Preliminary establishment and analysis of automated driving benchmark in China

Hui Xiong, Yinghan Wang, Qiang Ge, Ziqing Gu, Mingyang Cui, Jianqiang Wang, Keqiang Li and Qing Xu

State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Tsinghua Park NO. 1, Beijing, China

Abstract. In order to promote the localization of Automated Driving (AD) in China, it is necessary to collect large-scale traffic scene data with Chinese characteristic for future analysis. In this paper, we propose the methodologies and rules of establishing AD benchmark involving how to configure sensors, how to design the collection schema to show Chinese traffic characteristics and the rules of elaborating distinctive scenes and routes, what to label, and it is also demonstrated that the benchmark can support the future application of extended AD research. Data collection lasted about one month covering diverse scene data such as campus, highway, park, etc. from three representative Chinese cities and driving data from 30 different drivers. Moreover, some statistical results and analyses are produced in accordance with the designed methodologies as instances of potential application. Up to now, the dataset contains about 7,000 labelled image frames and corresponding LiDAR, GPS and Controller Area Network (CAN) data. Labels cover scene type, road user, traffic sign, traffic light, and lane marker. This benchmark can help researchers better understand Chinese traffic situation in aspects of environmental perception, driving behavior analysis, risk assessment, automated vehicle decision and control.

1 Introduction

In the research on the adaptability of AD by identifying and analysing Chinese unique complex traffic environment, also in order to find more key research topics in the field of AD, firstly researchers need to establish benchmark standards and build the collection platform. Afterwards a road environmental survey on scene features, road structures, and road user behaviours is launched. This paper makes a statistical analysis, and based on which, a conclusion on scene characteristics and road parameters such as the radius of road curvature, lane width, vehicle-following time gap, which could demonstrate the potential application of the dataset. To find out the behaviors of road users and drivers, we also focus on the relative distance, velocity, intention of other road users, which could deduce the native driving style for the automated vehicle to learn. We hope that data collected from real drivers would guide the research on automated driving.

Inspired by sustained progress in naturalistic driving benchmark [1], [2], [3], a preliminary establishment and analysis of automated driving research platform for specified scenarios in China have been carried out. Its primary goal is to provide a precise instruction on developing relevant algorithms for well-rounded research fields of AD including the perception, fusion, decision, and control parts. Specifically, the original image data from the monocular camera, detection output from the functional camera, and the 3D information of two full-

view LiDARs could be used for environmental perception [4] and sensor fusion[5]. Besides, the recorded different driving behaviors and the vehicle dynamics from CAN assisted by Inertial Measurement Unit (IMU) and Global Positioning System (GPS) would be useful for automated decision and control research [6], [7].

In the rest of this paper, the overview of approach architecture for AD benchmark and the vehicle platform are introduced in Section 2. The basis of scene design is explained, as well as the covered scenario illustration and route exhibition in Section 3. Next section introduces the details of data acquisition and labelling. The potential application analysis is discussed in Section 5. Finally, we sum up our work and propose the next research steps.

2 System architecture

2.1 Approach architecture

As shown in Figure 1, firstly we determine dataset category so that the sensor configuration can be confirmed. Then scene collection schema is designed to show Chinese characteristics. To cover these scenes, it is economical to select some representative cities, especially in different regions of China. Route design should consider the sizes of cities and hardware foundation. In labelling stage, it is necessary to make sure that we can cover all scenes' types; we also need to consider what and how to label to support AD research.

Finally, we can make analyses on the labelled data to demonstrate the potential application.

Figure 1. The architecture of AD benchmark establishment.

2.2 Acquisition platform

Our testbed is established based on a modified electric NISSAN Leaf. The sensor configuration for our AD benchmark in the test vehicle includes two LiDARs, GPS, IMU and three types of camera--monocular camera, stereo camera and functional Mobileye camera in Figure 2. The information of vehicle dynamics and driver's operation can also be acquired from vehicle's CAN. The basic rule of sensor configuration is to support full-view perception and comprehensive understanding of decision/control parts with effective cost.

