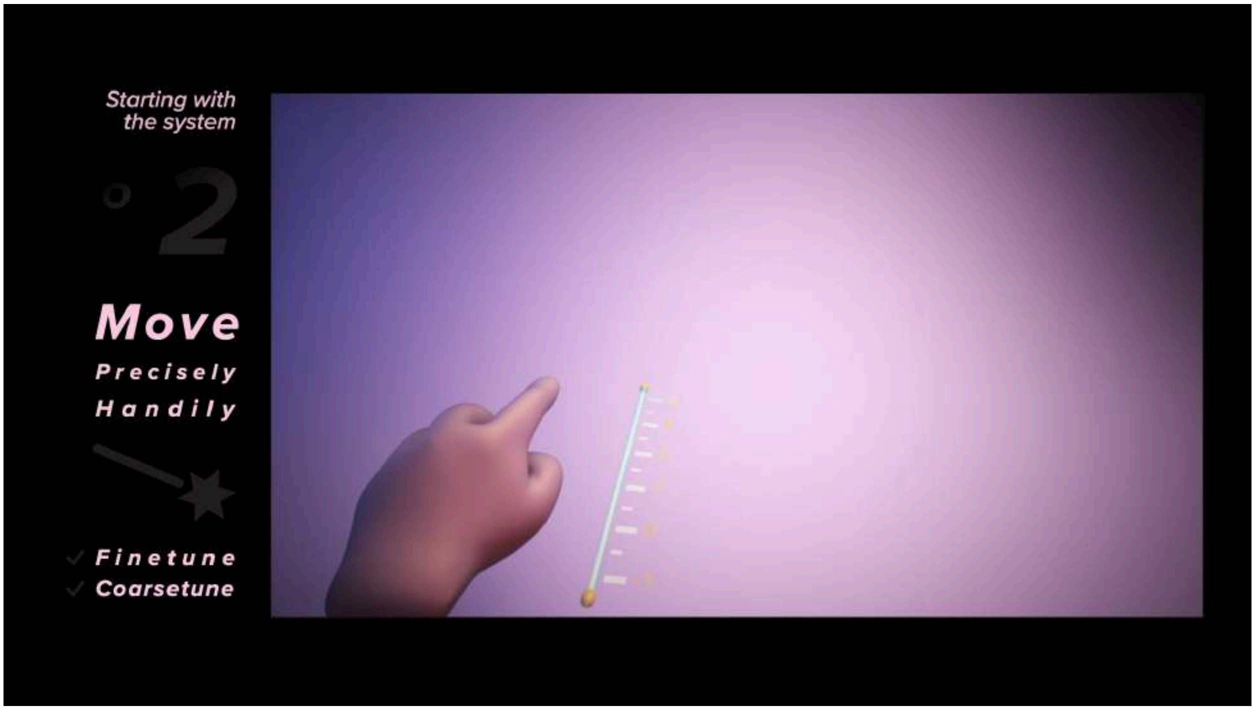
[8] "Augmented Reality - System Capabilities - Ios - Human Interface Guidelines - Apple Developer". 2021. Developer.Apple.Com. https://developer.apple.com/design/human-interface-guidelines/ios/system-capabilities/augmented-reality/.

