Motivation Letter

My name is Shiyu Gao (See Homepage.), this letter demonstrates my motivation, qualification and enthusiasm for your EDIC PhD programme.

Having completed my four-year bachelor of engineering program in 2019, I've been working in 3D vision for three years. Computer vision enables machines to see, observe and understand with visual inputs such as images and videos, while computer graphics try to "draw" visual output and make it differentiable and explainable. But there are still many puzzles waiting to be solved, for example, how to represents a 3d scene, how to use the powerful representation capabilities of neural networks to construct a robust mapping between 2d and 3d, taking into account both photometric consistency and geometric consistency, as well as the shape and texture of the object? In recent years, Nerf[1] present Neural Radiance Fields for scenes representation. But nerf does not give an answer on how to parameterize the scene and make it interpretable and editable. There is still a lot to explore.

My experiences at college and research labs not only equip me with immense theoretical knowledge and extensive experiment experiences, but also inspire my burning passion for further research in related fields.

I got my bachelor's degree in Detection Guidance and Control Technology from Harbin Institute of Technology (HIT, qs ranking #217) in 2019. During my undergraduate studies, I learned from the theories behind the electronic products to core control technologies. I gradually learned to simulate and control dynamic systems with numerical analysis and optimization methods. (See project: PID controller of Quadrotor.) With strong interests in how machines see things, I joined the HIT Robotics Team and developed an infantry robot vision system for ROBOMaster Robotics Competition. (See project: Vision System.) Our team developed an algorithm to detect the light bars on the target armors and calculated rotation and translation matrix from camera to target by solving PnP Problem. This competition experience has ignited my curiosity towards computer vision and motivated me to pursue further study in computer vision.

Because of outstanding academic performance (Weighted GPA: 89/100 and ranking at top 3) in my undergraduate program, I was admitted to a straight-to-PhD program in the Institute of Computing Technology (ICT), University of Chinese Academy of Sciences (qs ranking #112), where I start my CV journey and finished my first conference paper. The past three years in ICT has been the luckiest and unluckiest time to me because it confirmed my passion and talent in computer vision while witnessed my confusion and struggle in the group. As part of my direct doctorate program, I learned courses about computer science as well as computer vision (GPA: 3.4/4.0). Having been equipped with sufficient theoretical knowledge, I start working on 3D vision. I worked with Dr. Zhaoxin Li on multi-view stereo and other 3D vision projects (See Project: Car-Simulation & Human Rendering UE4.) from Sep. 2020 to Mar. 2022.

Apart from doing some engineering projects， I have published conference paper independently. My research interests lie in the field of 3D vision, especially multi-view stereo, 3D reconstruction and 3D detection. Basically, we are looking for an optimal mapping from 2D to 3D, which makes machine to perceive 3D world from semantic as well as geometric information of 2D inputs. Our work: "Cost Volume Pyramid Network with Multi-strategies Range Searching for Multi-view Stereo" (See arxiv preprint.) has been accepted by CGI2022. In this work, we propose multiple strategies to fully exploit information in cost volume and improved the completeness of reconstructed point clouds. In the process of doing research, I have gained a deep understanding of multi-view geometry.

In this year, I started my internship at Baidu Research (A technology company known for its autonomous driving and search engine). I've been working on knowledge distillation of transformer (See My blog) for about 3 months until July.

After that, I've been working on 3D segmentation & detection (See [My blog](#)) till now. My research experience in Baidu enables me to stay on the cutting edge of academia, thus I can follow a new topic in related fields (such as ViT, nerf, and detection) in a short time.

I'm open to various topics in CV & CG in my future work, because many problems are related, for example, how to capture the semantic and geometric information of 3D world from 2D images while having a strong capability of generalization is the central problem shared by both 3D Detection and multi-view stereo (See SOLOFusion [2]).

Having been through three years of harsh doctoral training, I find myself still passionate about research and still so determinate to get my PhD degree. But I realized that I need more open-minded collaborators as well as creative atmosphere. EPFL provides a world-class, highly collaborative international research environment, which is the key to bring about innovative and solid works.

Empowered with theoretical knowledge, strong programming skills, research experience and countless failures, I'm confident and determined to complete the PhD program under the supervision of Prof. Wenzel Jakob, thus sincerely submitting my application for your consideration. Thank you very much for your reading and evaluation.