Figure 2. Main sensors installed on our testbed.

Here list the details of each sensor:

- Mono-Camera: 1×IDS UI-5250CP-C-HQ, 1,920,000 pixels, 35.6fps, global shutter, CMOS.
- Stereo-Camera: 1×Stereolabs ZED stereo camera, max to 1080x720, field of view: 110°, up to 100fps.
- Mobileye-Camera: Mobileye 5-Series, the output of lane detection, velocity, position of target.
- LiDAR: 2×Velodyne VLP-16, 16 Channel, range: 100m, FOV: 360° (horizontal), ±15° (vertical), accuracy: ±3cm.

We achieve the camera calibration and LiDAR calibration using sensor drivers in Robot Operating System (ROS) [8] individually. The open source tool of joint calibration [9] is used, as presented in Figure 3. The camera and LiDAR are calibrated to the same coordinate system in the head of vehicle.

Figure 3. Joint calibration of camera and LiDAR sensor.

3 Scene design

Vast area and unbalanced developments in both infrastructure and economy bring China great challenges to promote AD. Scenes vary in all aspects: modern/traditional blocks; structured/unstructured road; main road users; truck and car on highway. Varied scenes then breed more behaviors on the road, mostly violating traffic rules, such as jaywalking, converse driving and illegal parking. Based on these features, this section elucidates design basis, describes covered scenarios, illustrates concrete route exhibition.

3.1 Design basis

A large scale of high-quality natural driving dataset is indispensable for AD development and testing. Therefore, it is necessary to design concise and representative scenarios and routes to build the dataset that meets the needs. We divide the currently general driving datasets into two categories: one is the universal dataset proposed by the field of pure computer vision, only including the element of VRUs like cyclists [14]; the other is AD dataset [10], [11], including IMU, GPS and other information which can be acquired from the driving simulator or the real-natural driving which is our concern.

We investigate the features of typical datasets under the real-natural driving scene finding that the dataset concerning the AD is mostly used in environmental perception and related fields in computer vision, ignoring the integration of drivers' behavior sets. Correspondingly, our proposal is to assist the development and testing of the overall AD process. Compared with current existing AD datasets (Table 1), our work covers more kinds of traffic scene (shown in next part), which varies with three Chinese cities, whereas previous papers focus more on climate variable [11] or limited type of scene [12], [13].

Table 1. Features of typical AD dataset.

Dataset	Features
KITTI[10]	In a mid-size city; In rural areas and on highways; real-world computer vision benchmarks
Oxford RobotCar[11]	One route; Over one year; In all weather conditions, including heavy rain, night, direct sunlight and snow
Cityscape[12]	Cover 50 cities; Several months (spring, summer, fall); Daytime; Good/Medium weather conditions; urban roads
Comma.ai[13]	7.25 hours of driving data of highway

3.2 Covered scenarios

Our dataset mainly contains ten natural scenarios (Figure 4) in three cities that are commonly used in the AD development, including *low-speed urban road*, *urban expressway*, *urban overpass*, *urban undercrossing tunnel*, *main road downtown*, *industrial park without signal light*, *original urban road*, *highway*, *campus* and *community*. They can reflect different traffic conditions in different types of cities in China under the same scenario and guide the adaptive development of AD research.

Figure 4. Normal scenarios.

Route in Beijing can be taken as a good example, which contains: Tsinghua University (*campus*), Wudaokou (*congested intersection, from low-speed urban road complex conditions*), Wanquanhe Bridge (*urban overpass*), Xiaojiahe Bridge (*urban expressway*), G7 (*highway*), Software Park (*industrial park without signal light*), Xi'erqi (*original urban road*), Zhongguancun (*commercial district, from main road downtown, large crossing*) and Nanlou (*community*).

