Introduction

As a master's candidate in the Robotics, Cognition, and Artificial Intelligence program at the Technical University of Munich (TUM), I have built a strong foundation in machine learning and computer vision. As I approach graduation, I am eager to pursue a PhD position in Professor Nießner's group within the Department of Informatics at TUM. I am deeply passionate about inverse rendering and neural reconstruction, which facilitate the reconstruction of 3D scenes and estimation of plausible material and scene parameters from a collection of images that depict the scene.

Research Experience

Before delving into computer vision, I conducted research as a physics student at the Technical University of Berlin. Under Prof. Dopfer's guidance, I performed spectroscopy experiments on isolated bioflavonoids in the gas phase. This work led to one first-author publication and two additional co-authored published works.

My fascination with computer science and machine learning emerged during my time in Prof. Dopfer's lab when we aimed to optimize experimental procedures. We transitioned from analog controls for our electrostatic lenses to a centrally controlled digital voltage controller. At that point, I was unaware that my manual adjustments of lens voltages—lenses that guided ions through the experiment—constituted an empirical optimization process within a high-dimensional problem space. Switching to digital voltages enabled us to store voltage states for high ion yields and to adopt a genetic algorithm that autonomously explored feasible settings, allowing us to focus on more complex tasks.

During my research fellowship in Prof. Nießner's lab, I focused on reconstructing complex scenes from single RGB images by training a neural field to predict colored 3D meshes from individual images. To achieve this, I trained a neural network to estimate a signed distance field and color for every point in a canonical volume, using local and global image features extracted from a Convolutional Neural Network as conditioning input. In my subsequent M.Sc. thesis, I explored the use of auxiliary image features, such as depth and normal maps, to train Neural Radiance Fields (NeRFs) more effectively. This investigation led to the development of an innovative loss function for depth supervision, which can be directly derived from a Gaussian likelihood of the transmittance function for a ray traversing a neural radiance field. This approach surpassed the prevailing state-of-the-art methods on widely-used ScanNet and Blender scenes.

Current and Future Research

Moving forward, I am particularly keen on advancing my studies on NeRFs, analyzing trends and the current rate of progress in the field, while taking into account the growing capabilities of devices. Although much work remains and the precise mechanisms of neural radiance fields are yet to be fully understood, it appears that the rapid improvements in NeRF accuracy and efficiency will soon lead to static NeRFs becoming commonplace, seamlessly integrating into daily life. This development necessitates a focus on dynamic scenes, which, despite their increased complexity, offer the potential to significantly broaden the scope of real-world applications. A converging trend is the emergence of mid-tier smartphones, released from 2020 onwards, equipped with high-quality RGB-D sensors. When combined with neural radiance fields, these sensors enable even amateurs to create top-quality virtual assets. Moreover, for image collections lacking depth data, monitoring trends in networks predicting auxiliary 2D image information — such as monocular depth, normal maps, and semantic maps — suggests that these could provide high-quality inputs for reconstructing more accurate neural radiance fields.

The research opportunities lie in furthering the general understanding of neural radiance fields, and I have identified two avenues that I am particularly interested in exploring:

1. The inherent 3D consistency of 2D constraints for multi-view stereo reconstruction with NeRFs is a direct result of dense and accurate supervision. However, there is insufficient work being done to ensure 3D consistency of monocular inputs when supervision is sparse. Efficient use of readily available auxiliary 2D information could dramatically decrease the data and supervision signal needed to produce high-quality NeRFs. The key to tackling this challenge lies in the intelligent use of 2D priors. I plan to build upon my previous work by devising mechanisms that ensure 3D consistency of auxiliary 2D image signals, such as machine-generated monocular depth maps and semantic maps, to match and co-locate plausible keypoints for 3D reconstruction in dynamic scenes. Progress in this area would anticipate the availability of abundant smartphone RGB-D data and could lead to substantial quantitative and qualitative improvements. Furthermore, I believe that the fact that depth maps implicitly encode empty space is underappreciated and could result in a significant speed increase for training and inference of NeRFs.
2. In the optimization process, we assume no errors in our given parameters, namely the camera pose, camera intrinsics, and image pixel values (colors). It is worthwhile to consider partially erroneous inputs, for example, by determining how errors in pixels, camera pose, and intrinsics propagate to generated rays. Additionally, treating generated camera rays traversing the neural volume as statistical processes would allow for the quantification of network uncertainty for its outputs and the derivation of bounds and fuzzy constraints. This approach could potentially lead to a more robust formulation of NeRFs for real-world usage.

Conclusion

In conclusion, my solid foundation in machine learning and computer vision, coupled with my passion for inverse rendering and neural reconstruction, make me an exceptional candidate for a PhD position in Professor Nießner's group at the Technical University of Munich. My research experience in both physics and computer vision has provided me with a distinctive perspective on the challenges and opportunities in the field of Neural Radiance Fields. As I delve deeper into this area, I am dedicated to making significant contributions to the understanding and development of NeRFs, ultimately unlocking their full potential for real-world applications. With the guidance and support of Professor Nießner and the TUM Department of Informatics, I am confident that my research will have a lasting impact on the future of computer vision and its practical applications.