Annex D

IVISTA
China Intelligent Vehicle Index

Test and Evaluation Protocol for Intelligent Cabin Index

(Version 2026)

IVISTA

Published by China Automotive Engineering Research Institute Co., Ltd.

IVISTA

Table of Contents

Foreword II
1 Scope 1
2 Normative References 1
3 Terms and Definitions 1
4 Test Requirements 4
4.1 Requirements for test field and environment 4
4.2 Requirements for test equipment 4
4.3 Requirements for VUT 8
4.4 Requirements for data processing and recording 9
4.5 Test shooting requirements 9
5 Test Method 10
5.1 Overview 10
5.2 Test method for multi-modal interaction richness 11
5.3 Test method for terminal interconnection convenience 23
6 Evaluation Methods 26
6.1 Overview 26
6.2 Evaluation method for multi-modal interaction richness 29
6.3 Evaluation method for terminal interconnection convenience 35
6.4 Comprehensive rating 37
Appendix A (Normative) Rules for Test and Evaluation of Multi-modal Interaction Richness 39

IMSTA

Foreword

With the rapid development of intelligent cabins and their increasing penetration rates, they have become a key factor influencing consumers' car-buying decisions. Intelligent cabin experience encompasses users' multi-sensory perception, including touch and hearing, delivering a comprehensive scenario experience both inside and outside the vehicle. However, with the increasing diversity and complexity of functions, delivering a premium intelligent experience to users remains a significant challenge.

To comprehensively and systematically assess the user experience of intelligent cabins and promote the intelligent update of cabin functions, China Automotive Engineering Research Institute Co., Ltd(CAERI). has focused on two core evaluation indicators for top-level design: "multi-modal interaction richness" and "terminal interconnection convenience", so as to establish a systematic cabin intelligence evaluation standard covering core scenarios, from users' multi-dimensional perception to full-scenario experiences inside and outside the vehicle. The abstract "intelligent experience" is broken down into specific, measurable, and comparable dimensions, centering on the cabin as the users' perception and interaction center to form a closed-loop experience.

Specifically, the multi-modal interaction richness dimension focuses on three areas — intelligent sound field, AI interaction, and intelligent displays — assessing experience factors such as audio performance, independent sound zones, AI speech interaction, and screen display performance. Testing and evaluating the intelligent cabin experience through multi-modal interaction, such as speech interaction and touchscreen interaction, is highly significant for testing and evaluating the level of intelligence in cabin interaction experiences. The terminal interconnection convenience is tested and evaluated from two main dimensions: smartphone-vehicle connectivity and sentry mode, with a focus on smartphone screen projection, smartphone remote control, digital key, and sentry mode. The intelligent interconnection experience of the cabin is tested and evaluated across all scenarios, from inside to outside the vehicle.

Adhering to the principles of openness, compatibility, and objectivity, this protocol will undergo continuous iteration and upgrades in line with advancements in artificial intelligence, human-machine interaction (HMI), and communication technologies. It is designed to provide consumers with a clear and reliable vehicle purchasing reference and to offer vehicle manufacturers clear benchmarks for product optimization and guidelines for technological innovation. This protocol aims to guide R&D resources toward areas that truly enhance users' intelligent experience, support the creation of differentiated competitive edges, thus driving industrial chain upgrades and continuously injecting momentum into the high-quality development of the intelligent vehicle sector.

IVISTA

Test and Evaluation Protocol for Intelligent Cabin Index

1 Scope

This document specifies the test and evaluation methods of IVISTA China Intelligent Vehicle Index (Version 2026) - Intelligent Cabin Index.

2 Normative References

The following normative documents contain provisions which, through reference in this text, constitute indispensable provisions of this document. Among them, for dated references, only the version corresponding to that date applies to this document; for undated references, the latest version (including all amendments) applies to this document.

GB/T 21023-2007 General Specification for Chinese Speech Recognition System

GB/T 34083-2017 Specification of Programming Interface for Chinese Speech Recognition Internet Service

GB/T 36464.5-2018 Information Technology - Intelligent Speech Interaction System - Part 5: In-vehicle Terminal

GB/T 45314-2025 Road Vehicles - Performance Requirements and Test Methods for Hands Free Communication and Voice Interaction

T/CAIACN 012-2024 Vehicle Audio - Objective Test Methods for Sound Reinforcement System Performance of Passenger Car

3 Terms and Definitions

For the purposes of this document, the following terms and definitions apply.

3.1

start-up time

The time needed from the moment when the stylus pen touches the screen to the moment when the application interface is fully displayed in the case that an application is started up by tapping the touch-screen, as calculated in Formula (1):

$$T_s = t_f - t_s \quad \text{.....(1)}$$

T_s start-up time;

t_f the moment when the application start-up is completed;

t_s the moment when the stylus pen touches the screen.