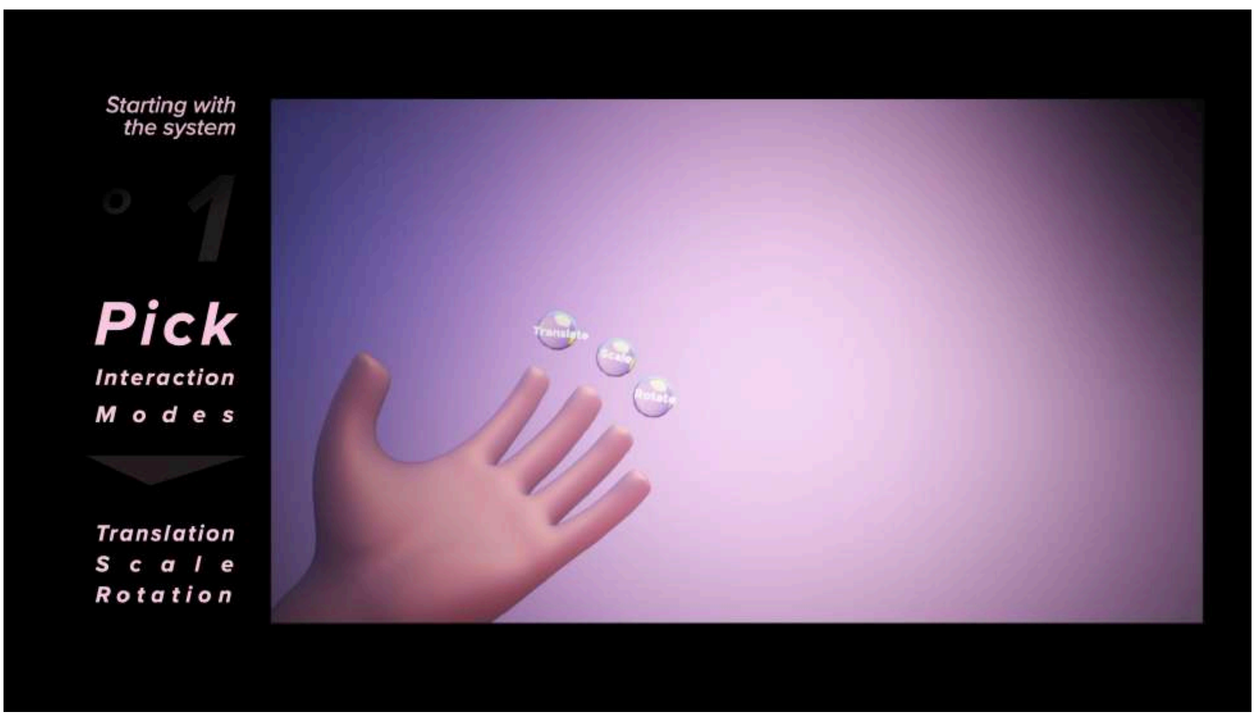
Appendix A: Bibliography and References

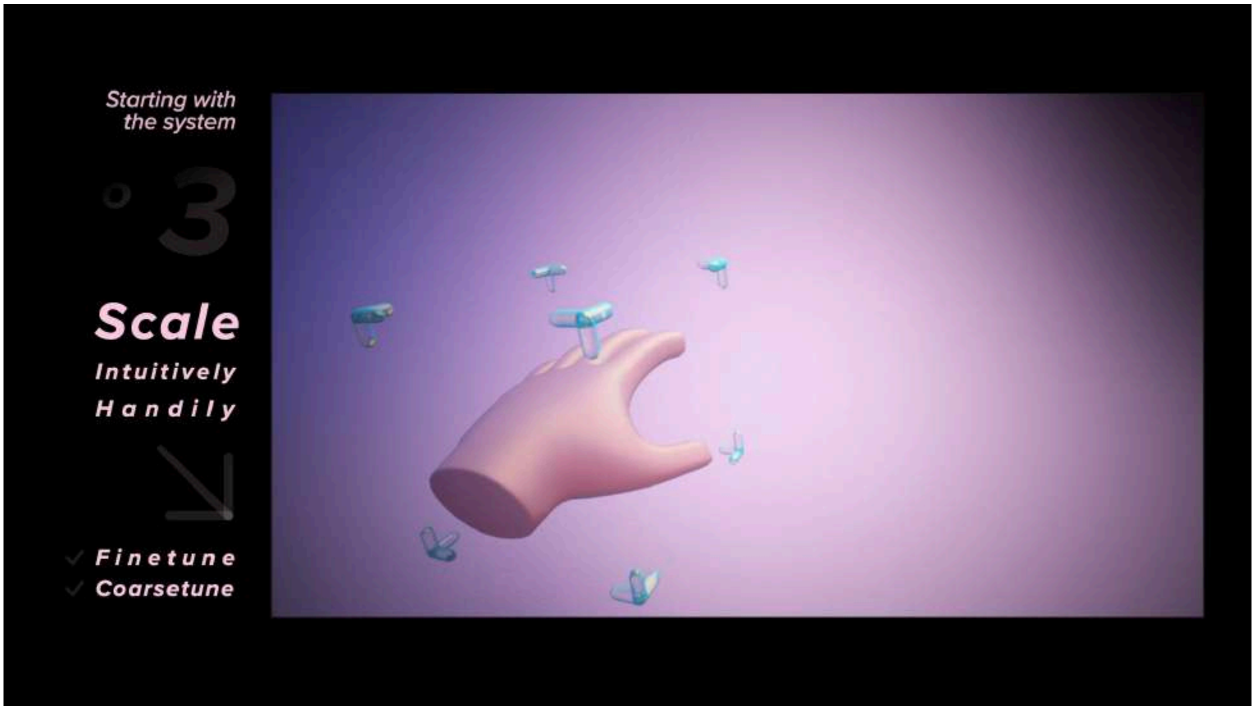
1. "Mixed Reality UX Elements - Mixed Reality". 2021. *Docs.Microsoft.Com*. https://docs.microsoft.com/en-us/windows/mixed-reality/design/app-patterns-landingpage.
2. Wu, Shengzhi. 2020. "为什么我不看好AR眼镜？". *Zhihu.com*. https://zhuanlan.zhihu.com/p/138203496.
3. "STRETCHING AND FLEXIBILITY - Normal Ranges Of Joint Motion". 1996. MIT. https://web.mit.edu/tkd/stretch/stretching_8.html.
4. "Control Of Muscle Tension | Boundless Anatomy And Physiology". 2021. Lumenlearning. Accessed May 8. https://courses.lumenlearning.com/boundless-ap/chapter/control-of-muscle-tension/.
5. Namwongsa, Suwalee, Rungthip Puntumetakul, Manida Swangnetr Neubert, Sunisa Chaiklieng, and Rose Boucaut. 2018. "Ergonomic Risk Assessment Of Smartphone Users Using The Rapid Upper Limb Assessment (RULA) Tool". PLOS ONE 13 (8): e0203394. doi:10.1371/journal.pone.0203394.
6. "AR Home". 2021. Google AR & VR. https://arvr.google.com/ar/.
7. "腾讯地图AR步行导航设计总结". 2021. 微信公众平台. https://mp.weixin.qq.com/s/fdjI9-GsahXVtHrPFuLGAA.
8. "Augmented Reality - System Capabilities - Ios - Human Interface Guidelines - Apple Developer". 2021. Developer.Apple.Com. https://developer.apple.com/design/human-interface-guidelines/ios/system-capabilities/augmented-reality/.
