# Research Statement

Zhenyi He (Ph.D. candidate, Future Reality Lab,
Department of Computer Science, New York University)

My research aims to address how to enable and improve users' interaction in immersive environments. I've created a series of work to explore how sketching facilitates creative collaborative Virtual Reality (VR), how to accelerate productivity in VR / AR through different input systems like handwriting and in-place typing, as well as how robotic graphic – providing timely haptic feedback – improves distant VR collaboration. I envision that people are able to communicate and collaborate through creative content and productive input system in immersive environments. I will summarize my previous work and describe what I plan to contribute in the future.

**keywords**: human-robot interaction; collaboration; text entry; sketch;

### *Research #1: CollaboVR and Creative Authoring*

Writing or sketching on whiteboards is an essential part of collaborative discussions in business meetings, reading groups, design sessions, and interviews. However, prior work in collaborative VR systems has rarely explored the design space of multi-user layouts and interaction modes with **virtual whiteboards**. I designed CollaboVR [1], a reconfigurable framework for both co-located and geographically dispersed multi-user communication in VR. Our system unleashes users' creativity by sharing freehand drawings, converting 2D sketches into 3D models, and generating procedural animations in real-time. To minimize the computational expense for VR clients, I leverage a cloud architecture in which the computational expensive application (Chalktalk) is hosted directly on the servers, with results being simultaneously streamed to clients. I have explored three custom layouts – integrated, mirrored, and projective – to reduce visual clutter, increase eye contact, or adapt different use cases. We have investigated CollaboVR on teaching and collaborative authoring.

As one branch of collaboration research, we further explore a **non-VR collaborative set-up** for the increasing requirements of working remotely. Considering that video conferences play a significant role in working remotely, we research whether and how to facilitate video conferencing by bringing back missing nonverbal cues, including gaze and spatial information. We implemented LookAtChat, a web-based video conferencing system, which empowers remote users to identify eye contact and spatial relationships in small-group conversations. Leveraging real-time eye-tracking technology available with ordinary webcams, LookAtChat tracks each user's gaze direction identifies who is looking at whom, and provides corresponding spatial cues. We explored the design space of eye contact visualization in both 2D and 3D layouts and evaluate LookAtChat in three conditions: baseline layout, 2D directional layout, and 3D perspective layout.

As another branch of creative collaboration, we extend the capability of Chalktalk as our **creative content** platform, which has been configured as the cloud-based application in CollaboVR framework. Chalktalk [6] is a paradigm for creating drawings in the context of a face-to-face brainstorming session that is happening with the support of VR or AR. Participants draw their ideas in the form of simple sketched simulation elements, which can appear to be floating in the air between users. Those elements are then recognized as animated and interactive "sketches" by a simple AI recognition system and can be interactively incorporated by users into an emerging simulation that builds more complex simulations by "linking" together these simulation elements in the course of the discussion. All "sketches" are designed to have an input and output protocol so "sketches" can be "linked" to each other like Lego blocks. Additionally, the input and output values help shape the behavior, e.g. once we connect the thermometer to a fan, the fan will be turned on when the temperature is higher than the threshold. The idea of Chalktalk was further explored in the recognition system, interaction method (phone movement, touchscreen, and controllers), surface drawing, and collaborative modes. The concept of Chalktalk is not only behaving as a live presentation tool, but "sketches" that have personalized behaviors could be "linked" for combined behaviors as well.

### *Research #2: TapGazer and Handwriting*

VR / AR is known as a good fit for entertainment use. Besides, they have strong potential for serious work like what was presented in "office of the future" from UNC and "sixth sense" from MIT. I've implemented an early version of hand-based user interface [2] and considered to bring immersive environments with more productivity through text input: handwriting and typing.

We implemented a handwriting approach with haptic feedback in VR for surface writing when interning in Oculus Research Pitt (now FRL). Users can either hold a regular pen or conduct a pinch gesture to perform writing tasks. Gestures are detected based on the input of head-mounted cameras with a revised convolutional pose machines (CPM) framework. The accuracy of **keypoint detection** reaches 98.3%.

Alternatively, we propose a typing solution that maintains users' standard QWERTY kills. We envision that smartphones may be replaced by inexpensive and lightweight eyeglasses capable of displaying augmented reality. Ideally, those future glasses will incorporate gaze detection, and will also track hand and finger movements. Wearers of those glasses should then be able to type text simply by tapping their fingers on any surface. Thus, we implemented TapGazer, a text entry system in which users type by tapping their fingers, **without needing to look at their hands** or be aware of their hand position. Ambiguity is resolved at the word level, by using gaze for word selection. In the absence of gaze tracking, disambiguation can be effected by additional taps. User studies show that TapGazer works with different devices and is easy to learn, with beginners reaching 52.17 words per minute on average, achieving 77% of their QWERTY typing speed with gaze tracking and 58% without. Currently, we are working on investigating whether and how TapGazer affects standing or walking while typing.

