Statement of Purpose

It is often said that the uniqueness of human intelligence lies in its ability to learn and use tools. This is embodied in how humans, as intelligent agents, understand the physical world, interact with it, and ultimately solve real-world problems. While physics seeks to distill the laws governing the physical world into simple principles, computer science explores how to leverage these principles to acquire information about the real world, uncover patterns, and create new realities. The use of various neural networks in AI to perceive and understand the world has become highly sophisticated, yet the domain of AI learning to interact with the physical world still holds immense potential for further development. For me, the most captivating aspect is how agents can learn new skills and transform the world through interaction with the physical environment.

Therefore, my intellectual interests focus on embodied AI, which combines reinforcement learning (RL) and data-driven methods to enable agents to acquire skills through real-world interactions. This technology holds significant potential for applications in robotics, where it can assist humans, and in graphics, where it can enhance character animation and even physical simulation to generate richer and more realistic content.

Through my research experiences in the Visual Computing and Learning (VCL) Lab at Peking University and during my research internship at the University of Illinois Urbana-Champaign (UIUC), I have delved into intriguing problems within this field, gained valuable insights into research methodologies and forming my perspectives. My previous work has focused on using RL for physics-based control to extract physical information from rich data sources such as motion capture and video, enabling the generation of physically plausible character animations, including realistic interaction with external objects and environments. Notably, while existing video generation models often deliver visually striking effects, they frequently violate fundamental physical principles. I firmly believe that integrating research with real-world physics exploration can yield more robust results in generation, also with potential applications in robotics to enable agents to perform diverse tasks.

If given the opportunity to join UIUC's graduate program, I aspire to deepen my exploration of the learning agents' capabilities and learn from the exceptional faculty and peers in the program. I am confident that UIUC's abundant resources and vibrant academic atmosphere will support my pursuit of advancing embodied AI for physical-world interactions.

My earlier physics education sparked my understanding of intelligent agents in the physical world, while my transition to computer science and conducting research in the field equipped me with the skills to approach these questions from the perspective of computing and AI. Initially, I majored in physics as an undergraduate student at Peking University, where I studied fundamental topics such as theoretical mechanics.

Inspired by computer graphics' ability to enrich simulations of the physical world, I developed a keen interest in using computer science to tackle these challenges. This interest led me to transfer to the computer science department, where I will graduate with a degree in computer science, having also built a solid foundation in computer science knowledge and coding skills.

At the computer science department, I cultivated my interest in graphics through courses like Character Animation and Physics-based Simulation, taught by Professor Libin Liu. I also joined the VCL lab under Professor Liu's guidance. In the course, I implemented kinematics-based animation techniques to achieve real-time locomotion control. My research expanded upon this foundation, focusing on extracting physical properties from motion capture and video data to develop control strategies for humanoid agents interacting with physical environments.

One of my initial projects explored character fatigue effects during skill performance. Leveraging RL architectures enhanced with generative models like GANs and VAEs, I aimed to constrain joint torques with physical fatigue principles to produce novel and physically consistent motion results. The RL-trained policies demonstrated motions reflecting fatigue dynamics, with the agent adapted its responses based on the complexity of tasks, aligning with my goal of integrating physical characteristics into animation. This project honed my skills in RL and generative model training, providing a strong foundation for future work.

Building on this success, my discussions with Professor Liu and lab peers inspired us to extend our approach to broader motion styles. Instead of solely relying on kinematics-based machine learning models to summarize stylistic patterns, I pursued a physics-based perspective, aiming to uncover interpretable motion styles with controllable physical properties. Adopting a policy fine-tuning structure inspired by AdaptNet, and aligned with stylization methods like AdaIN and CycleGAN, we designed an architecture for multi-task, multi-style physics-based generation. This research is still ongoing, with promising potential for future breakthroughs.

In the summer of 2024, I had the privilege of joining Professor Yuxiong Wang's group at UIUC to explore human-object interaction (HOI) control. Recognizing HOI as a critical aspect of agent-physical world interactions, I aimed to leverage RL methods enhanced by generative models to learn HOI control policies. My initial efforts involved applying GAN-based architectures to learn HOI tasks like ball-holding and dribbling, yielding successful results with policies that dynamically adapted to unseen situations. Currently, we are investigating compliance control in HOI, focusing on how torque influences policies to achieve more natural motions and safer, more stable task executions for real-world robotics applications.

These research experiences have deepened my understanding of physical-world interactions and the broader concept of embodied AI. I have also benefited from interdisciplinary discussions with researchers in adjacent fields, learning to incorporate advancements from LLMs, LVMs, and diffusion models into my research. These experiences have strengthened my abilities to independently conceptualize and execute research projects while collaborating effectively with diverse teams. Although these projects inevitably came with conceptual challenges, I view research as a journey of continuous learning and creative exploration."

The opportunity to collaborate with leading researchers at UCSD, particularly working with Professor Xiaolong Wang and Professor Hao Su, excites me greatly. Professor Wang's innovative research on humanoid robot control and reinforcement learning, such as his recent work on "Mobile-TeleVision: Predictive Motion Priors for Humanoid Whole-Body Control," resonates deeply with my experience in HOI motion control. The PMP and CVAE approach in this research, enabling multi-task interaction for humanoid robots, aligns closely with my interests in leveraging physics-based methods for general interaction control generation.

Similarly, Professor Hao Su's groundbreaking contributions to 3D reconstruction and its role in enabling agents to comprehend 3D worlds strongly align with my vision for embodied AI. I believe advancements in 3D reconstruction are fundamental to building intelligent agents capable of interacting seamlessly with their environments.

UCSD's research-focused master's program, which combines rigorous coursework with abundant opportunities for cutting-edge exploration, offers the ideal environment for me to advance these interests. I am confident in my ability to thrive in UCSD's computer science master's program and look forward to contributing meaningfully to advancements in robotics, embodied AI, 3D understanding, and related fields.