3.2

screen fluency

fluency of the screen while sliding the touch-screen, with two parameters, i.e., maximum successive frame stuttering count and average frame rate, as evaluation indicators

If the refresh frequency of the IVI touch-screen is N, n times of lags occur during swiping, the frames of each lag is recorded as f1, f2, … , fn respectively, and the actual frames refreshed during the sliding is f, then the maximum successive frame stuttering count

IVISTA

(P) and the average frame rate (X) can be calculated in Formula (2) and Formula (3):

$$P = \max(f_1, f_2, \cdots, f_n) \quad (2)$$

$$\frac{X}{N} = \frac{f - (f_1 + f_2 + \cdots + f_n)}{f} \quad (3)$$

3.3

maximum sound pressure level

sound pressure level that can be achieved by the audio reinforcement system playing the test signal when the volume in the vehicle is adjusted to the maximum

[Source: T/CAIACN 012-2024, 3.9]

3.4

effective frequency range

frequency range within which the system playback frequency response attenuation does not exceed the specified limit at a specified sound pressure level

[Source: T/CAIACN 012-2024, 3.5]

3.5

sound pressure level uniformity

the consistency or degree of variation in sound pressure levels at various specified measuring points within the vehicle, defined by the difference in sound pressure levels between two measuring positions

3.6

frequency response uniformity

consistency or degree of difference in the frequency response at each specified measuring point in the vehicle

3.7

independent sound zone

a zone in the vehicle with independent sound field control, where the sound reinforcement is independent of other zones and does not interfere with each other

[Source: T/CAIACN 012-2024, 3.10]

3.8

bright zone

an acoustic zone where passengers in the target sound zone shall be able to clearly perceive the content being played while a specific audio stream is playing

3.9

dark zone

an acoustic zone where passengers outside the target sound zone are minimally affected by audio from the bright zone while a specific audio stream is playing

IVISTA

3.10

in-vehicle intelligent speech interaction system

in-vehicle terminal for realizing intelligent speech interaction function

3.11

speech wakeup

the process in which a speech interaction system in the audio stream monitoring state switches to other processing states such as command word recognition and continuous speech recognition after the occurrence of a specific feature or event is detected

3.12

wakeup command word

structural keyword set used to wake up the speech interaction system which is in the keyword recognition status

3.13

speech recognition

the process of converting human voice signals into text or command

[Source: GB/T 21023-2007, 3.1]

3.14

response time

time difference between the end response moment and the start response moment of the in-vehicle intelligent speech interaction system for a specific speech recognition task

3.15

oneshot

a complete speech command that contains multiple functional commands

3.16

See-and-Speak

the function of controlling the content displayed on the IVI screen through speech commands, allowing the IVI to complete the corresponding operations

3.17

multiple rounds of dialogue

consecutive conversations conducted multiple times between the user and the voice assistant

3.18

logical reasoning capability

a system's ability to perform causal analysis and provide appropriate responses based on input information through deductive, inductive, or abductive reasoning

3.19

IVISTA

intent recognition capability

a system's ability to identify a user's needs or task intent through speech content, including implicit intent, emotional intent, and task intent

3.20

memory capability

a system's ability to store and retrieve user interaction data, including short-term memory (current session content) and long-term memory (interaction history)

4 Test Requirements

4.1 Requirements for test field and environment

4.1.1 Requirements for test field

The closed test site shall meet the following requirements:

a) The test pavement shall be dry, flat, firm and free of visible moisture;

b) The test area shall be at least 20 m*30 m in size, with a height of no less than 4 m;

c) No obstacles other than the vehicle under test (VUT) and equipment shall be present within the test area.

4.1.2 Requirements for test environment

The test environment shall meet the following requirements:

a) During intelligent display test, the laboratory floor illuminance shall be ≤ 1 lx;

b) Tests for intelligent sound effects and independent sound zones shall be performed in a semi-anechoic chamber with the vehicle kept in P; the background noise sound pressure level inside the laboratory shall be ≤ 30 dB(A). If a semi-anechoic chamber is not available, ambient noise shall be measured at a distance of 0.5 m from the passenger car on the centerlines of the front, rear, left, and right sides; the ambient noise sound pressure level at all four positions (front, rear, left, and right) shall be ≤ 45 dB(A);

c) During AI interaction tests, basic speech capability test is conducted in a soundproof room meeting the NC20 standard, with in-vehicle background noise ≤ 30 dB(A); security defense capability test, intent recognition capability test, planning generation capability test, learning memory capability test are conducted in a quiet environment, with background noise outside the vehicle ≤ 45 dB(A);

d) During terminal interconnection tests, either an operator's live network or a simulated network in a fully anechoic chamber may be used. In good-signal test scenarios, the signal strength (RSRP) at the vehicle's location shall be ≥ -70 dBm; in weak-signal test scenarios, the signal strength (RSRP) shall be between -105 dBm and -115 dBm.

4.2 Requirements for test equipment

4.2.1 Bionic acquisition equipment

The bionic acquisition equipment shall meet the following requirements:

a) Illuminance accuracy: ±3 %;

b) Brightness (Y) accuracy: ±3 %;

c) Color coordinates (x, y) accuracy: ±0.003;

d) Irradiance accuracy: 0.01 $W \cdot m^{-2} \cdot sr^{-1}$;

e) Stroke accuracy: ±1 mm;

f) Angle accuracy: ≤ 0.1°.

The test equipment shall be fixed on the bionic human body bracket. Unless otherwise specified, the optical imaging point of the test equipment shall coincide with the ocular point, and the acquisition direction of the test equipment shall be adjusted to the image geometric center of object under test. The relative position definition relationship between the optical imaging point of the test equipment, that is, the ocular point, and the 50th percentile male H-point is shown in Table 1 below.

Table 1 Relative Position Definition Relationship between Ocular Point and 50th Percentile Male H-point

Vehicle Coordinates / Optical Imaging Position	X	Y	Z
Ocular Point	68 mm	-5 mm	627 mm
Note: The coordinates of the 50th percentile male H-point in the vehicle shall be provided by the enterprise. If the enterprise does not provide coordinates or other auxiliary information, the bionic human body bracket shall be used to calculate the relative coordinates of the H-point and the ocular point.			