9. "Airbus Uses Microsoft Hololens 2 To Improve Production Of Aircraft - Mspoweruser". 2019. Mspoweruser. https://mspoweruser.com/airbus-uses-microsoft-hololens-2-to-improve-production/.
10. Politansky, Mathew. 2021. "How Microsoft's Hololens Increases Productivity - The Mobile Apps Man". The Mobile Apps Man. https://themobileappsman.com.au/microsofts-hololens-increases-productivity/.
11. Baro, Victor. 2015. "Recreating Apple'S Rubber Band Effect In Swift". Medium. https://medium.com/thoughts-on-thoughts/recreating-apple-s-rubber-band-effect-in-swift-dbf981b40f35.

Appendix B: Thesis Defense Presentation




Starting with
the system
2
Move
Precisely
Handily
Finetune
Coarsetune


Starting with
the system
1
Pick
Interaction
Modes
Translation
Scale
Rotation
Translate
Scale
Rotate


Starting with the system
°3
Scale
Intuitively
Handily
Finetune
Coarsetune


Starting with the system
°4
Rotate
Precisely
Handily
Finetune
Coarsetune

Dissolving ___
the system

Interaction Model

1

Divide the work by dimension

Only handle one direction at a time.

2

Assign the work to both hands

Non-dominant hand decides interaction modes, dominant hand performs interactions.

Dissolving ___
the system

Ergonomics

1

Utilise muscle memory for mobile

People have adapted to raising arms for mobile phones.