Figure 5. Special scenarios.

a) T-shape crossing	b) Y-shape crossing
c) Right-turning lane split from straight lane	d) Complicated scene, easy to get crowded
e) illegal road occupancy	f) Vogue lane marks
g) spotted lane mark	h) Patched road surface
i) Road under construction with lane width changed	j) Old-fashioned block and market, one-way lane
k) Mixed traffic flow on narrow road	l) Narrow road outside, few people or vehicles

m) Main road with unique left-turning lane occupancy	n) Urban road contains roundabout
o) Interior roads in hospital with narrow one-way lanes, and many parked cars aside	p) Undercrossing tunnel downtown, white balance problem occurs when entering/exiting tunnel

Furthermore, there are sixteen special scenarios as typical representations in Figure 5, which can well reflect features of the road with Chinese characteristics. For example, no matter in the large city or small city, illegal road occupancies are very common. The T-shape or Y-shape crossing is unique leading to more complex traffic flow. Also, there are some roads under construction with lane width changed. These unconventional datasets with local characteristics of traffic scenes can better guide the AD technology to land in China.

3.3 Selected routes

We choose three relatively developed and representative cities located in the northern, central and southeast region of China -- Beijing, Luoyang and Suzhou (see Figure 6). As the capital of China, Beijing is the most representative first-tier city. Compared with the national average level, the management of pedestrians and non-motor vehicles is more standardized. Luoyang is a second-tier city located in central China, the reconstruction of the Old City and construction of the New District are underway, which produces the mixture of structured road and unstructured road scenes. Suzhou, located in the southeast, different from the two cities mentioned before. Many township roads, bridges, newly-developed commercial areas, high-speed entrances and exits could effectively supplement previous scenes.

Figure 6. Three representative cities in China.

c) Route of Beijing

Figure 7. Designed routes in Suzhou, Luoyang and Beijing.

Due to limited storage and cruising mileage, design principle of the route is to cover all scenarios around the city center, and as many local features of traffics as possible in 2~2.5 hours' journey covering 45~50 km, so that we can drive to collect twice a day each city, as depicted in Figure 7.

4 Data acquisition and labelling

4.1 Data acquisition and extraction

To achieve the data acquisition function of independent sensors, we use the top-down design idea to design the data-collecting program in the ROS platform, and all the procedures apply hierarchical modular design to improve the portability and scalability of the software in Figure 8.

Figure 8. Software platform in ROS nodes.

Because the collected sensor data is saved in a unified ROS bag format, we need to extract the image file, point cloud file, CAN information, IMU/GPS information and Mobileye-camera results from the generated bag files. As far as the timestamp synchronization is considered, the integrity of the data is guaranteed using combined shell and python script by loading the complete message information from the bag file, and the concrete sensor information of the real traffic environment is extracted, as shown in Table 2.

Table 2. Extracted sensor data.

Item	Format	Description
Mono-Camera	PNG	Provide the frontal environment information
Mobileye Camera	CSV	Provide the detection information
VLP16 LiDAR	PCL	Provide 360-degree perception
CAN	CSV	Provide vehicle dynamics information
GPS&IMU Data	CSV	Provide the position and vehicle posture
Scene Type	CSV	Provide different scene types

4.2 Labelling contents and rules

4.2.1 Labelling content design

For subsequent AD research regarding to the environmental perception such as multiple object detection, scene segmentation [15], tracking [16], even intention prediction [17] and risk assessment [18], we design to label all the road users, all the traffic signs and lights, as well as any road markers and road users' abnormal behaviors. In addition to common park-highway-city AD testing scene roadmap from closed to open road, we also consider cluster campus scene as a semi-closed road. Moreover, considering varied movement model, air resistance and illumination condition, the 2nd level scenes -- uphill, downhill, on or under the bridge, and tunnel scene are designed. The labelling contents are summarized as follows:

Table 3. Labelling content design of scenes and objects.

Scenes	1 st level	Campus	Park	City	Highway
	2 nd level	Uphill	Downhill	On or Under Bridge	Tunnel
Objects	Road user		Traffic sign	Traffic light	Road marker

To understand the scene type, here shows some 1st scenes:

Figure 9. The example of the 1st scene types.

Table 4. Labelling content design of abnormal behaviors.