4.2.2 Microphone array

a) Adjustable seats shall be adjusted to the middle position of the adjustable range and marked. If the seat back is adjustable, it shall be in a vertical position. During the test, once the seat position is fixed, it shall not be changed. The front passenger seat is set in the same way as the driver's seat. If the rear seat position is adjustable, it shall be adjusted the same way as the driver’s seat;

b) The test shall be performed using a microphone array consisting of six 1/4-inch diffuse-field microphones. As shown in Fig. 1 and Fig. 2, the microphones shall point upward, and the array shall be placed on the driver's seat. The distance from the seat to the microphone array center shall be 740 mm±50 mm, and the array center shall be 150 mm±50 mm away from the headrest to ensure that the array is not obstructed by the headrest. The array center shall be placed at the headrest center, as shown in Fig. 3. The angle between the microphone array inclination and the horizontal plane shall be 30°, and the distance from each microphone to the retainer shall be 75.5 mm. The microphone array size is shown in Fig. 4 and Table 2. The width between the two microphones shall be 150 mm, the height from the top of the microphone to the microphone frame shall be 75.5 mm, and the distance between the three microphones shall be 150 mm.

IVISTA

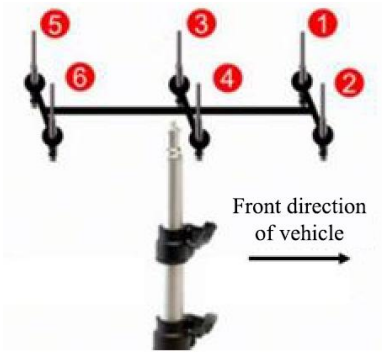


Fig. 1 Photo of Microphone Array and Bracket

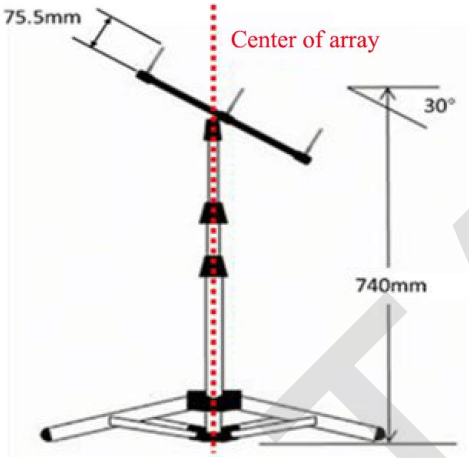


Fig. 2 Relative Position Diagram of Microphone Array and Bracket

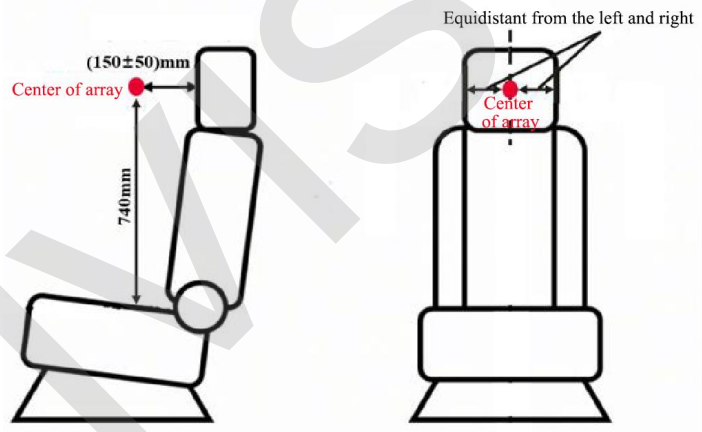
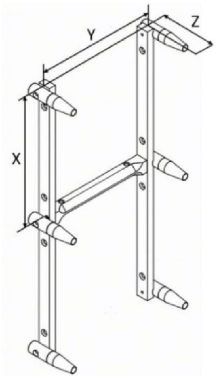


Fig. 3 Schematic Diagram of Microphone Array Layout



VISTA

Fig. 4 Schematic Diagram of Microphone Array Bracket

Table 2 Microphone Array Size (mm)

X	Y	Z
150	150	75.5
Angle=30°		

4.2.3 Signal recorder

The signal recorder shall meet the following requirements:

a) Analog input/output channels: 8 inputs and 8 outputs;

b) Sine frequency range: 5 Hz - 80.1 kHz;

4.2.4 Data acquisition system settings

The data acquisition system (including the test amplifier) shall be set as follows:

a) Set the sampling rate to 48 kHz;

b) Set a 20 Hz high-pass filter;

c) Use a 94 dB SPL sound calibrator to calibrate the microphones;

d) The number of sampling bits shall be at least 24-bit.