2

Utilise social affordance formed by mobile

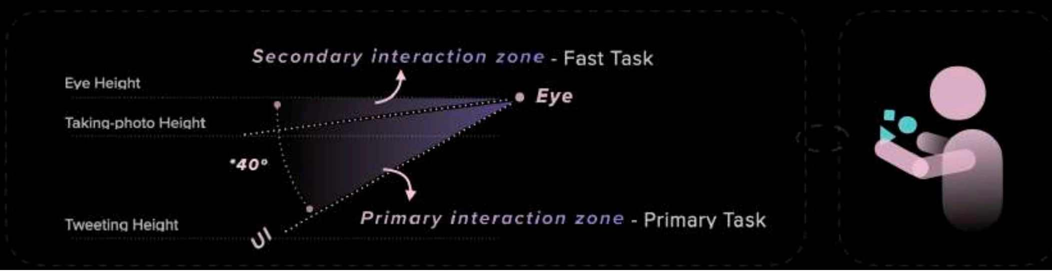
People have adapted to seeing others raising their arms for mobile phones.

*Dissolving ___
the system*

Ergonomics

1 **Place interactions at mobile level.**

2 **Limit behaviours within specific extent.**



*Dissolving ___
the system*

Visual Cues

1

Consistent with task complexity

2

Clarify information ambiguity +-

Dissolving the system

Visual Cues

1 **Number (distance, angle, etc)**

2 **Direction**

3 **Interaction Mode**

Dissolving the system

Focal Movement

1

Place visuals where interactions go

When performing interaction, users tend to pay attention to where interaction happens.

2

Follow hands' trends as rubber band

When hands move, users tend to see dynamic elements move along hands.

