2020-01-5219 Published 30 Dec 2020

Design and Implementation of Test Scenario Library Data System for Autonomous Vehicles

Chen Cheng TMRI, China

Citation: Cheng, C., "Design and Implementation of Test Scenario Library Data System for Autonomous Vehicles," SAE Technical Paper 2020-01-5219, 2020, doi:10.4271/2020-01-5219.

Abstract

In recent years, with the development of autonomous driving technology, the unsafe, unreliable and unstable of autonomous driving systems have gradually been exposed. It is necessary to build test scenarios through a large amount of data in the real driving scene database to conduct comprehensive and strict tests and evaluation. There is a lack of an autonomous vehicle test scenario library data system that provides unified standard data. This paper designs and implements a test scenario library data system for autonomous vehicles. It explains the build steps of data system and analyses the methods of the implementation. It explains the build steps of data system and analyses the methods of the implementation. The data system contains real driving scene management module and test scenario management module. The system is based on B/S structure and decomposed into front-end and back-end. It is designed and implemented hierarchically that is composed application layer, data layer and support layer. The real driving scene management module is responsible for data collection, it could validate the data format, and then process, combine and store submitted data. The test scenario management module integrates the functions such as data filtering, data generation and scenario rebuilding. It extracts scenario element action libraries and provides algorithm-based test datasets. At present, the basic function of the data system has been developed. The test scenarios can be generated and export by using the system directly. It solves the problems in data processing and maintenance of massive autopilot data, and can give out standard data to construct virtual simulation cases. It would make the data support for building a suitable test environment of autonomous vehicles.

Keywords

Autonomous vehicles, Testing and evaluation, Test scenarios, Database

Introduction

With the development of emerging technologies such as big data, cloud computing, artificial intelligence and 5G, autonomous vehicle technology has entered in a stage of rapid development. It becomes as an important trend in the future development of the automotive industry [1]. Countries around the world have carried out research on autonomous driving R&D, testing, management and application. At the same time, they are also considering formulating corresponding standards or regulations for autonomous vehicles.

However, in recent years, there have been many road traffic accidents causing casualties in autonomous vehicles all around the world. It shows that there are still many instabilities in the autonomous vehicle system. The system needs to conduct massive tests in a wider range of traffic environments. It should be fully considered the road environment and the various factors of traffic participants and should be verified the relevant functions. The autonomous vehicles test and evaluation system is based on certain specific conditions and indicators. However, the test scenarios are relatively limited. For complex scenarios such as autonomous driving and manual driving vehicles mixed scenario, extreme scenarios such as the simulate road traffic accidents scenario, it is unable to cover the test evaluate of autonomous vehicles on the situation of completing perception and decision-making tasks. Therefore, it is necessary to establish a test scenario library data system for autonomous vehicles that conforms to the actual traffic conditions in China. It would promote the scientific, effective, and standardized testing of autonomous vehicles.

This paper mainly explains the design and implementation of the test scenario library data system for autonomous vehicles. In the first part, it explains the process of design. It analyses the requirements, the working process and the architecture of the system and introduces the functions of each module of the test scenario library data system for autonomous vehicles. In the second part, it explains the system implementation methods including database design, deconstruction and rebuild technology of scenario data and construction of test scenarios. In the third part, it shows the research results. The test scenario library data system for autonomous vehicles is now being used. It predicts the response ability of autonomous vehicles in the face of typical scenarios, and provides a data foundation and system support for the standardized development of autonomous vehicle operational safety tests.

1 Related Work

The test scenarios are the important prerequisite for the testing and evaluation of autonomous vehicles [2]. In 2016, the PEGASUS project initiated in Germany. They planned to build a scenario library for system development and test verification by the end of 2019 [3]. In this project, they divided test scenarios into three levels: functional scenarios, logical scenarios and concrete scenarios [4].

In order to get more key scenarios, Zhao and Huang et al. [5, 6] used accelerated testing methods to speed up the evaluation process by finding the key influencing factors of test scenarios and applied importance theory. The acceleration rate is estimated by comparing the collision rate of a human-driven vehicle under acceleration conditions and natural conditions. However, this method only considers some limited influencing factors of specific driving assistance functions, which limits the scope of application.

Ma [7], Ungoren [8] and Kou [9] proposed extreme scenarios test method (WCSE) based on the database established by traffic accidents to create the most challenging test scenario based on vehicle dynamics and control analysis. Although the extreme scenario assessment method can identify the weaknesses of the vehicle and the vehicle control system, it does not consider the possibility of the worst-case. Andreas validated the emergency lane assist system based on these most dangerous scenarios [10] .

Li Huang [11] et al. proposed a method of scene generation and screening for L2 level autonomous vehicles. According to the relative position and relative movement direction between vehicles and surroundings, they extracted the important parameters and rearranged them. Through the analysis of the three principles of authenticity, the influence of the interference between vehicles and the operability, they put forward the concept of the importance of working conditions. They selected scenarios which retain normal driving conditions and pre-collision conditions with test value to form a scenario data-set.