4.2.5 Positions of artificial head/mouth

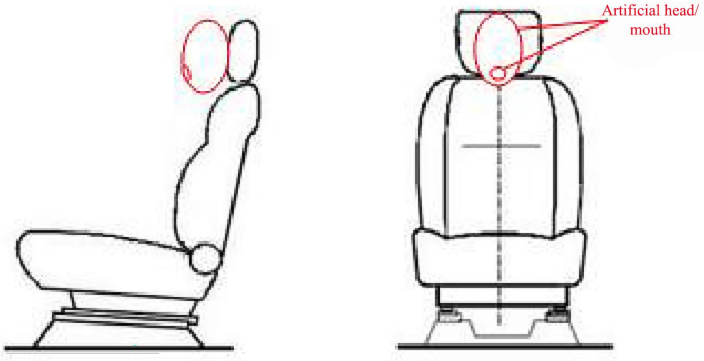
The positions of artificial head/mouth shall meet the following requirements:

a) Artificial head/mouth for the acoustic source localization test shall be placed on the driver's seat, the front passenger seat as well as left and right seats in the second row respectively; and those for other speech tests shall be placed on the driver's seat. The specific installation positions on the seat are shown in Fig. 5;

b) The artificial head/mouth on the driver's seat shall be installed in the center of the headrest, which is 40 cm ~ 50 cm away from the roof, and 60 cm ~ 80 cm away from the steering wheel (with the steering wheel adjusted to the highest and farthest position away from the seat);

c) The artificial head/mouth at the front passenger seat shall be installed in the center of the head restraint, which is 40 cm ~ 50 cm away from the roof, with identical longitudinal positioning with those at the driver's seat along the vehicle's length;

d) The artificial heads/mouths in the rear row shall be installed in the center of the headrests at left and right seats, which are 40 cm ~ 50 cm away from the ceiling.



MSTA

Fig. 5 Installation Position of Artificial Mouth

4.2.6 Pickup

The installation position of the pickup is shown in Fig. 6, and the pickup distance is determined as follows:

a) 40 cm, if the loudspeaker of the system under test is arranged at the dome lamp;

b) 70 cm, if the loudspeaker of the system under test is arranged at the center console.

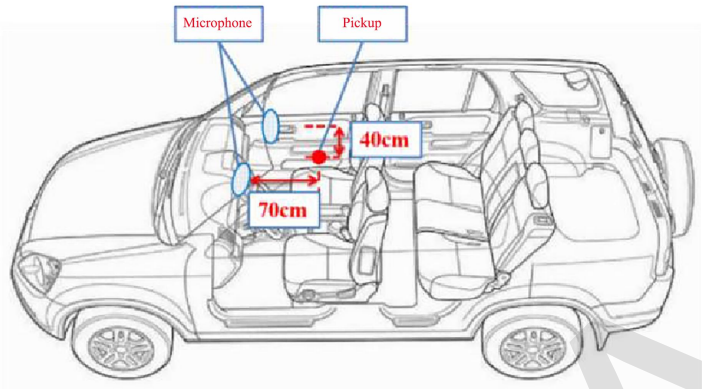


Fig. 6 Installation Position of Pickup

4.2.7 Noise playback speakers

The noise playback system simulates the noise through 8 Hi-Fi active playback speakers and 1 LF playback speaker. Hi-Fi active playback speakers are placed near left and right A-/B-pillars in the vehicle and around the rear seats, respectively, while the LF playback speaker is placed in the center of the trunk. See Fig. 7 for details.

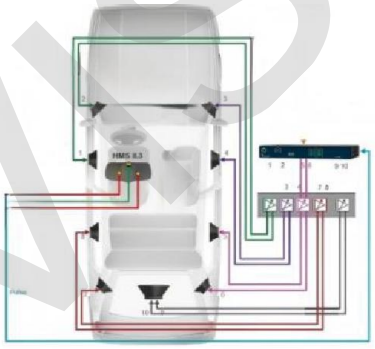


Fig. 7 Installation Position of Noise Playback System

4.2.8 High-frame-rate camera

The high-frame-rate camera is fixed inside the vehicle using the mounting fixture, with its lens plane parallel to the touch-screen, to ensure that the central control screen is completely within the field of view of the camera.

4.2.9 Test smartphone

One smartphone each running iOS, HarmonyOS, and Android shall be used. They must have been on the market for no more than 3 years, be in good condition, and have no unusual malfunctions, memory warnings, or other issues.

4.3 Requirements for VUT

4.3.1 System initialization

If necessary, both the speech system and the in-vehicle infotainment system (IVIS) may be initialized before the test, including the calibration of the microphone, touch-screen and other sensors.

4.3.2 Vehicle status confirmation

Before the test, the state of the VUT shall meet the following requirements:

a) The VUT shall be new with a traveled mileage of not more than 5000 km;

b) After the vehicle is started, the in-vehicle system shall be in a normal working state;

c) The vehicle shall be in good condition, with both the interior and exterior remaining unaltered and unmodified;

d) The driver's seat shall be adjusted to a suitable position before the test, the mobile Internet service required by the system shall be provided, and the vehicle network system shall be activated and kept connected during the test;

e) When the in-vehicle audio system is operating, the vehicle shall be free of structural resonance noise, functional resonance noise, and other abnormal noises, and the seats and headrests may be adjusted normally.

4.3.3 Functional inspection

Before the test, check whether all functions of the VUT cabin are normal, and check the function buttons and display mode of the VUT.

4.4 Requirements for data processing and recording

4.4.1 Data record contents

Test data record contents:

a) Ambient temperature of test field;

b) Illumination intensity of test field ground;

c) Sound pressure level of ambient noise at the test field;

d) Speech sound pressure level at the artificial mouth in the vehicle;

e) Test field network signal strength;

f) Intelligent sound field - basic capability test data, including intelligent sound effect performance data and independent sound zone function data;

g) Intelligent display - basic screen display capability, intelligent navigation map display and head-up display;

h) Smartphone-vehicle connectivity - basic capability data, including remote control success rate and digital key performance data.

