Statement of Purpose

My primary research objective and interest is in the area of robot perception. I am currently a research and design engineer in HTC VIVE, and I am actively involved in a research project that is developing object recognition for VR/AR. My work focuses on analyzing point cloud from stereo-scene to recognize objects and obstacles in the indoor environment and lining up a safe playing region for VR/AR players. This project links up my SLAM(Simultaneous Localizationa and Mapping) and Robotics background, and some ideas spring in my head. For example, “How to make point cloud(or other sensor data) become semantic for robot?” and “How to use dynamic objects or semantic objects in the map to benefit SLAM?” These ideas stimulate me to undertake further study through your research program.

Before joining the object recognition project, I was a SLAM researcher in “Advanced and Creative Team”(AC-Team) in HTC VIVE, focusing on localizing VR device. The special effects in VR’s virtual world depends on users’ motion in the real world. Though SLAM becomes well-studied in these years, the robustness of SLAM for VR devices are a still challenging topic. Unlike robot or self-driving car, the SLAM algorithm on VR device has to tolerate fast scene changing because human head move in this way. Therefore, in order to conquer it, we implemented tightly-coupled visual-inertial SLAM. The performance of IMU generally is powerful in fast moving situation. Also, IMU’s speedy update rate compensates camera frames’ duration. Furthermore, we developed “Multi-Frame System”, which means that our SLAM is able to handle multiple cameras and IMUs at the same time, and estimating an optimized state of VR device in the environment. I studied comprehensive SLAM and VIO papers in these years such as ORB-SLAM, DSO, LSD-SLAM and VINS, and combining their strengthens into our product. For example, for the front-end of SLAM, we used direct method to efficiently obtain camera’s state. Nevertheless, the feature-based(ORB) SLAM still be in our algorithm because the robustness of ORB help relocate the camera’ state. On the other hand, for the back-end, we also adopted IMU pre-integration to our state estimation. The Jacobian of IMU’s bias are considered. These research results are implemented to my career milestone “VIVE COSMOS”, which is the newest VR device in HTC VIVE.

HTC VIVE taught me that collaboration and communication are the disciplines to accomplish a mature product. I handled many projects in HTC VIVE, and in some of them, I was the leader or the core engineer. Developing a commercial product requires an immense of consideration. Finishing personal research is just a part of a product. Conversely, organizing a team to be productive is a big thing. For example, my SLAM algorithm has to be implemented on different VR/AR products. Different products have different requirements. In order to fit my SLAM into each product, I had to discuss with firmware team to secure CPU’s computing resource, evaluate sensors with multi-media team to make sure that the data are stably offered by sensors, and inform the mechanic team how important is the location of cameras for SLAM’s robustness. These members all comes from different background and different professionals. I prefer integrate my teammates by sorting the priority and explaining why. For instance, cameras’ positions on VR device are always critical because we have to make sure that the scene obtained by cameras is broad enough to keep VR devices tracking in any environment. In this case, I would like to tell my teammate that the camera position issue is at high priority and explain the reason why I do that instead of directly command them. And because they know the reason, different teams can communicate based on the reason and make decision based on the priority. I believe that sorting the priority can keep all members at the same direction, and explaining the reason make the direction clear.

The competition in HTC VIVE stimulates me to keep to improve my research ability and problem solving ability. The management of HTC VIVE prefers to allocate a topic to different teams in order to spread risk. That means only one team’s research result will be product and earn money. AC-TEAM is talents’ gathering. It is not easy to be outstanding in this environment. In order to competing these talents, I cultivate the habit of updating professional knowledge and improving my research ability. I take one hour every morning to watch top conferences videos or some technology news. If the video is strongly related to my professionals or my research interests, I will read its paper even download its source code and learn some from it. Sometime these knowledges can interact and develop new idea. This habit has been kept three years. In the first two years, I concentrated more on SLAM and robot perception. However, I gradually change my focus to machine learning and robot reasoning in this year because machine learning become powerful in computer vision in these decades. I consider that machine learning will be broadly applied in robot perception in the future.

My research journey started from Robotics Lab under Prof. Han-Pang Huang in the graduate school of Mechanical Engineering at National Taiwan University (NTU). Robotics Lab equipped me with diversity robot-related knowledge because it is the largest lab in the graduate school of mechanical engineering. Robotics Lab contain many fields, including manipulation robot, humanoid (biped) robot and mobile robot. The most attracting thing in Robotics Lab is that the members like to discuss and are glad to share research result and give suggestion. During the discussion, my robot knowledge is broadly extended, and some idea and research are figured out. For example, my master thesis combines SLAM and humanoid robot’s path planning. This research is beholden to the researchers of both humanoid (biped) robot and mobile robot. The experience in Robotics Lab constructed my research foundation. Until today, in HTC VIVE, I use Robotics knowledge on my work.

For these reasons, the Robotics Institute at Carnegie Mellon University is especially attractive to me. I am intrigued by several interesting research projects carried on by its faculty members. In particular, Professor Sebastian Scherer’s research in autonomous aerial robots is fascinating. His research recently tends to integrate robot perception and automatic control with machine learning. This integrated research is what I desire to do. For example, *“Visual Memorability for Robotic Interestingness via Unsupervised Online Learning”* introduces the concept of interestingness detection. We know that “how to decide keyframe or keypoint?” is always a critical issue for robot mapping, especially in repeated monotonous scene case. Air Lab beautifully conquered the problem by three learning strategies. In addition, Air Lab is good at fusing SLAM and path planning. *“DROAN - Disparity-space Representation for Obstacle AvoidaNce”* describes wonderful consensus of mapping and obstacle avoidance. Therefore, Air Lab is the reason why I apply the programe in CMU RI.

My career goal is to develop a robot system that senses the world like human beings. Robot perception becomes powerful in these years. However, there is still challenges. For example, human beings sense objects in the environment directly by objects’ location instead of point cloud or voxel. “How object recognition feedback to robot’s state estimation?” is not well-defined as I know. To study this question and the questions I mentioned in the first paragraph, I desire to do research in advanced degree, and I believe that the Robotics Institute at Carnegie Mellon University is good fit for me. The resources and breadth of robot research performed at CMU will provide me with the experience that I need to pursue my passion in robot perception.