In summary, there are some scene database or data-sets exist around the world, which have certain significance for the research of this paper.

2 Design of the System

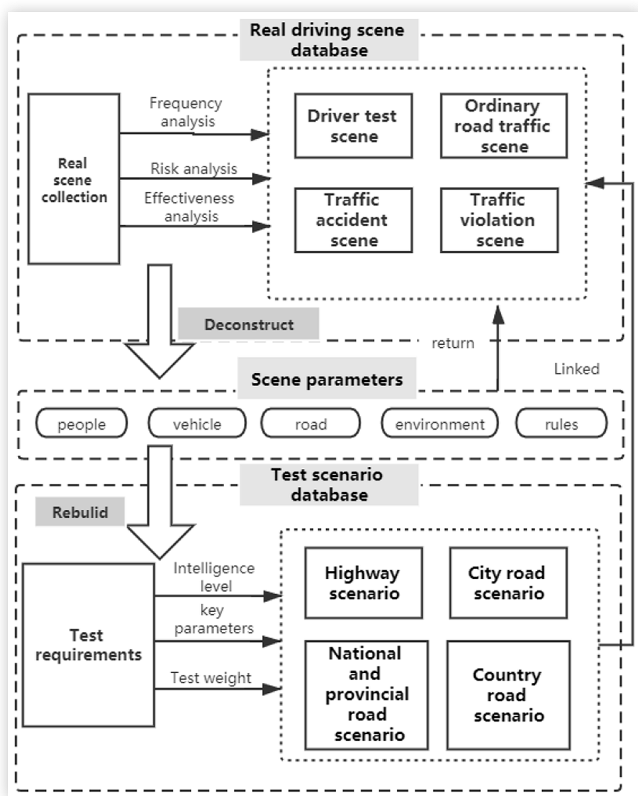
2.1 System Requirements Analysis

The focus of the establishment of the test scenario library data system for autonomous vehicles relies on the data processing procedures and methods.

First of all, we concluded the standards for road traffic management and road construction, and sort out basic scene data based on various structured and unstructured thematic data resources collected in the real driving scene database. Then, we performed format conversion, coordinate conversion, and spatialization of the subject data, and deconstructed the abstract multi-dimensional data into a unified format of two-dimensional data. Finally, according to the requirements of the level and test weight of the autonomous vehicles, the deconstructed real driving data is recombined according to the rebuild algorithm to form a test scenario database which is suitable for autonomous vehicles testing. The requirements analysis process of the test scenario database for autonomous vehicles is shown in Figure 1.

The whole system contains real driving scene database and test scenario database. In the real driving scene database, it composes the data of driver test content and methods, ordinary road traffic scene, traffic accident scene and traffic violation scene which are collected by front-end hardware equipment. Through the analysis of the occurrence frequency, risk factor and effectiveness of the scenarios, the typical real scenes that can be correlated which related to the autonomous vehicles test scenarios are filtered and refined. In the test scenario database, a test scenario contains five categories of people, vehicles, roads, environment and rules. The data which deconstructed based on the real driving scene database. Through analysis of the test requirements of the autonomous vehicle, considering the intelligence level, key parameters and test weight of autonomous vehicle, the four types of the test

FIGURE 1 The requirements analysis process of the test scenario database for autonomous vehicles



scenarios are generated, which including highway scenario, city road scenario, national and provincial road scenario, and country road scenario. The test scenarios of autonomous vehicles generated by the system have a certain correlation with the real driving scenes to ensure that the generated test scenarios are well-founded and reasonable.

2.2 Overall System Architecture

The system architecture of the test scenario library data system for autonomous vehicles is shown in Figure 2.

It is mainly including application layer, data layer and support layer. Among them, the application layer is composed of real driving scene management module and test scenario management module. The data layer includes subject data and basic data. The support layer provides technical support such as data interaction, data storage, and data analysis.

1. The application layer

The application layer is divided into two parts: real driving scene management module and test scenario management module. Relying on the technical foundation of the support layer, it realizes the functions of collecting, analyzing, and deconstructing real driving scene data, and rebuilding, generating, and instantiating test scenarios [12].

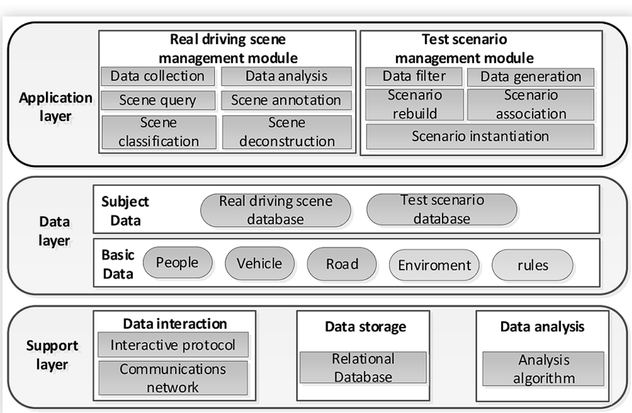
2. The data layer

The data layer is divided into two parts: subject data and basic data. The real driving scene library and the test scenario library constitute the subject data. Basic information such as people, vehicles, roads, environment and rules constitute basic data. Subject data and basic data provide data services for various systems in the application layer.