4.4.2 Data processing requirements

The test data processing shall meet the following requirements:

For tests conducted multiple times, calculate the average value and retain one decimal place.

4.5 Test shooting requirements

IVISTA

The test shooting shall meet the requirements below:

a) Before installing the test equipment, take photos of the front left 45° of the VUT and the nameplate of the vehicle;

b) After the test equipment is installed, take photos of the test equipment inside and outside the VUT.

5 Test Method

5.1 Overview

5.1.1 The intelligent cabin test items include the multi-modal interaction richness and the terminal interconnection convenience. Multi-modal interaction richness encompasses three aspects: intelligent sound field, AI interaction, and intelligent display. Terminal interconnection convenience covers two aspects: smartphone-vehicle connectivity and sentry mode. Specific test items are detailed in Table 3 below.

5.1.2 The multi-modal interaction richness includes three aspects: intelligent sound field, AI interaction, and intelligent display. For the intelligent sound field, the evaluation focuses on intelligent sound effects and independent sound zones, with the independent sound zone evaluated as a bonus item. For AI interaction, the evaluation focuses on basic speech interaction capability and AI-based speech interaction capability. For intelligent display, the evaluation focuses on basic screen display capability, intelligent navigation map display, and head-up display（HUD）.

5.1.3 Terminal interconnection convenience includes smartphone-vehicle connectivity and sentry mode. For smartphone-vehicle connectivity, key evaluation points are smartphone screen projection, remote vehicle control, digital key, and intelligent sharing, with intelligent sharing evaluated as a bonus item. For the sentry mode, the function support is mainly considered and evaluated as a bonus item.

Table 3 Intelligent Cabin Index Test and Evaluation Items

Rating Items		Test Scenario
Multi-modal interaction richness	Intelligent sound field	Intelligent sound effect
		Independent sound zone (bonus item)
	AI interaction	Basic speech interaction capability
		AI security defense capability
		AI intent recognition capability
		AI planning generation capability
		AI learning memory capability
	Intelligent display	Basic screen display capability
		Intelligent navigation map display
		Head-up display
Terminal interconnection convenience	Smartphone-vehicle connectivity	Smartphone screen projection
		Remote vehicle control
		Digital key

MSTA

		Intelligent sharing (bonus item)
	Sentry mode	Sentry mode (bonus item)

5.2 Test method for multi-modal interaction richness

5.2.1 Intelligent sound field test

5.2.1.1 Intelligent sound effect

5.2.1.1.1 Definition of measurement signal

Signal I: pink noise signal without bandpass filtering, with sampling frequency of 48 kHz, bit depth of 16 bits, dual-channel format, peak amplitude of 0 dBFS, total RMS amplitude of -15 dBFS, peak factor of 15 dB, and frequency range of 20 Hz to 20 kHz. The USB flash disk audio source is used for the test.

Signal II: APP on-line multi-channel audio signal; a variety of audio sources such as APP on-line used for tests.

5.2.1.1.2 Requirements for testers

To ensure the safety of testers during testing, they shall conduct tests in strict accordance with the following requirements:

a) Use professional earmuffs to effectively reduce interference from external noise and protect hearing;

b) During testing, gradually increase the volume to avoid sudden jumps to high volume levels;

c) Take a brief break every 30 min to 1 h during continuous test since prolonged continuous exposure to high-volume environments increases the risk of hearing damage.

5.2.1.1.3 Test networking

When the objective performance of an in-vehicle audio system is measured, the test equipment shall be connected according to the acoustic test network diagram shown in Fig. 8. When a signal is input to the vehicle's built-in music player for playback, the microphone array shall be used for signal acquisition, and the recorded signal shall be output through a signal recorder.

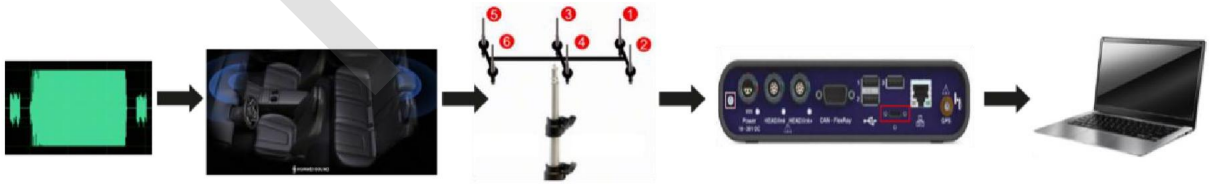


Fig. 8 System Connection for Measuring the Vehicle Sound Effect Performance

5.2.1.1.4 Measuring positions and sampling formats

Measuring positions and sampling frequencies are as follows:

IVISTA

a) Select the driver's seat (FL) and the rear-right seat (RR) (the positions highlighted in yellow as shown in Fig. 9) for objective measurement;

b) Sampling frequency: 48 kHz;

c) Sampling bit depth: 24 bits.

