The Journal of Open Source Software

TraViA: a Traffic data Visualization and Annotation tool in Python

Olger Siebinga[1]

1 Human-Robot Interaction group, Department of Cognitive Robotics, Faculty 3mE, Delft University of Technology, Mekelweg 2, 2628 CD Delft, the Netherlands

DOI: 10.21105/joss.03607

Software

- Review
- Repository
- Archive

Editor: Daniel S. Katz

Reviewers:

- @rusu24edward
- @tomalrussell

Submitted: 28 July 2021
Published: 21 September 2021

Summary

In recent years, multiple datasets containing traffic recorded in the real world and containing human-driven trajectories have been made available to researchers. Among these datasets are the HighD, pNEUMA, and NGSIM datasets. TraViA, an open-source Traffic data Visualization and Annotation tool was created to provide a single environment for working with data from these three datasets. Combining the data in a single visualization tool enables researchers to easily study data from all sources. TraViA was designed in such a way that it can easily be extended to visualize data from other datasets and that specific needs for research projects are easily implemented.

Statement of need

The combination of drones, cameras, and image recognition techniques might sound like a recipe for a spy movie. But actually, this combination allows for the collection of rich traffic datasets. The recipe is straightforward: hover a drone above a location with traffic, record a video, and use image recognition to generate bounding boxes for all vehicles. The result is a dataset containing human-driven trajectories at the location of interest that can be used for many scientific purposes, e.g., to study traffic flow, model human behavior, or design autonomous vehicle controllers.

Because the required ingredients are easily accessed all over the world, multiple such datasets have been published in recent years. In Germany, the highD project (Krajewski et al., 2018) recorded all traffic at 6 different high-way locations; in Athens, Greece, all traffic in the city's business district was recorded using 10 drones for 5 days in the pNEUMA project (Barmpounakis & Geroliminis, 2020); and American highway traffic was recorded using fixed base cameras in the NGSIM project (U.S. Department of Transportation Federal Highway Administration, 2016). Combined, these datasets span different countries, types of vehicles, and environments, a combination valuable for researchers with different backgrounds. Example usages of these datasets are validating human behavior models (e.g., by Talebpour et al. (2015) and Treiber et al. (2008)) or testing autonomous vehicle controllers (e.g., by Schwarting et al. (2019)).

Currently, it is difficult to leverage the powerful combination of multiple datasets because all the datasets come in different formats, and it is often difficult to get a good and real-time visualization of the data. Some visualization tools exist (one is provided with the highD data (Krajewski et al., 2018) and another example for NGSIM data can be found in Sazara (2017)) but they are specifically made for only one of these datasets and are very basic in the sense that they provide little control over simulation time and no insight in raw values per vehicle per frame. In addition to difficulties with visualization, finding, and annotating situations of interest in these massive datasets is a time-consuming task and keeping track of the annotations for the different datasets requires some bookkeeping skills.

TraViA was developed to provide a solution for these problems. TraViA can be used to visualize and annotate data from highD, pNEUMA, and NGSIM and uses generic vehicle objects to store the state of vehicles at a specific time. This makes it possible to validate and test models or controllers on multiple datasets in parallel, without having to cope with the different dataset formats.

Software Functionality

TraViA is written in Python 3 and has a graphical user interface developed in PyQt5. A screenshot of TraViA is provided in Figure 1. This screenshot shows the capabilities of TraViA in a single image. The main features of TraViA are:

- Advanced information display based on raw data for every vehicle in every dataset by leveraging generic vehicle objects
- Dynamic visualization of the traffic scene with possibilities to zoom, pan, and rotate for an optimal view
- Exporting the visualization to a video or single image
- An interactive timeline that shows dataset annotations, which are saved as python objects for easy manipulation

