

Mission Statement

When I had the chance to try Tesla's autopilot, I was more than enthusiastic. After all, I studied computer vision for nearly two years to understand how difficult autonomous driving is and what state-of-the-art techniques can achieve. However, on the busy highways around Frankfurt am Main, the autopilot was barely able to be active for a minute without human intervention.

This experience motivated me to work towards fully autonomous cars by enabling machines to develop a human-like understanding of the world around us – specifically, to research high-level 3D scene understanding. I believe that 3D scene understanding is a key component in autonomous systems – not only autonomous cars – with the need to operate in real environments.

I am confident that I can truly contribute to the field of autonomous driving by doing original research in computer vision. In the future, I want to play an important role in making autonomous cars reality both regarding the technical aspects as well as the necessary public discussion. I identified a PhD as the ideal framework to become a thought leader in computer vision by devoting my full attention to research.

Background

At RWTH Aachen University, through seminars on neural networks and deep learning (see [1]) as well as two teaching opportunities in mathematics, I got introduced to the academic world and began considering a PhD as challenging opportunity after graduation.

Advised by Prof. Bastian Leibe, I wrote my bachelor thesis on "Superpixel Segmentation using Depth Information" [2]. Parts of the results were published at the German Conference on Pattern Recognition (GCPR) 2015 as student paper [3]. In addition, a journal paper providing a comprehensive comparison of available superpixel algorithms [4] is in preparation. Both publications spurred my ambition to practice computer vision research on a full-time basis and equipped me with valuable skills including academic writing and reproducible research. Above all, this experience taught me perseverance regarding my research goals and interests.

An exchange semester at the Georgia Institute of Technology further strengthened the idea of pursuing a PhD in computer vision. Thanks to Prof. Bastian Leibe, I was welcomed in Prof. Irfan Essa's group and worked side-by-side with his PhD students. Addressing video segmentation and intrinsic images/video, I was able to devote more time to research than before. This experience not only benefited my intercultural skills, but also contributed to my understanding of original research. I learned to question my beliefs of what is possible and to value one-to-one advising in the academic environment.

In industry, two experiences are particularly related to my career goals. First, I worked with Dr. Stefan Holzer and Dr. Alexander Trevor at Fyusion on computer vision problems including line segment and keypoint tracking, statistical clustering and pedestrian detection. Second, supervised by Dr. Thomas Guthier from MOBIS, I was introduced to advanced driver assistance systems working on deep learning for pedestrian detection. Based on the exchange of experiences, I started to recognize the value of a PhD as opportunity to immerse myself deep into challenging topics. Moreover, I became interested in autonomous driving, especially going beyond pure (pedestrian) detection and towards higher-level scene understanding.

Currently, I am working in Dr. Andreas Geiger's group at the MPI in Tübingen. Motivated by his work and through his supervision, I became particular interested in 3D scene understanding. As result, my master thesis focusses on the fundamental problem of 3D shape completion. Specifically, given LiDAR data from a moving autonomous car, I want to learn how to complete the 3D shapes of detected cars without supervision. The thesis is intended as a foundation for my future PhD studies – regarding 3D scene understanding as long-term research goal as well

as the necessary research skills. So far, the work in Dr. Andreas Geiger's group helped me to develop concrete research interests and taught me how to organize my research and develop as well as communicate my own research ideas. The interaction with other motivated PhD students additionally allowed me to continuously reflect upon and improve my approach to research.

Research Interests

In general, 3D scene understanding is a challenging task due to the complexity of real-world environments – be it cluttered indoor scenes or crowded traffic scenes – as well as difficulties arising from 3D data. Recent advances in deep learning have already had major impact on tasks such as optical flow, object detection and tracking, pose estimation or semantic segmentation. However, most of the approaches are limited to image data. Additionally, as indicated by Prof. Amnon Shashua [5] for the case of autonomous driving, scene understanding involves additional semantic layers on top of these tasks to infer useful information – in other words, the individual tasks need to be combined in a reasonable way.

3D scene understanding has been tackled in the past. In the case of autonomous driving, Dr. Andreas Geiger's work [6] combines several individual tasks in order to infer a higher-level understanding of traffic scenes. I want to revisit this problem by extending deep learning techniques to 3D data and jointly learning the involved tasks end-to-end. In contrast to Prof. Amnon Shashua's believe [5], I am convinced that jointly learning these tasks is crucial to leverage context and obtain a complete understanding of real-world scenes. Based on the (possibly limited) overview as a student, I believe that this approach to scene understanding is both promising and considerably new.

Max Planck Institute for Intelligent Systems

The contact to researchers from the MPI in Tübingen had significant influence on my decision to specifically apply for this PhD program. During my exchange semester, I got to know Prof. Michael Black and Dr. Naejin Kong as part of a short collaboration. Later, talking to Dr. Peter Gehler gave me further, valuable insights regarding PhD program structure, work environment and international orientation. Finally, since January, I am working on my master thesis in Dr. Andreas Geiger's group. His work and guidance as well as close contact to his PhD students and collaborators has considerable influence on my research interests and career goals. After three months in Tübingen, I determined the MPI in Tübingen as perfect fit for my PhD studies.

Conclusion

In conclusion, I am motivated to begin my career with a PhD at the International Max Planck Research School for Intelligent Systems. Equipped with a unique mindset and valuable skills from industry and academia, I want to shape the impact of deep learning on 3D scene understanding for autonomous driving by collaborating with world leading researchers and like-minded students.

References

- [1] D. Stutz. Personal Website & Blog. <http://davidstutz.de>.
- [2] D. Stutz. *Superpixel Segmentation Using Depth Information*. Bachelor Thesis, 2014.
- [3] D. Stutz. *Superpixel Segmentation: An Evaluation*. German Conference on Pattern Recognition, 2015.
- [4] D. Stutz, A. Hermans, B. Leibe. *Superpixels: An Evaluation of the State-Of-The-Art*. CVIU (accepted). ArXiv Preprint: <https://arxiv.org/abs/1612.01601>.
- [5] A. Shashua. *What Goes into Sensing for Autonomous Driving?* Goldman Sachs Cars 2025, 2016.
- [6] A. Geiger. *Probabilistic Models for 3D Urban Scene Understanding from Movable Platforms*. Ph.D. Thesis, 2013.