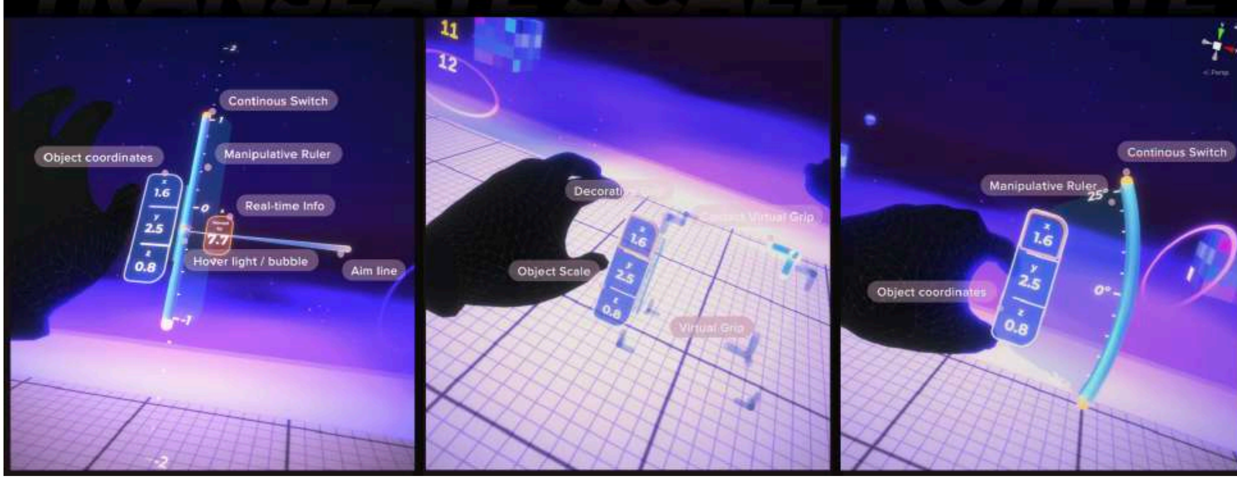
Validating _
the system

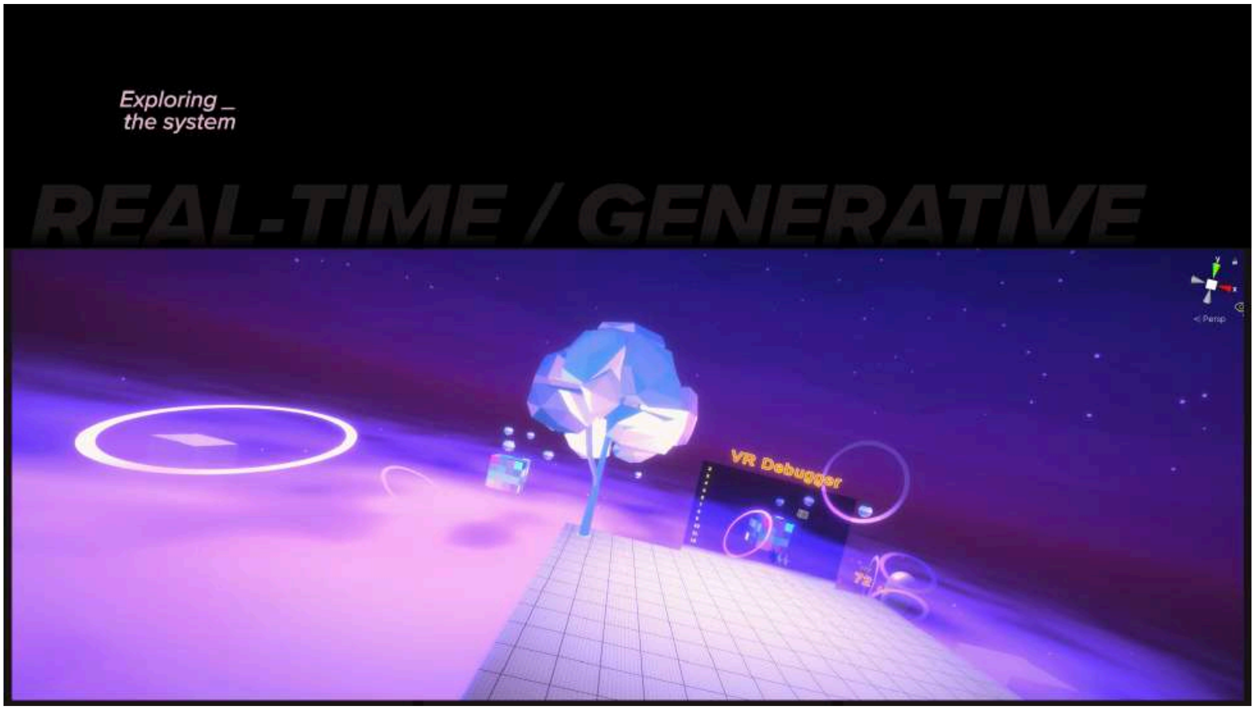
Unity Demo →

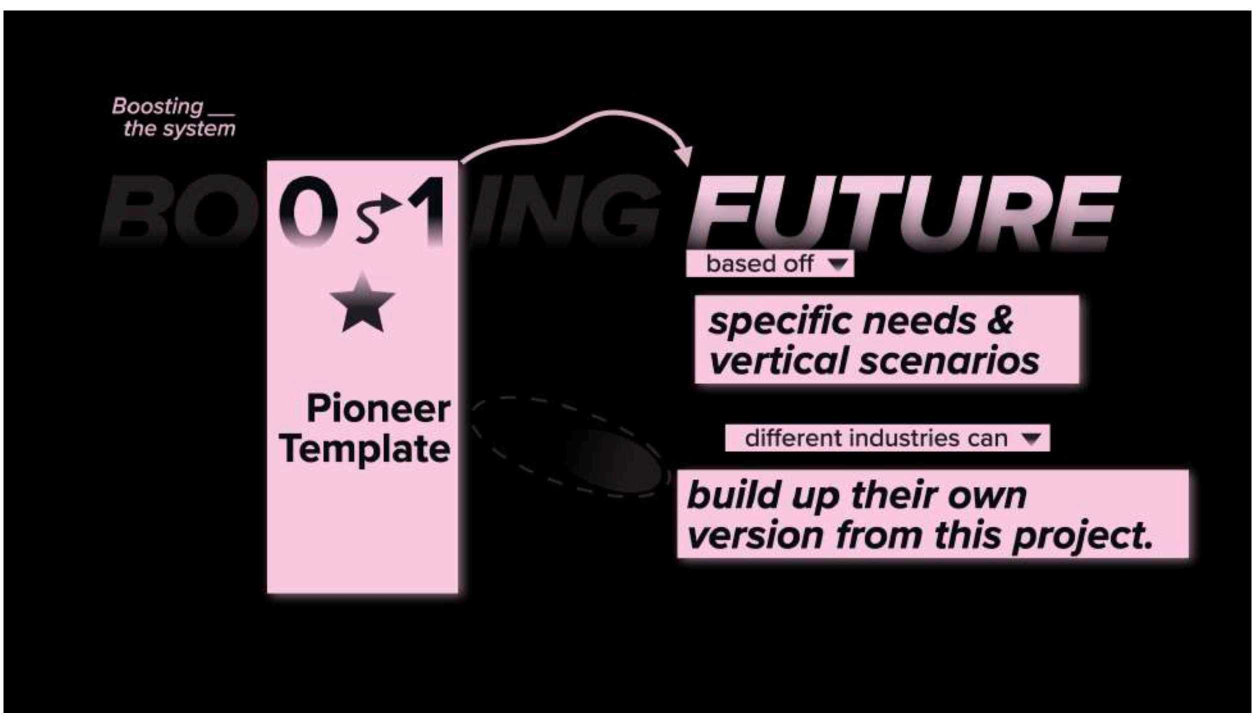
Xindi Liu AR/VR Designer from Nreal	This can work. Actually the design thinking here inspires me a little. However, the system might be limited when the hands move outside the tracking range.