Abnormal behaviors	Description
abparking	Illegal parking in wrong place
abLanechange	Illegal lane change
abovertaking	Illegal overtaking
jaywalk	Jaywalk
abDrive	Driving between two lanes
accident	Exist accident
converse	Driving in the oppsite direction
disobey	Drive through red light

Non-motor examples of road user types namely Vulnerable Road Users (VRUs) are shown:

Figure 10. The non-motor example of road users.

4.2.2 Labelling rules

In order to better refine the workflow of the labelling, there are some specific regulations listed below, according to the requirements of future AD research:

For scene labelling, each image could have one and only one 1st level scene type, and one or more 2nd level scene labels.

For object labelling, line segments are used for labelling lane and the boundary of the road, while bounding boxes are used for road users, traffic signs and traffic lights. The lane marking should be marked from "near" to "far", and other lanes should be marked in the same order. Lane markers less than two pixels or of which starting points and endpoints are nearly connected should be ignored. With regard to the traffic signs, indicate signs and guide signs are distinguished, so is the background color of them.

Besides, in order to cover special situations in China, we also categorize abnormal behaviors [22] of motor and non-motor vehicles labelled with bounding boxes. These behaviors involve abnormal parking, abnormal lane changing, abnormal overtaking especially for the motorcycle, and jaywalking or crossing the road for non-motor vehicles. Also, some common abnormal behaviors may exist for both of them, such as abnormal driving, converse, disobey and accident.

5 Dataset application analysis

In this section, some potential applications of our dataset will be discussed. There are various types of data in our dataset which we hope can support the automated driving research in the future.

5.1 Road condition analysis in China

5.1.1 Radius of the road curvature

Road curvature can be estimated by the lane or the boundary of the road in the image, which labelled as several segments defined as $[(x1, y1), (x2, y2)]$ in the image coordinate system [19]. All segments are classified based on whether they belong to the same lane. After

warping the image to the birds-eye-view image, it is necessary to interpolate some points between $(x1, y1)$ and $(x2, y2)$ to avoid under-fitting in the second order polynomial fitting. Finally, road curvature radius can be calculated by fitted formula.

Figure 11. The distribution of the radius in three cities' scenes.

Three different scenes are selected to show road radius. Comparing the distribution of radius, Suzhou Yupan road is the straightest followed by Luoyang Xindian expressway and Beijing industrial park, as shown in Figure 11.

This example proposes a simple method to fit road lane and calculate road curvature, which shows that our dataset can support the research on road environment perception for AD, such as the research of lane detection model.

5.1.2 Lane width statistics

In the collected images, we label a straight line in the lane from the left to the right. The width of lane W can be estimated by $W=wd/f$, where w is the length of the labelled line, d is the distance from the camera to the position of labelled line and f is the focal length. The distance d can be estimated by $d=Hf/l$, where H is the height of the camera and l is the vertical distance between the labelled line and the vanishing point.

Figure 12. The average width of lane in three scenes.

Typically, the width of lane in China is between 3.25m and 3.75m. But it may have some changes in different cities, different scenes or different types of roads. In our statistics, the lane width of expressway is the largest, as given in Figure 12.

Lane width is an important factor, as varied as the scenes, in the lateral control as the lane helps drivers operate vehicles closer to the center of the lane [20]. This example implies the collected data can support the research on lateral control and decision for AD.

5.2 Vehicle – following distance

Vehicle-following distance is an important factor to evaluate the traffic and the driving manner. Statistics is focused on the time gap [21] calculated by the velocity of self-vehicle and the distance from the front labelled vehicle, which can be obtained from CAN data and estimated by the method in Section 5.1.2 separately.

Figure 13. The time gap curves in two scenes.

Two Beijing scenes are selected to show vehicle-following characteristic. In Wudaokou, our vehicle followed the front vehicle until reaching the intersection. In the industrial park, the front vehicle is far away from our vehicle at the beginning. The driver tried to catch up with it and keep a safe following distance. The change of the time gap is shown in Figure 13.