***Research #3: VR Representation***

In the meantime of researching on the HCI side, I researched the representation of participants and bystanders that may affect users' subjective feedback and behaviors in VR.

When interning in Facebook People AI Face Tracking team, we investigated how **avatar representation** affects users' communication and collaboration in videoconferencing scenarios. I implemented a 2D video image representation and a 3D avatar representation for comparison. To solve the uncanny feeling of the 2D video image in VR, I designed no-change, cropped, segmented, frame, and portal. Similarly, we presented outpost [5] in Siggraph this year that presents a distributed social VR experience on an alien planet. Participants are allowed to choose different avatar representations for the main experience. The data was collected to figure out how different attributes including representation affects communication efficacy.

When VR provides immersive experiences in the virtual world, it may potentially reduce or even remove users' awareness of their physical surroundings, resulting in safety concerns and psychological discomfort. Hence, there is a clear need for ambient information design to increase users' **situational awareness** (SA) of critical elements (people and objects) in those parts of the physical space that are reachable in the virtual world. We explored the efficacy of five representational fidelities (indexical, symbolic, and iconic with three emotions) of ambient entities in VR [4]. Results of an empirical study showed that all five designs can evoke SA, but to various extents, while consistency with the immersive environment can mitigate breaks in presence.

***Research #4: PhyShare***

Although VR provides a realistic visual experience of the virtual world and the opportunity of using limited physical resources to establish haptic communication between users and the virtual environment, the absence of effective haptic feedback will strongly detract from the suspension of disbelief needed to bridge the virtual and physical worlds. To address such tradeoff, we present PhyShare [3], a new **haptic user interface** based on actuated robots. PhyShare presents multiple mappings between physical robots virtual proxies, such that the bots can represent one or multiple virtual objects, or even collaborate to represent the same scenario in different locations. We designed a tabletop configuration for board-game scale interaction, such as tic-tac-toe, and a floor-based configuration for room-scale interaction, such as "escape the room". It is highly confirmed that haptic feedback is beneficial for VR interaction, besides, we further researched the topic from the collaboration perspective and provided an alternative for latency issues introduced by the hardware.

**Future Research Agenda**

I am passionate at facilitating collaboration that requires creative content creation, productive interaction between users and system, and design space in collaboration. CollaboVR is a typical of my approach.

- Creative Content Creation
  An easy-to-use, friendly, and creative tool for content creation is beneficial for users to share their ideas or even imagination. A system that enables to inspire people during creating content can bring fun to the result and procedures as well.

- Cross-device Interaction
  More and more mobile devices and sensors are available now. A workspace that allows various devices or sensors to adapt to input (touch on surface or gaze) or output (shape display) is beneficial for people to interact with the system.

- Collaborative AR / VR Design
  Design becomes complex and useful when the system allows for collaboration. Additionally, combining sociability with asymmetric or time-travel manner worth exploring.

# References

[1] **He, Zhenyi**, Ruofei Du, and Ken Perlin. "CollaboVR: A Reconfigurable Framework for Creative Collaboration in Virtual Reality". In: *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2020).

[2] **He, Zhenyi** and Xubo Yang. "Hand-based interaction for object manipulation with augmented reality glasses". In: *Proceedings of the 13th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*. 2014, pp. 227–230.

[3] **He, Zhenyi**, Fengyuan Zhu, and Ken Perlin. "Physhare: Sharing physical interaction in virtual reality". In: *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST)*. 2017, pp. 17–19.

[4] **He, Zhenyi**, Fengyuan Zhu, Ken Perlin, and Xiaojuan Ma. "Manifest the Invisible: Design for Situational Awareness of Physical Environments in Virtual Reality". In: *Arxiv Preprint* (2018).

[5] Kris Layng, Michael Gold, Benjamin Ahlbrand, **He, Zhenyi**, and Ken Perlin. "The outpost". In: *ACM SIGGRAPH 2020 Immersive Pavilion*. 2020, pp. 1–2.

[6] Ken Perlin, **He, Zhenyi**, and Fengyuan Zhu. "Chalktalk vr/ar". In: *International SERIES on Information Systems and Management in Creative eMedia (CreMedia)* 2017/2 (2018), pp. 30–31.