Exploring _
the system

TRANSLATE SCALE ROTATE

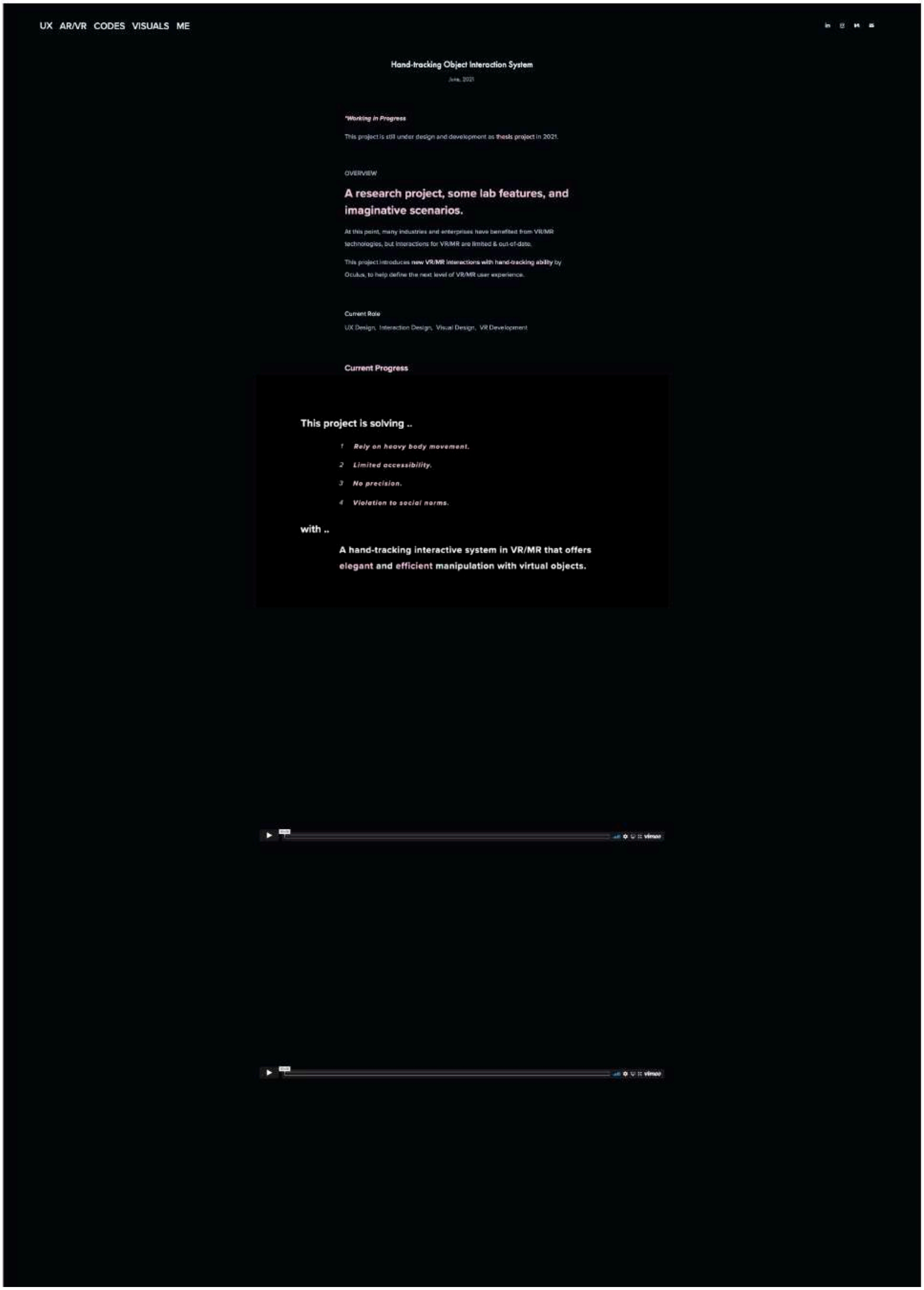



Exploring _
the system
REAL-TIME / GENERATIVE
VR Debugger


Boosting __
the system
BO
0
1
ING
FUTURE
Pioneer
Template
based off
specific needs &
vertical scenarios
different industries can
build up their own
version from this project.


Thank you

Appendix C: Screen Capture of Thesis Website



Dissolving __ the system

Interaction Model

1
Divide the work by dimension

Only handle one direction at a time.

2
Assign the work to both hands

Non-dominant hand decides interaction modes, dominant hand performs interactions.

Dissolving __ the system

Ergonomics

1
Utilise muscle memory for mobile

People have adapted to raising arms for mobile phones.

2
Utilise social affordance formed by mobile

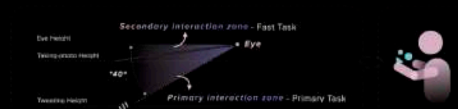
People have adapted to seeing others raising their arms for mobile phones.

Dissolving __ the system

Ergonomics

1 Place interactions at mobile level.

2 Limit behaviours within specific extent.

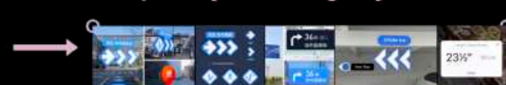


Dissolving __ the system

Visual Cues

1
Consistent with task complexity

2
Clarify information ambiguity +-



Dissolving __ the system

Dissolving __ the system

Visual Cues

[1] Number (distance, angle, etc)

[2] Direction

[3] Interaction Mode

Dissolving __ the system

Focal Movement

1

Place visuals where interactions go

When performing interaction, users tend to pay attention to where interaction happens.

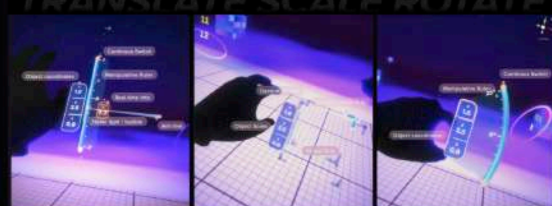
2

Follow hands' trends as rubber band

When hands move, users tend to see dynamic elements move along hands.

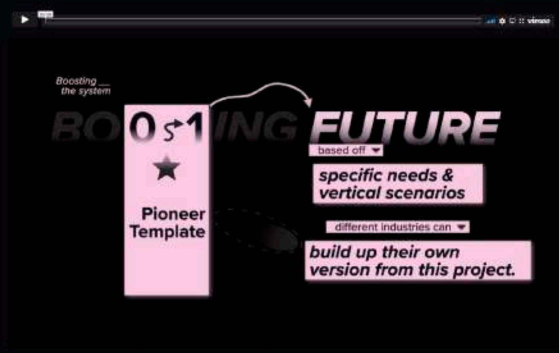
Exploring _ the system

TRANSLATE SCALE ROTATE



Boosting __ the system

BO 0→1 ING FUTURE



Background

• • •

Industries of healthcare, vehicle, manufacture, aerospace ... already benefit a lot from VR/MR.

statistics

- ***Engineers*** work ***30%*** faster and ***96%*** more accurately.
- ***Warehousing*** productivity increases from ***15%*** to ***46%***.
- ***Airbus*** reduces inspection time from ***3 weeks*** to ***3 days***.
- ***Ford*** drops ***90%*** engineering issues.
- ***Boeing*** increases ***40%*** productivity.