3. The support layer

The support layer contains the software and hardware environment of the entire autonomous vehicle test scenario library management system. It has the functions of data interaction, data storage and data analysis. The data interaction includes the deployment of communication networks and the formulation of interaction protocols to realize the interaction between various parts of the system and system users. Data storage mainly relies on relational databases (RDBMS) [13] to realize the storage of structured and unstructured data [14]. Data analysis includes the analysis of structured data using algorithms, as well as, using distributed computing to analyze unstructured data such as pictures.

FIGURE 2 The system architecture of the test scenario library data system for autonomous vehicles



2.3 System Function Modules

The test scenario library data system for autonomous vehicles is displayed in the form of front-end software, which is divided into functions such as data collection, scenario query, data analysis and statistics, scenario generation.

1. Data collection

The sources of real driving scenario data acquisition include:

- Based on driver test, including test content, test methods, road traffic rules, safe driving and civilized driving content.
- Based on road traffic scenes, it is mainly for collection of dynamic scenarios. The scenario data is divided by actions and stored in the database.
- Based on traffic accidents and illegal situations. Through the collection and analysis of traffic accident files, we analyse the causes of accidents in depth and extract typical illegal behaviors.

The data can be insert into database by forms or templates which are set before by system. When entering a scenario, the scenario type should be marked.

2. Scenario Query

Scenario query is applicable to both real driving scene and test scenarios, including retrieval of information such as scenario type and scope of application, and supports single query, combined query, and schema query of scenario data.

3. Data analysis and statistics

The system can count multiple indexes of various scenario data, including scenario type, risk factor, frequency factor, effectiveness, etc., to provide an effective data reference basis for test scenario generation. The system can provide the function of the data analysis. It mainly includes scenario cluster analysis, scenario comparison analysis, etc. It is displayed in the form of visual analysis charts, which can provide data support for the rebuild of test scenarios.

4. Scenario generation

The generation of test scenarios is divided into automatic and manual operations. The system will automatically generate a list of test scenarios for users to view and filter by setting the scope of application of the scenario test requirements [15].

3 System Implementation Methods

3.1 Real Driving Scene Analysis and Filtering

In the process of autonomous vehicles testing, we focus on verifying whether the vehicle can perceive changes in the environment. It is necessary to verify the reliability and safety of the autopilot system functions in various environments. When constructing the test scene, the content of the real scenes will be the main data source. The real scenes that we collected are diverse, so we need to analyse and filter the effective scene data according to the testing requirements of autonomous vehicles. The data selection follows the following evaluation criteria.

1. Based on Driving Behavior

We defined the normal driving behaviors that frequently occur in vehicles as typical driving behaviors, including following lanes, following cars, changing lanes, overtaking and changing lanes, cutting in, cutting out, turning left, turning right, reversing, and parking etc.

For accident scenes, if the accident is caused by factors such as exceeding the vehicle limit, the driver's violation of the law (e.g. drunk driving, fatigue driving, running the red light), or the vehicle function failure, it will be excluded from the data source of the scenario construction. The accident scenes that we concerned are caused by "human factors" such as the driver's unawareness, the vehicle's failure to avoid collisions, or the driver's operational errors. So we defined them as typical scenes.

2. Based on Road Structure

Autonomous vehicles should be able to perceive different road structures and perform corresponding actions. The typical real scenes should cover straight roads, bends, tunnels, crossroads, intersections, bridges, roundabouts, parking lots and other representative special scenes.

3. Based on Traffic Facilities

In the composition of scenario data, roadside traffic facilities and signs are the indispensable parts. To test whether autonomous vehicles can drive normally on the road, compliance with traffic rules is an important evaluation. The traffic rules of each country are different. We need to establish a test scenario library suitable for the road traffic environment in China. So the scenes covering traffic facilities were defined as typical scenes.

4. Based on Special weather

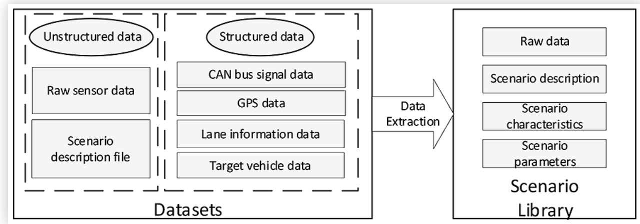
When analyzing real scene data, we found that severe weather is one of the important factors that induce traffic accidents, especially for the highway scenes. It is also necessary to consider typical weather factor, especially for rain, snow, fog, and cross wind.