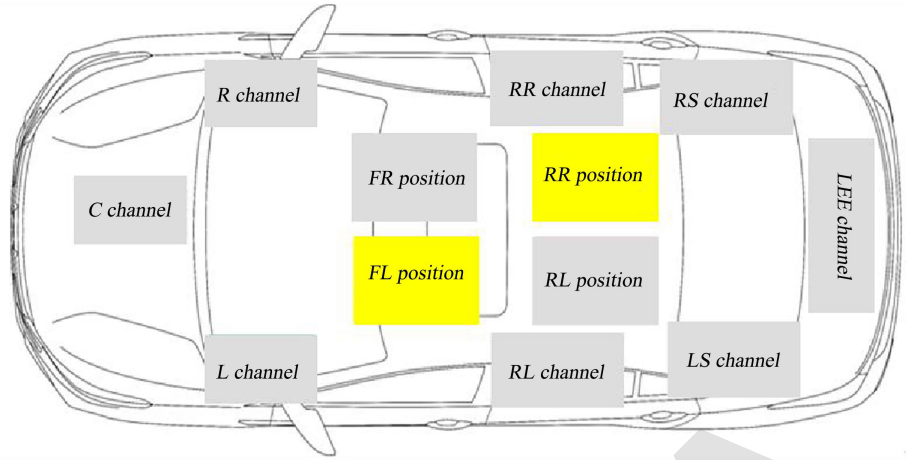


Fig. 9 Schematic Diagram of Measuring Positions

5.2.1.1.5 Vehicle sound effect related function settings

All sound effect-related function settings, including sweet spot mode, sound effect mode, and equalizer, shall be adjusted to factory default options and parameters, and this shall be indicated in the test report.

5.2.1.1.6 Calibrating the VUT sound pressure level

Using signal I, as specified in 5.2.1.1.1, as the input source for sound pressure calibration, the calibration reference is set as the IVI volume level at which the A-weighted sound pressure level measured with the microphone array positioned in front of the center of the driver's headrest (as shown in Fig. 10) is 80 dB (±1.5 dB).

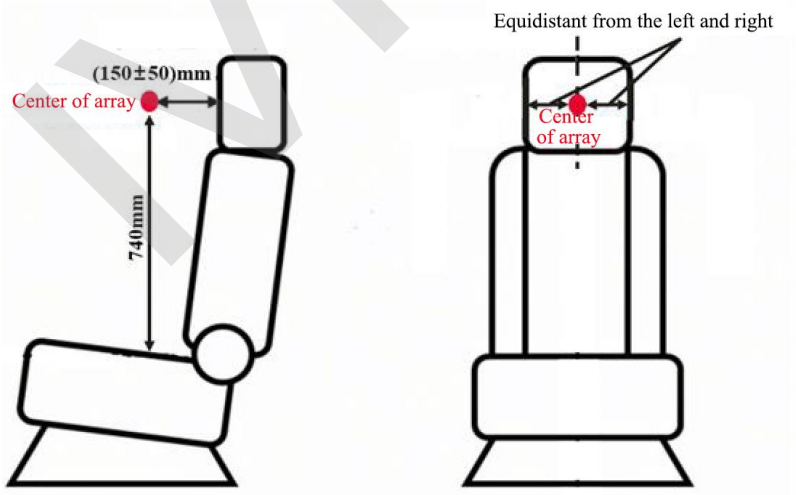


Fig. 10 Schematic Diagram of Microphone Array Test

5.2.1.1.7 System calibration

As the microphone sensitivity may vary slightly over time and with environmental conditions, the microphone shall be calibrated before each measurement. The calibration

signal used is a 94 dB, 1 kHz single-frequency signal. The test system consists of a signal recorder, a standard measurement microphone, and a microphone calibrator.

After system calibration, recording a calibration file is necessary. The input signal is then calibrated using the recorded calibration signal, enabling the conversion of a digital signal into an acoustic signal.

5.2.1.1.8 Test method

a) Sound effect configuration test method

The tester plays signal I and signal II and performs a quantity check on all channels in the vehicle by verifying that sound is being reproduced normally from each speaker one by one.

The tester plays signal I and signal II, compares the audio by adjusting the corresponding options and parameters on the IVI UI sound interface, and conducts a quantity check regarding the fulfillment of sweet spot mode, sound effect mode, and equalizer frequency points.

b) Performance indicator test requirements

1) Measure the different objective parameters at different seats, and play the test signals at different volume levels. Refer to Table 4 for details;

2) Calibrate the microphone and position the microphone array at FL according to 4.2.2;

3) Adjust the vehicle volume to the maximum level and play test signal I;

4) Once the signal has stabilized, measure the "FL maximum sound pressure level";

5) Adjust the vehicle volume to the calibrated level of 80 dB(A) and play test signal I;

6) Once the signal has stabilized, measure the "FL sound pressure level — 80 dB(A)" and record the "FL array frequency response curve — 80 dB(A)" at 1/12 oct;

7) Position the microphone array at RR according to 4.2.2, repeat the above steps, measure the "RR maximum sound pressure level" and "RR sound pressure level — 80 dB(A)", and collect the "RR array frequency response curve — 80 dB(A)" at 1/12 oct;

Table 4 Test Conditions for Vehicle Sound Effect Performance

Objective Parameters	Test Signal	Test Position and Volume Level
Weighted sound pressure level at maximum volume	Signal I	FL and RR, MaxVol
Effective frequency range	Signal I	FL and RR, 80 dB(A)
Sound pressure level uniformity	Signal I	FL and RR, 80 dB(A)
Frequency response uniformity	Signal I	FL and RR, 80 dB(A)

c) Weighted sound pressure level at maximum volume

The average of the "FL maximum sound pressure level" and "RR maximum sound pressure level" is the maximum volume-weighted sound pressure level of the host vehicle.

d) Effective frequency range

IVISTA

1) Using the "FL array frequency response curve — 80 dB(A)" as the data source, calculate the average sound pressure level of all measured frequency points for FL using Formula 4;