By comparison, although these two scenes have similar average speed, the following time gaps are different. In the industrial park, the time gap is nearly two times of that in the Wudaokou. This result can reveal the strategies for vehicle following would vary in varied scenes. For vehicle-following research, we should consider not only the related distance and velocity, but also the scene types.

5.3 Road user behavior analysis in China

5.3.1 The actions of violation

Not only the objects are labelled in the dataset, but also some actions of violation are listed in Table 4. We select some scenes and count the violation in Figure 14.

Figure 14. Statistics on abnormal traffic acts.

When conducting research on AD, some violations are worth considering. It can help researchers better understand the real situation in China. In the statistics, illegal parking is the most common violation in China followed by driving in other lanes, jaywalking and

converse driving, especially on the road without digital surveillance (e.g. Changchun Market).

5.3.2 Driving manner in expressway

The dataset can also be used to analyze driving manner of different drivers. We selected Suzhou Su-Shao expressway to analyze driving manner of three drivers:

Table 5. Information of three drivers.

Driver	Gender	Miles/Year	Self-estimated driving manner
Driver A	Female	10000km	Passionate
Driver B	Male	8000km	Moderate
Driver C	Male	300km	Moderate

Driving velocity is used to demonstrate the driver manner, which can be extracted from CAN data. Figure 15 shows the change of velocity for each driver -- Driver A keeps velocity under 90km/h all the way, showing more consistence although her self-estimated driving manner is passionate; Whereas, driver B seems to be more passionate as his driving velocity keeps around 110km/h for a long time; Compared with driver B, driver C shows different driving manner, as his velocity was relatively smaller-in-amplitude but fluctuated frequently. This may be due to lack of driving experience as he only drives 300km per year on average.

Figure 15. The velocity curves of three drivers.

Learning from collected data for decision-making is very effective, such as using deep learning or some anthropomorphic algorithms. However, driver characteristics are complex factors and how to model them is also a challenge [23]. This example to some extent reflects that not all drivers' data would be appropriate in that some may be not representative. Therefore, it is necessary to choose excellent drivers for data collection to provide typical driving data.

6 Conclusion and future work

In this paper, the whole process of the establishment of an automated driving benchmark in China is introduced, ranging from layout of sensors, scene design to labelling objects selection. Furthermore, from specific scenarios of 7,000 labelled frames, both road features and naturalistic driving manners are analyzed. Not only road parameters could be extended to other potential automated driving scene information, but also naturalistic driving manners might be utilized as a learning target for automated

vehicle. With all the methodologies and the datasets yielded, researches on automated driving can acquire more supports.

Concretely, based on the work above, at least following researches can be carried on -- moving target detection, e.g. car, pedestrian and other VRUs; road marking recognition, like lane markers, traffic lights and signs; road segmentation based on boundary and lane line labelling; depth information prediction with point cloud data or multi-sensor fusion; identification and study of drivers' intention based on CAN data from different drivers' data. Moreover, the underlying research of risk assessment, automated vehicle decision and control will be in progress.

7 Acknowledgment

This research is supported by the project NO.20173000179 from Tsinghua and NISSAN Joint Research Center for Intelligent Mobility. The authors would like to thank all partners within for their cooperation and valuable contribution.