3.2 Database Design

1. Database content

The data in the data layer mainly includes structured data and unstructured data. In the database, unstructured data mainly refers to raw sensor data such as pictures, videos, point cloud images (PCD), and scenario description files. Structured data mainly refers to the vehicle CAN bus signal, absolute positioning, lane information and target vehicle data obtained after the analysis of autonomous vehicles systems. Through the collected data, a structured and unstructured data-sets are formed. The specific database content is shown in Figure 3.

FIGURE 3 Database content



By completing the steps of multi-source data consistency verification, vacancy value verification, and unification of standards and measurement units, we complete data extraction tasks such as data matching, data cleaning, and data unification to enhance the data dimension and data accuracy information to improve data quality, and form a highly reliable scenario library. In the scenario library, it contains raw data, scenario description, scenario characteristics and scenario parameters.

2. Data structure

How to store the collected data? How to export the data when generating the scenarios? The structure of data storage plays a decisive role. From the analysis of the collected real scenario data, it is concluded that the traffic scenario on the actual road is formed by a combination of factors such as people, vehicles, roads, environment, and traffic management rules. Therefore, the design of the data structure is also based on these elements, which are specifically divided into static elements and dynamic elements. Static elements include roads, environment, management information. When generating a scenario, a static model can be built in advance. Dynamic elements include people and vehicle information. When generating a scenario, the action behavior of participants and vehicles can be set according to the test requirements, and added to the static model, thereby we can form a test scenario that meets the actual road test requirements.

The data structure is shown in Figure 4 and Figure 5.

Data items such as pedestrian, vehicles, roads and environment are shown in Table 1 to Table 4 [16]. Due to the large amount of data content, only part of the content is displayed

FIGURE 4 Static elements data structure

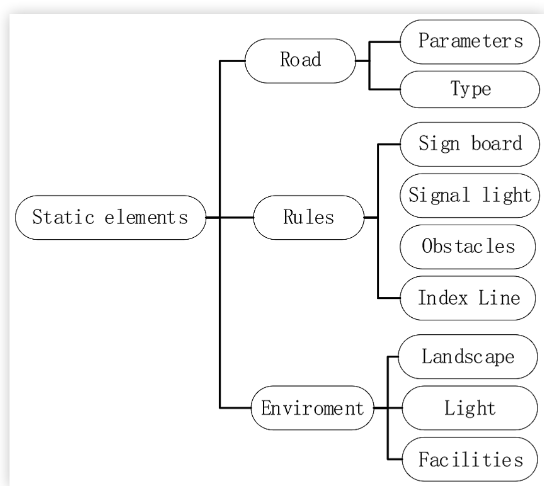


FIGURE 5 Dynamic elements data structure

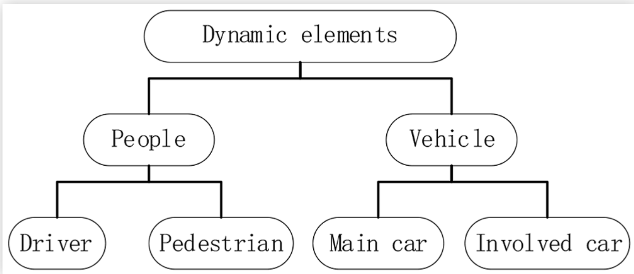


TABLE 1 Pedestrian Information Table

Pedestrian Information Table	
Table Name	**AEV_SCENE_PEDESTRIAN**
Fields	Data Content
TYPE	1-adult, 2-child, 3-disabled
SPEED	signed number, unit: km/h, relative to the speed of the collection point
SPEED_ RANGE	1-stationary, 2-slow travel (less than 5km/h), 3-normal (about 5km/h), 4-fast travel (5~10km/h), 5-run (greater than 10km/h)
DIRECTION	1-forward, 2-left, 3-right, 4-back [17]
STATE	1-normal, 2-crossing crosswalk, 3-crossing road, 4-overpassing isolation facilities, 5-walking in the motor vehicle lane, 6-playing on the road, 7-working on the road, 8-staying in the motor vehicle lane [17]

TABLE 2 Vehicle Characteristic Parameter Table

Vehicle Characteristic Parameter Table	
Table Name	**AEV_SCENE_MOTOR**
Fields	Data Content
TYPE	1-Large passenger cars, 2-Medium passenger cars, 3-Small passenger cars, 4-Mini passenger cars, 5-Heavy trucks, 6-Medium trucks, 7-Light trucks, 8 -Mini trucks, 9-three-wheeled vehicles, 10-low-speed trucks, 11-special operation vehicles, 12-tramcars, 13-motorcycles, 14-trailers [18]
STATE	1-Normal, 2-Fault
COLOR	1-white, 2-gray, 3-yellow, 4-pink, 5-purple, 6-green, 7-blue, 8-red, 9-brown, 10-black [18]
SPEED	signed number, unit:km/h
DIRECTION	1-offset to the left, 2- follow the line, 3- offset to the right
ACCELERATION	signed number, unit:m/s^2
DRIVE_STATE	1-go straight, 2-turn left, 3-turn right, 4-overtake, 5-change lane, 6-start, 7-stop, 8- U-turn, 9-avoid obstacle, 10-other
TURN_LIGHT_ STATE	1-Left turn light is on, 2- Right turn light is on, 3-double jump, 4-off
LIGHT_STATE	1-off, 2- high beam, 3- low beam
BRAKE_LIGHT_ STATE	1-on, 2-off