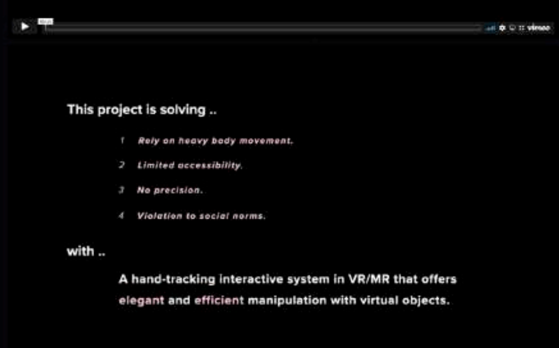
Problem

!!.

Rely on heavy body movement.

exhausting* | *unhealthy* | *low-efficiency* | *anti-social-norms

- Heavy muscle tension, physiologically unhealthy for long-term usage.
- A lot repetitions of interactions, low productivity.
- Violating social norms, embarrassing in public.



Visual hint to hover   Visual hint to grab / pinch   Dynamic UI

UX Principle - Ergonomic Gestures

Ergonomic and social norms adaptation.

Refer to mobile's ergonomics

- Utilise muscle memory trained by mobile.

  People have adapted to raising arms for mobile phones.

- Utilise social affordance formed by mobile.

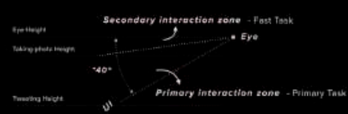
  People have adapted to seeing others raising their arms for mobile phones.

UX Principle - Ergonomic Gestures

Ergonomic and social norms adaptation..

Refer to mobile's ergonomics

- Place interactions at mobile level.
- Limit behaviours within specific extent.



UX Principle - Visual Cues

Real-time UI..

Consistent with task complexity

- Number (distance, angle, ...)
- Direction
- Manipulation (word / icon)

Clarify information ambiguity

- Do not use '+' or '-' that could be ambiguous.
- Use indicator to point out direction.

UX Principle - Feedback Cues

System color..

Geek Blue

Most common color in futuristic scenes.

Orange

Activated and can be cancelled anytime.

Pink Purple

Feels like "overcooked" or "overshoot".

Chartreuse

Feels like "poisoned" or "withering".

UX Principle - Eyeflow

Minimise focal point movement..

Functional affordance

- When performing interaction, users tend to pay attention to where interaction happens.

  Change visuals & place UI at where interaction happens.

- When hands move, users tend to see dynamic elements move along hands.

Change visuals & place UI at where interaction happens.

- When hands move, users tend to see dynamic elements move along hands.

  Interactive visual elements follow hands' trends.

Real-time Generative Cloud via Shader Graph

- Use shader components to generate cloud patterns with color combinations and fading effect.

Immersive monitor visuals via Shader

- Use shader code to create a unique visual effect via a unique blending mode for debugger's monitor.

Working in Progress

Coding, coding & coding..

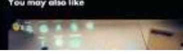
- ***VR Debugger***
- ***Gesture Detector***
- ***Universial Easer***
- ***Hand-tracking Object Interactive System***

  Hand-tracking Object Interactor
  Hand-tracking Object Translator & UI Manager
  Hand-tracking Object Scaler & UI Manager
  Hand-tracking Object Rotator & UI Manager

Part of UniversalEaser.cs

Thank you for scrolling :)

You may also like



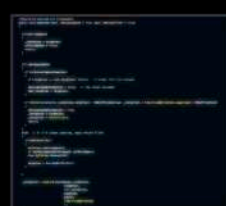
Immersive monitor visuals via Shader

- Use shader code to create a unique visual effect via a unique blending mode for debugger's monitor.

Working in Progress

Coding, coding & coding..

- ***VR Debugger***
- ***Gesture Detector***
- ***Universial Easer***
- ***Hand-tracking Object Interactive System***
  - Hand-tracking Object Interactor
  - Hand-tracking Object Translator & UI Manager
  - Hand-tracking Object Scaler & UI Manager
  - Hand-tracking Object Rotator & UI Manager



Part of UniversalEaser.cs

Thank you for scrolling :)

You may also like