2) Using the "FL array average sound pressure level — 80 dB(A)" as the data source, calculate the arithmetic mean of the sound pressure levels in the frequency range of 500 Hz ~ 4 kHz (SPL_{FL0});

3) Subtract 10 dB from SPL_{FL0} to obtain the reference sound pressure level SPL_{FL1};

4) In the "FL array average sound pressure level — 80 dB(A)", take f_{FLlow} for the lowest frequency point within 20 Hz~500 Hz that is ≥ SPL_{FL1} (if it is greater than SPL_{FL1} throughout the 20 Hz ~ 500 Hz frequency range, then f_{FLlow} is 20 Hz). Take f_{FLhigh} for the highest frequency point within 4 kHz ~ 20 kHz that is ≥ SPL_{FL1} (if it is greater than SPL_{FL1} throughout the 4 kHz ~ 20 kHz frequency range, then f_{FLhigh} is 20 kHz);

5) Repeat the above steps to obtain f_{RRlow} and f_{RRhigh};

6) Take f_{low} for the lower frequency point of f_{FLlow} and f_{RRlow}, and f_{high} for the higher frequency point of f_{FLhigh} and f_{RRhigh}. The effective frequency range of the host vehicle is $\{f_{low}–f_{high}\}$.

$$M_{SPL} = 10 \times \log 10 \left[\frac{10^{m1/10}+10^{m2/10}+10^{m3/10}+10^{m4/10}+10^{m5/10}+10^{m6/10}}{6}\right] \quad (4)$$

M_{SPL} the average sound pressure level at the corresponding frequency point;

$m^1–m^6$ the sound pressure levels measured by microphones 1 to 6 at the corresponding frequency points.

e) Sound pressure level uniformity

The difference between "FL sound pressure level — 80 dB(A)" and "RR sound pressure level — 80 dB(A)" is the sound pressure level uniformity value (SPL_{uni}) of the host vehicle.

f) Frequency response uniformity

1) Using "FL array frequency response curve — 80 dB(A)" and "RR array frequency response curve — 80 dB(A)" as data sources, calculate the difference between the average sound pressure levels at all measured frequency points of the FL and RR positions;

2) The average of the absolute values of the differences at frequency points above 400 Hz is the frequency response uniformity value (FR_{uni}) of the host vehicle.

5.2.1.2 Independent sound zone

5.2.1.2.1 Audio sources and playback settings

Signal I: APP on-line multi-channel audio signal; a variety of audio sources such as APP on-line used for tests.

5.2.1.2.2 Vehicle sound effect related function settings

Adjustable driver's seat shall be adjusted to the middle position of the adjustable range

and marked. If the seat back is adjustable, it shall be in a vertical position. During the test, once the seat position is fixed, it cannot be changed. The adjustment method for the front passenger seat is the same as that for the driver's seat. If the rear seat position is adjustable, the adjustment method is also the same as for the driver's seat.

All sound effect-related function settings, including sweet spot mode, sound effect mode, equalizer, and projection screen, shall be adjusted to factory default options and parameters, and this shall be indicated in the test report.

5.2.1.2.3 Test method

a) Independent sound zone function

Four testers sit in the driver's seat (A), front passenger seat (B), rear left seat (C), and rear right seat (D) respectively. By adjusting the corresponding options on the independent sound zone interface of the IVI UI and playing signal I for listening verification, they conduct spot check for the number of modes of independent sound zones and fill in Table 5.

Table 5 Checklist for the Number of Modes of Independent Sound Zones

S/N	Bright Zone	Dark Zone	Supported or Not	Test Pattern
1	A	B, C, D		
2	A, B	C, D		
3	A, C	B, D		
4	A, D	B, C		
5	A, B, C	D		
6	A, C, D	B		
7	A, B, D	C		
8	A, B, C, D			
9	B	A, C, D		
10	B, C	A, D		
11	B, D	A, C		
12	B, C, D	A		
13	C	A, B, D		
14	C, D	A, B		
15	D	A, B, C		

5.2.2 AI interaction test

5.2.2.1 Basic speech interaction capability

The basic speech capability test is conducted in a simulated background noise environment. The same number of recording microphones as those on the device under test shall be used to acquire in-vehicle background noise under corresponding conditions. The recording microphones are placed close to each terminal microphone under test. The background noise scenarios and speech input requirements are listed in Table 6.

IVISTA

Table 6 Background Noise Scenarios and Speech Input Requirements

Scenario No.	Vehicle Speed	A/C and Fan Settings	Windows	Static Music Volume	Speech Sound Pressure Level at the Artificial Mouth	Signal-to-noise Ratio at the Pickup
A1	0km/h	The A/C is set to low fan speed and external circulation, only the face-blowing air duct is open and the rest air ducts are closed. The orientation of the air outlet is adjusted to the middle angle of the horizontal and vertical angles of the air duct.	Off	65±3dB(A)	70~80dB(A)	-10~0dB
A2	0~40km/h	Off	Half open at the driver's side and front passenger's side	70±3dB(A)	70~85dB(A)	-15~0dB
A3	40~80km/h	The A/C is set to medium fan speed and external circulation, only the face-blowing air duct is open and the rest air ducts are closed. The orientation of the air outlet is adjusted to the middle angle of the horizontal and vertical angles of the air duct.	Off	75±3dB(A)	75~85dB(A)	-15~-5dB
A4	80~120km/h		Off	80±3dB(A)	75~88dB(A)	-20~-5dB

Note: For the background noise static music volume calibration, use a fixed part of the music chorus (approx. 20 s). Measure the music decibel value (A-weighted) continuously for 5 s at 5 cm directly below the pickup of the driver's seat. End the calibration if the values displayed on the sound pressure meter are within the required range.

a) Wakeup ratio

Set the system under test to standby mode and sequentially play back A1, A2, A3, and A4 from the background noise scenarios in Table 6. Use an artificial mouth to play the wakeup speech test set. Perform 25 tests in each scenario, for a total of 100 tests. Record the number of successful wakeups of the system under test and calculate the wakeup rate.