TABLE 3 Road Information Table

Road Information Table	
Table Name	**AEV_SCENE_ROAD**
Fields	Data Content
RADIUS	signed number, unit:m
DECLIVITY	signed number, unit:degree
HIGHWAY_TYPE	1-high speed, 2-level I, 3-level II, 4-level III, 5-level IV, 6-level V [19]
HIGHWAY_LEVEL	1-National Road, 2-Provincial Road, 3-County Road, 4-Country Road, 5-Others [19]
CITY_ROAD_TYPE	1-express road, 2-general urban road, 3-unit community self-built road, 4-public parking lot road, 5-public square road, 6-other urban roads [19]
IS_CROSS	1-round intersection, 2-ramp, entrance and exit, 3-section (intersections with only two directions), 4-T-shaped intersection, 5-Y-shaped intersection, 6-staggered T-shaped intersection, 7-staggered Y-shaped intersection, 8-cross intersection, 9-diagonal intersection, 10-multiple intersection, 11-others [19]

TABLE 4 Environment Information Table

Environment Information Table	
Table Name	**AEV_SCENE_ENV**
Fields	Data Content
TEMPERATURE	signed number, unit:celsius
VISIBILITY	0-50M, 1-100M, 2-200M
LIGHT	1-Sunny day (above 30000lux), 2-cloudy day (10000lux-30000lux), 3-rainy day (3000lux-10000lux), 4-will black (0.5lux-3000lux), 5-dark night (0.001lux-0.02lux), 6 -Moon night (0.02lux-0.1lux), 7-star light (less than 0.001lux), 8-street lighting (0.1lux-0.5lux)
RAINFALL	Level 1-8
WIND	Level 1-16
NOISE	1-levelI(30~59dB), 2-level II(60~89dB), 3-level III(90~120dB)
WEATHER	1-clear, 2-cloudy, 3-smog, 4-rain, 5-snow, 6-sand dust, 7-hail, 8-smoke, 9-gale
traffic_flow	1-fast, 2-medium, 3-slow

in the database table. The definition of data content refers to relevant road traffic standards in China.

3.3 Deconstruction and Rebuild Technology of Scenario Data

Since the number of scenes in real driving is unlimited, only limited information can be obtained from data sources such as driver test scenarios, natural driving scenarios, traffic accident scenarios, and traffic violation scenarios. Therefore, the way to solve this problem is to deconstruct the acquired data. For existing scenarios, we carried out the scenario division and element extraction to analyze the influence of different dimensional elements in different scenarios, including single influence and coupling influence between different factors.

On this basis, we parameterized and reorganized the scenario elements to supplement a large number of test scenarios with unknown working conditions. They could cover the blind spots of automatic driving function testing effectively. The composition is not a purely random combination. It has to consider the laws and constraint relationship between the elements. For example, the restrictions on road monitoring regulations and restrictions on traffic laws.

1. Scenario feature parameter extraction

The extraction process of scenario feature parameters mainly used mathematical statistical algorithms such as principal component analysis and cluster analysis. The method of principal component analysis uses the idea of dimensionality reduction. Through the study of the internal structure of the indicator system, multiple indicators are transformed into independent comprehensive indicators. The new comprehensive variables can reflect the main information provided by the original variables, thereby simplifying the data structure and Find the relationship between variables [20].

When rebuilding an autonomous vehicle test scenario, there are various factors that affect the applicability of the scenario. It is not enough to evaluate its quality based on a single factor. Therefore, the principal component analysis method is used to evaluate the impact factors of the test scenario in order to obtain effective scenes. We did the analysis on the danger coefficient, frequency coefficient, occurrence area, participating objects, vehicle behavior and other indicators of collected real driving scenes and extracted important parameters. Then, we combined a large amount of random discrete real traffic scenario data and use dimensionality reduction and clustering methods to group data with the same attributes and generated a test scenario with certain typical characteristics. As shown in Figure 6, the scene parameters were extract from different traffic scenes. After cluster analysis, we divided them into three test scenarios. S1 is for traffic accident scene. S2 is for regular road traffic scenes and S3 is for driver test scenes. The test scenario comes from real traffic scenes. The amount of data in real traffic scenes is much larger than that in test scenarios.

Here is an example for cluster analysis. Based on collected 133 motor vehicle and non-motor vehicle accidents data, we extracted the vehicle's motion parameters and road environment information of accidents and performed statistical analysis. The variables with the same characteristics are put in one group to obtain the typical accident scenes in the mixed environment of motor vehicles and non-motor vehicles. The variables we concerned about are as follows.

