

Vehicle Trajectory and Gap Estimation for Conflict Prediction

Data Collection

What data will you collect or create?

Surveillance video of vehicles in traffic situations. We will try to re-use existing data at CMU of the intersection of Forbes and Craig St, as well as other data available from the Federal Highway Administration research facility. In addition, there are existing publicly available vehicles in traffic videos that have been used by researchers to establish benchmarks and capabilities. We expect the data to be video recorded at 25 - 30 frames per second with standard RGB colors in common video formats such as avi, mp4, etc. These are the most common and widely used formats which can be decoded by almost any system. We will also use data we collected at the Pittsburgh Airport parking facility, and have already made publicly available. If we need to collect our own data, we will also make this data available through the CMU Library data sharing Kilthub. No personal identifiable information will be collected. However, we currently do not expect to collect new data during this research.

How will the data be collected or created?

The data will be collected from existing repositories by other researchers and our own existing data sets which have been made public. Similar to the data collection at the PIT airport parking lot, the cameras will be observing vehicle from light or traffic signal poles, but no license plates or individual faces will be identifiable. Video data will be organized by camera, date and time. All cameras will be calibrated for extrinsic and intrinsic parameters using standard calibration techniques.

Documentation and Metadata

What documentation and metadata will accompany the data?

The camera type and calibration parameters will be stored in a file, together with the cameras location information. Together with the date and time information, this will allow anyone with a video player to access the video and understand the context.

Ethics and Legal Compliance

How will you manage any ethical issues?

The positioning of the cameras will not allow identification of vehicle license plates or individuals. Thus there will be no need for anonymization. No audio was recorded in any of the currently targeted video data sets. If we do need to collect our own new data, we expect this to fall under the category of "not human subject research" as was determined in the past.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

While the data will be owned by various researchers that have published their work, we expect to similarly have a license to use this data for research and publications. Any newly collected data will be property of CMU, but we intend to make it available to other not-for-profit research.

Storage and Backup

How will the data be stored and backed up during the research?

The data will initially be stored on redundant RAID drives in the SCS machine room, which already exist for our other projects. Backup services will be provided by SCS facilities according to their established schedules.

At the end of the project the data will be given to the CMU Libraries as part of the KiltHub data dissemination and preservation work.

How will you manage access and security?

Access to the data will be password restricted to members of the research group only. The SCS machine rooms are only physically accessible to authorized personnel.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

We expect the videos to have some long-term value, although this will diminish over time. In the past videos of this type tended to lose their value after about 10 years, as better recording technologies became available.

What is the long-term preservation plan for the dataset?

The long term preservation plan is to utilize the CMU Library KiltHub services for preservation and dissemination.

Data Sharing

How will you share the data?

Our data and analysis results will be freely shared with others. We will publish and disseminate our research findings and the resulting data at conferences and in journal articles.

Are any restrictions on data sharing required?

We currently don't expect to put any restrictions on the data sharing.

Responsibilities and Resources

Who will be responsible for data management?

Alex Hauptmann as PI of the project will be responsible for data management.

What resources will you require to deliver your plan?

We currently have resources in form of a cluster of GPU servers with 2-4 NVIDIA 1080TI and 2080TI GPUs available, which should be sufficient for most of the computational processing. We also have a dedicated storage server in the machine room to save data and results during the project duration.