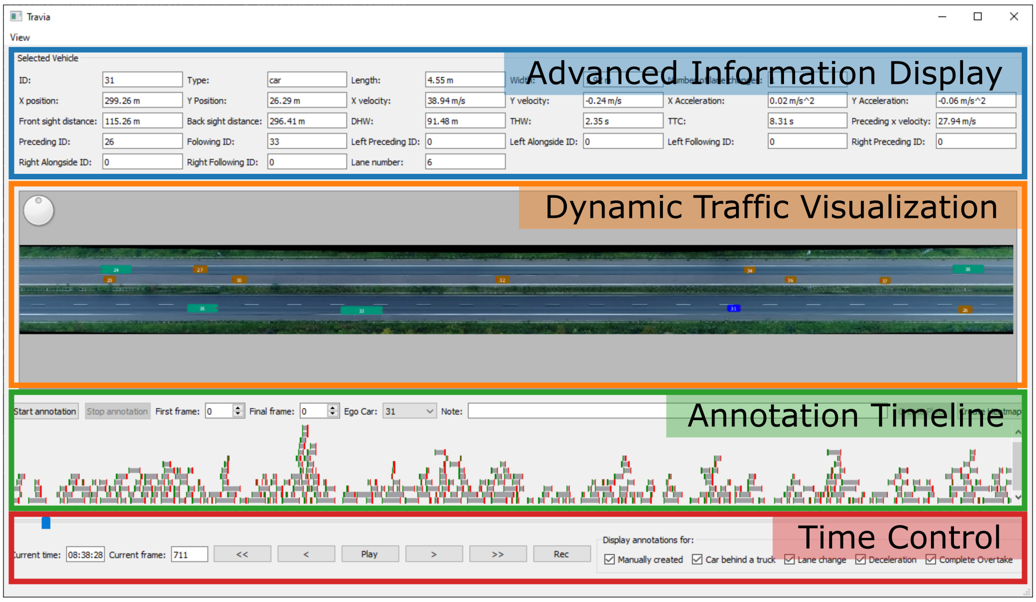

Figure 1: A screenshot of the TraViA software visualizing a frame of the highD dataset. The main features of TraViA are highlighted in this image.

TraViA was designed for use as a stand-alone program. It uses abstract classes as a basis for all dataset-specific objects to enable easy implementation of new datasets (for a class diagram and more information on how to do this, please see the readme file in the repository). It was specifically created to serve as a tool for generic visualization and annotation such that it can be used by researchers from different fields. To show the capabilities of TraViA and to provide a starting point for other researchers that want to use TraViA for their work, three example implementations of tools for specific purposes are included with TraVia. The first example is the functionality to automatically detect and annotate specific scenarios (e.g., lane changes), the second is functionality to plot specific vehicle signals over the course of an annotation, and the third is a function to plot a heatmap overlay for use in autonomous vehicle reward function development. All of these example tools are only implemented for use with the highD dataset.

Usage of TraViA in Science and Education

Currently, TraVia is being used by the author for model validation of an inverse-reinforcement-learning-based driver model. A publication on this validation is currently being prepared for submission. Besides that, TraViA is used for educational purposes, allowing students at TU Delft to explore big naturalistic datasets by providing them with an accessible, GUI-based starting point.

Acknowledgements

I thank Nissan Motor Co. Ltd. for funding this work, and I also thank my supervisors David Abbink and Arkady Zgonnikov for their valuable help and advice. Finally, I thank Joris Giltay for his help with testing the instructions in the readme file.

References

Barmpounakis, E., & Geroliminis, N. (2020). On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment. *Transportation Research Part C: Emerging Technologies*, *111*(November 2019), 50–71. https://doi.org/10.1016/j.trc.2019.11.023

Krajewski, R., Bock, J., Kloeker, L., & Eckstein, L. (2018). The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2018-Novem*, 2118–2125. https://doi.org/10.1109/ITSC.2018.8569552

Sazara, C. (2017). *NGSIM-trajectory-animation*. https://github.com/cemsaz/NGSIM-trajectory-animation.

Schwarting, W., Pierson, A., Alonso-Mora, J., Karaman, S., & Rus, D. (2019). Social behavior for autonomous vehicles. *Proceedings of the National Academy of Sciences*, *116*(50), 24972–24978. https://doi.org/10.1073/pnas.1820676116

Talebpour, A., Mahmassani, H. S., & Hamdar, S. H. (2015). Modeling lane-changing behavior in a connected environment: A game theory approach. *Transportation Research Part C: Emerging Technologies*, *59*, 216–232. https://doi.org/10.1016/j.trc.2015.07.007

Treiber, M., Kesting, A., & Thiemann, C. (2008). How Much does Traffic Congestion Increase Fuel Consumption and Emissions? Applying a Fuel Consumption Model to the NGSIM Trajectory Data. *Transportation Research Board*, *August*.

U.S. Department of Transportation Federal Highway Administration. (2016). *Next Generation Simulation (NGSIM) Vehicle Trajectories and Supporting Data. [Dataset]*. https://doi.org/10.21949/1504477