We extracted scenes with a frequency of more than 10% for analysis and aggregated them into three typical scenes: a motor vehicle collided with a non-motor vehicle without decelerating at an intersection (22.6%), a non-motor vehicle illegally changed lanes and adjacent lanes of motor vehicles collided (17.3%), non-motor vehicles crossed the road at the brunches of the road section, but the motor vehicle in front had no time to avoid collisions with non-motor vehicles (12.8%).

FIGURE 6 Test scenario library extraction process

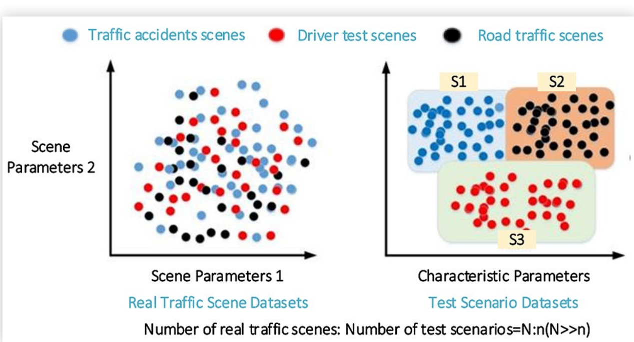


TABLE 5 Accident Data Statistical analysis Table

No	Type	parameter	value	Size	Frequency
1	Non-motor vehicle type	Two-wheel electric bicycle	0	105	78.9%
2		Three-wheel electric bicycle	1	17	12.8%
3		Bicycle	2	11	8.3%
4	Vehicle behavior	Follow the line	0	44	33.1%
5		Go straight at the intersection	1	21	25.6%
6		Turn right	2	44	15.8%
7		Turn left	3	34	9.0%
8		Merge	4	21	6.8%
9	Road characteristics of non-motor vehicle	Go straight at the intersection	0	37	27.8%
10		Follow the line	1	27	20.3%
11		Crossing the road	2	27	20.3%
12		Merge	3	20	15.0%
13		Crossing crosswalk	4	8	6.0%
14	Road characteristics	No traffic lights	0	32	24.1%
15		Traffic lights	1	31	23.3%
16		Straight lane	2	28	21.1%
17		Ramp	3	18	13.5%
18		Crosswalk	4	10	7.5%
19		Branch	5	6	4.5%
20	Driver's vision	Blocking	0	49	36.8%
21		Unblocking	1	84	63.2%
22	Traffic condition	Crowded	0	14	10.5%
23		Normal	1	107	80.5%
24		Uncrowded	2	12	9.0%
25	Weather	Sunny	0	112	84.2%
26		Cloudy	1	7	5.3%
27		Rain and snow	2	14	10.5%
28	Lighting	Normal	0	110	82.7%
29		Low	1	21	15.8%
32		Dark	2	1	0.8%
33	Vehicle type	Car	0	130	97.7%
34		Truck	1	3	2.3%

2. Rebuild technology based on schema matching

Madhavand et al. proposed a new method for schema matching, Cupid [21], which used element name, type and structure information to complete the matching. In the process of rebuilding, we used a similar method. The specific manifestation of the autonomous driving test scenario is a strong random coupling of traffic elements such as people, vehicles, roads, environment, and rules. Schema matching scenario rebuild refers that the system can provide one or more reasonable and executable test scenarios by changing the factors that affect the test results according to the given test requirements. That is, extracting the main components of people, vehicles, roads, environment, and rules from the original data and making sure that the original variable information are retained as much as possible but not related to each other. Usually, it makes a linear combination of the original indicator as a new generation scenario in the way of the mathematical treatment.

The schema matching logic rule can be expressed by the following mathematical expression Equation 1.

$$F_s = K_1\beta_{people} + K_2\beta_{vehicle} + K_3\beta_{road} + K_4\beta_{environment} + K_5\beta_{rules} \quad (1)$$

Where, K_i represents the influence weight of the element on the test scenario. The range value is 0-1. β is the normalized value of the original variable. It is divided into several types such as people, vehicle, road, environment and rules. Through the combination of K_i and β, we can get F. The higher the value of F, the more complex the scenario and the more difficult the test.

3.4 Construction of Test Scenarios

The test scenario is an interactive dynamic system composed of people, vehicles, roads, environment, rules and other dynamic and static traffic elements in a specific time and space.

It mainly has the following characteristics:

- Representative typical scenarios obtained through data analysis and processing of real traffic scenarios.
- It can highly summarize the main information of the traffic behavior. It is a high-level summary of the objects, processes, and results involved in traffic scenarios. The test scenario does not require specific descriptions of insignificant factors in the results produced by the traffic scenario.
- The test scenario is derived from the real traffic scenario and is more representative than the real traffic scenario.