References

1. A. Geiger, P. Lenz, C. Stiller, R. Urtasun. Vision meets robotics: The KITTI dataset. *International Journal of Robotics Research*, **32(11)**, 1231-1237 (2013)
2. J. Ziegler, P. Bender, M. Schreiber, et al. Making Bertha Drive — An Autonomous Journey on a Historic Route. *IEEE Intell. Transport. Syst. Mag.*, **6(2)**, 8-20 (2014)
3. A. Rangesh, K. Yuen, R. K. Satzoda, et al. A Multimodal, Full-Surround Vehicular Testbed for Naturalistic Studies and Benchmarking: Design, Calibration and Deployment. *arXiv preprint arXiv: 1709.07502*. (2017)
4. B. Romera-Paredes, P. H. S. Torr. Recurrent Instance Segmentation. *European Conference on Computer Vision*, Springer, Cham, 312-329 (2016)
5. C. Stiller, F. P. León, M. Kruse. Information fusion for automotive applications – An overview. *Information Fusion*, **12(4)**, 244-252 (2011)
6. J. Suh, H. Chae, K. Yi. Stochastic Model Predictive Control for Lane Change Decision of Automated Driving Vehicles. *IEEE Transactions on Vehicular Technology*, **67(6)**, 4711 (2018)
7. D. Silver, J. A. Bagnell, A. Stentz, Learning Autonomous Driving Styles and Maneuvers from Expert Demonstration. *Experimental Robotics*, Springer, Heidelberg, 371-386 (2013)
8. M. Quigley, K. Conley, B. Gerkey, et al. ROS: An open-source Robot Operating System. *ICRA workshop on open source software*, **3(3.2)**, 5 (2009)
9. M. Vefas, M. Španěl, Z. Materna, et al. Calibration of RGB Camera with Velodyne LiDAR. In *Proc. Conference on Computer Graphics, Visualization and Computer Vision* (2014)
10. A. Geiger, P. Lenz, R. Urtasun. Are we ready for autonomous driving? The KITTI vision benchmark suite. *IEEE Conference on CVPR*, 3354-3361 (2012)
11. W. Maddern, G. Pascoe, C. Linegar, et al. 1 year, 1000 km: The Oxford RobotCar dataset. *The International Journal of Robotics Research*, **36(1)**, 3-15 (2017)
12. M. Cordts, M. Omran, S. Ramos, et al. The Cityscapes Dataset for Semantic Urban Scene Understanding. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 3213-3223 (2016)
13. E. Santana, G. Hotz. Learning a Driving Simulator. *arXiv preprint arXiv: 1608.01230* (2016).
14. X. Li, F. Flohr, Y. Yang, et al. 2016. A new benchmark for vision-based cyclist detection. In *Proc. Intelligent Vehicles Symposium*, 1028-1033 (2016)
15. V. Badrinarayanan, A. Kendall, R. Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. on PAMI*, **39(12)**, 2481-2495 (2017)
16. Y. Chen, X. Yang, B. Zhong, et al. CNNTracker: Online Discriminative Object Tracking via Deep Convolutional Neural Network. *Applied Soft Computing*, **38**, 1088-1098 (2016)
17. R. Quintero, I. Parra, D. F. Llorca, et al. Pedestrian Intention and Pose Prediction through Dynamical Models and Behaviour Classification. In *Intelligent Transportation Systems (ITSC), IEEE 18th International Conference on*, 83-88 (2015)
18. D. Dominic, S. Chhawri, R. M. Eustice, D. Ma. Risk Assessment for Cooperative Automated Driving. In *Proc. 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy ACM*, 47-58 (2016)
19. R. Song, H. Chen, X. Xiao, et al. Lane Detection Algorithm Based on Geometric Moment Sampling. *Scientia Sinica*, **47(4)**, 455-467 (2017)
20. S. Mecheri, F. Rosey, R. Lobjois. The effects of lane width, shoulder width, and road cross-sectional reallocation on drivers' behavioral adaptations. *Accident Analysis & Prevention*, **104**, 65-73 (2017)
21. A. Tordeux, S. Lassarre, M. Roussignol. An adaptive time gap car-following model. *Transportation Research Part B Methodological*, **44(8)**, 1115-1131 (2010)
22. D. Gaither, H. Do, B. R. Bryant. Toward Detection of Abnormal Behaviors in Timing and Security Requirements. In *Proc. 24th Asia-Pacific Software Engineering Conference Society (APSEC)*, IEEE, 707-712 (2017)
23. K. C. Dey, L. Yan, X. Wang, et al. 2016. A Review of Communication, Driver Characteristics, and Controls Aspects of Cooperative Adaptive Cruise Control (CACC). *IEEE Intell. Transport. Syst. Mag.*, **17(2)**, 491-509 (2016)