1. Classification and construction of test scenarios according to different test purposes, test scenarios can be constructed into the following three type scenarios:

- Functional scenarios. Functional scenarios are operational scenarios described by semantics. It describes the entities in the domain and the relationships between entities through language scenario symbols.
- Logical scenario. The logic scenario presents the operation scenario in the state space. It expresses entity characteristics and relationships between entities by defining the parameter range of variables in the state space. The parameter range can be determined by the probability distribution.
- Specific scenarios. The specific scenario is represented by a certain parameter value in the state space and the relationship between the entities. Each logical scene can be converted into a specific scene by selecting a specific value from the parameter range [22].

2. Test scenario construction process

A complete test scenario is a summary description of the object, process, and result of an event, including at least the following five aspects of information. That is functional semantic scenario, test scenario reference diagram, typical analysis, logical scenario, scenario source description.

The first step is the construction of functional semantic scenarios. The complete semantic description scene contains at least there primitive scenes, that is the starting state of the scene, the situation when the scene occurs, the impact and possible results at the end of the scene. The starting state of the scene refers to the steady state. At the beginning of the scene, defining the road environment and the status of the main vehicle and the target vehicle. For example, if the main vehicle is driving on an urban express road, the lane speed limit information is 60km/h—80km/h. The situation of the scene is unstable, mainly used to describe the state of transition from one incident to another traffic incident. For example, while the main vehicle is driving strictly, a target vehicle in the adjacent lane ahead drives into the main vehicle lane. The impact and result at the end of the scene refer to the transition from one steady state to another steady state. When the transition state is unstable, traffic conflicts will occur. For example, the main vehicle collides with the target vehicle.

The second step is generation of test scenario reference graph. This is a supplement to the semantic scenario, which mainly describes entities in the domain and an interoperability relationship between entities in the form of simplified diagrams.

The third step is typical analysis. It is used to describe the probability of occurrence of the scene and risk assessment.

The fourth step is construct logical scenes. That is, it is used to define the spatial distribution and value range of the state parameters in the semantic scene.

The fifth step is the description of the scenario source. Generally, including vehicle driver test scenes, traffic accident scenes, traffic violation scenes, and typical road traffic scenes.

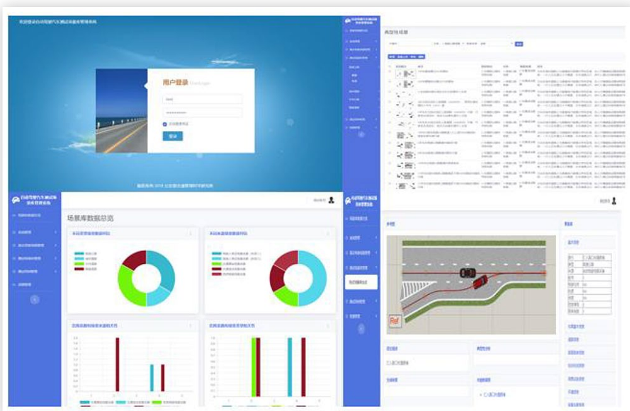
4 Implementation of Application Software

The automated driving test scenario database software system adopts a service-oriented SOA architecture. It provides the functions of the IT system to users or other services in the form of services to adapt to changing needs. The display and

TABLE 6 System implementation technology

Implementation part	Technology
Front-end interface	Bootstrap+Vue.js (an adaptive HTML5 front-end framework, can be compatible with most modern browsers based on Webkit core on desktops, IOS, and Android.)
Backend framework	Spring Boot + JPA
Operating environment	CentOS 7
Middleware	Tomcat 8, Nginx 1.14
Database	Oracle 10g

FIGURE 7 The software interface of the self-driving car test scene library



operation interface of each functional module of the system are presented in the form of Web, which can be browsed and operated normally on mainstream Web browsers. The implementation technology used as in Table 6.

The software system consists of two parts. The front-end has the function of uploading, downloading, querying, and deconstructing real driving scene data, and automatic rebuilding, generation, and display of test scenarios. The back-end is for database management, user management and other functions. The interface of the whole system is shown in Figure 7.

In the test scenario library data system for autonomous vehicles, it contains 22943 real driving scenes that covering the area of vehicle driver test scenes, traffic accident scenes, traffic violation scenes and typical road traffic scenes. The test scenario library is divided into two sections: the typical test scenario and generated test scenario. In the next step, we will create the corresponding test case for simulation and application.

Conclusion

Building a library of autonomous driving test scenarios covering typical traffic environment in China is conducive to accelerating the safety testing of autonomous vehicles. However, the scene on the road is uncertain, including dangerous driving, extreme driving conditions, special driving scenes, dense traffic flow, mixed traffic, complex terrain and road structures. How to construct effective test scenarios for actual road testing, closed field testing, or simulation testing, etc., still needs to be coordinated to improve and optimize the iterative automatic driving test scenario library.

In this paper, through limited mapping of complex driving environment, extract typical test scenarios based on the real driving scene database, and use computer technology to develop an automated driving test scenario database. It could do the configuration of scene variables flexible during testing and use limited scenarios to map test scenarios under multiple working conditions. In this way, it predicts the response ability of autonomous vehicles in the face of typical scenarios, and provides a data foundation and system support for the standardized development of autonomous vehicle operational safety tests.

References

1. Wachenfeld, W., Winner, H., Gerdes, J.C. et al., "Use Cases for Autonomous Driving," Presented at in *Autonomous Driving*, 2016, doi:10.1007/978-3-662-48847-8_2.s.
2. Najm, W.G., Smith, J.D., and Yanagisawa, M., "Pre-Crash Scenario Typology for Crash Avoidance Research," (DOT HS 810 767), National Highway Traffic Safety Administration (NHTSA), 2007.
3. Form, T., "General Introduction to PEGASUS & Opening of the Exhibition," in *PEGASUS Symposium*, 2017.
4. Menzel, T., Bagschik, G., and Maurer, M., "Scenarios for Development, Test and Validation of Automated Vehicles," in *IEEE Intelligent Vehicles Symposium*.
5. Zhao, D., Huang, X., Peng, H., Lam, H. et al., "Accelerated Evaluation of Automated Vehicles in Car-Following Maneuvers," *IEEE Transactions on Intelligent Transportation Systems* 19(3):733-744, 2018.
6. Huang, Z., Lam, H., LeBlanc, D.J., and Zhao, D., "Accelerated Evaluation of Automated Vehicles Using Piecewise Mixture Models," *IEEE Transactions on Intelligent Transportation Systems* 99:1-11, 2017.
7. Ma, W.-H. and Peng, H., "A Worst-Case Evaluation Method for Dynamic Systems," *J. Dyn. Syst. Meas. Control* 121(2):191, 1999.
8. Ungoren, A. and Peng, H., "An Adaptive Lateral Preview Driver Model," *Veh. Syst. Dyn.* 43(4):245-259, Apr. 2005.
9. Kou, Y., "Development and Evaluation of Integrated Chassis Control Systems," The University of Michigan, 2010.
10. Eidehall, A., "Tracking and Threat Assessment for Automotive Collision Avoidance," *Behavior*, 2007.
11. Li, H., Xia, Q., Xie, F., Xiu, H.-L. et al., "Study on the Test Scenarios of Level 2 Automated Vehicles," in *2018 IEEE Intelligent Vehicles Symposium (IV)*.
12. Musa, S., Shahzad, A., Aborujilah, A. et al., "Secure Security Model Implementation for Security Services and Related Attacks Base on End-To-End, Application Layer and Data Link Layer Security," in *Proceedings of the 7th International*

Conference on Ubiquitous Information Management and Communication. Jan. 2013, doi:10.1145/2448556.2448588.

13. Bakalash, R., Shaked, G., Caspi, J. et al., "Relational Database Management System Having Integrated Non-Relational Multi-Dimensional Data Store of Aggregated Data Elements," United States Patent Application 20090271384, 2009, doi:US8463736 B2.
14. Indeck, R.S. and Indeck, D.M., "Method and System for High Performance Integration, Processing and Searching of Structured and Unstructured Data Using Coprocessors," United States Patent Application 7660793, 2010, doi:US7660793 B2.
15. Bagschik, G., Menzel, T., and Maurer, M., "Ontology Based Scene Creation for the Development of Automated Vehicles," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018.
16. Breiman, L.I., Friedman, J.H., Olshen, R.A. et al., "Classification and Regression Trees (CART)," *Biometrics*, 1984.
17. Public Safety Industry Standard of the People's Republic of China (GA), "Information Codes for Road Traffic Management," GA/T 16-2017.
18. Public Safety Industry Standard of the People's Republic of China (GA), "Types of Motor Vehicle-Term and Definitions," GA 802-2014.
19. Industry Standards of the People's Republic of China (JTG), "Technical Standard of Highway Engineering," JTG B01-2014.
20. Keenan, D.F., Valverde, J., Gormley, R. et al., "Selecting Apple Cultivars for Use in Ready-to-Eat Desserts Based on Multivariate Analyses of Physico-Chemical Properties," *LWT-Food Science and Technology* 48(2):308-315, 2012.
21. Madhavan, J., Bernstein, P.A., and Rahm, E., "Generic Schema Matching with Cupid," *Proc 27th Int Conf on Very Large Data Bases* (San Francisco, CA: Morgan Kaufmann Publishers Inc., 2001).
22. Menzel, T., Bagschik, G., and Maurer, M., "Scenarios for Development, Test and Validation of Automated Vehicles," in *IEEE Intelligent Vehicles Symposium*, 2018, doi:10.1109/IVS.2018.8500406.

This paper is based upon a presentation at the *3rd International Forum on Connected Automated Vehicle Highway System through the China Highway & Transportation Society.*

ISSN 0148-7191