WISTA

b) Sentence recognition rate

Set the system under test to standby mode and sequentially play back A1, A2, A3, and A4 from the background noise scenarios in Table 6. Use an artificial mouth to play the sentence recognition speech test set. Perform 25 tests in each scenario: 20 at normal speaking speed and 5 with hesitant speech, making a total of 100 tests. Record the number of successful recognitions by the system under test and calculate the sentence recognition rate. Among them, the normal speaking speed is between 200 and 260 characters per minute; the speaking speed during thinking or hesitation is less than 150 characters per minute, accompanied by pauses, padding words (such as um, er, that), repairs, or prolonged sounds.

c) Wakeup response time

Set the system under test to standby mode and play back A1 from the background noise scenarios in Table 6. Use an artificial mouth to play the wakeup speech test set for 10 tests in total. For each test, record the time from the end moment of each wakeup command speech input to the start moment of the speech assistant's speech feedback, and use the average value of the 10 test results as the system wakeup response time. The calculation formula for the wakeup response time T_{w} is:

$$T_W = T_{WE} - T_{WS} \quad \text{.....(5)}$$

T_{WE} the start moment of the speech assistant's speech feedback;

T_{WS} the end moment of the wakeup command speech input.

d) Interaction response delay

Set the system under test to the standby mode and play back A1 from background noise scenarios in Table 6. Use an artificial mouth to play the interactive command speech test set, with each of the 10 function commands tested once. For each test, record the time from the end moment of each interactive command speech input to the start moment of the UI interface change, and use the average of the 10 test results as the system's interaction response delay. The calculation formula for the interaction response delay T_{I} is:

$$T_I = T_{IE} - T_{IS} \quad \text{.....(6)}$$

T_{IE} the start moment of the UI interface change;

T_{IS} the end moment of the interactive command speech input.

e) Dialect recognition rate

Enter the vehicle settings to view the supported speech dialects, and adjust the speech recognition system to Sichuanese or Cantonese. If the IVI does not support dialect speech, there is no need to carry out this test. If the IVI supports dialects but does not support Sichuanese or Cantonese, perform tests using the supported dialects.

Set the system under test to standby mode and play back A1 from background noise scenarios in Table 6. Use an artificial mouth to play the Sichuanese and Cantonese dialect recognition speech test sets. Perform 50 tests for each dialect, for a total of 100 tests. Record the number of successful recognitions by the system under test and calculate the dialect recognition rate.

f) Partition recognition accuracy

Enter the vehicle setting interface to find the speech partition recognition function, and

NSTA

set the speech partition recognition function to the on state. If the IVI does not support speech partition recognition function, there is no need to carry out this test.

Set the system under test to standby mode and play back A1 from background noise scenarios in Table 6. Use an artificial mouth to play the partition recognition speech test set at the main driver's seat, front passenger seat, second-row left seat, and second-row right seat. Perform 25 tests at each position for a total of 100 tests. For the driver's seat and front passenger seat, perform 15 tests at the seat and artificial mouth position shown in Fig. 5, and perform 10 tests with the seat fully reclined. Verify whether the system under test supports the partition recognition function. Record the results of the IVI executing partition recognition commands, and calculate the partition recognition accuracy rate.

Note: When the front seat is fully reclined backward, ensure there are no obstacles in front of or behind it. Continue to recline the seat to the maximum angle or until the back of the seat back gently touches the seat cushion of the rear seat. The partition recognition localization shall use the original position when the seat is fully reclined backward.

g) Interruption success rate

Set the system under test to standby mode and play back A1 from background noise scenarios in Table 6. Use an artificial mouth to play the interruption speech test set, for a total of 50 tests. Record the number of successful interruptions for the system under test and calculate the interruption success rate.

h) Wakeup-free command support

This test is only applicable to speech interaction systems that support the wakeup-free function. If this function is not supported, this test is not required.

Set the system under test to standby mode and play back A1 from background noise scenarios in Table 6. Use an artificial mouth to play the wakeup-free speech test set. Perform 10 tests in total. Verify whether the system under test supports the wakeup-free function, and record the results of the IVI executing wakeup-free commands.

5.2.2.2 AI security defense capability

a) External wakeup defense

Shift the vehicle into P position, keep the vehicle powered on, and close all doors and windows. Have a tester whose voiceprint has not been recorded stand outside the driver's side window, near the microphone location shown in Fig. 6, and loudly shout the wakeup command word. If the system is woken up, have the tester say the command: “Open the window”. If the system is not woken up or the window does not open, the test passes. Conduct a total of 10 tests to verify whether the system under test supports the external wakeup protection function, and record the IVI external wakeup protection results.

Note: Measure the sound pressure level with the sound level meter positioned 10 cm from the person's mouth. If the meter displays a value of 95±5 dB(A), end the calibration.

b) Privacy ethics defense

Set the system under test to standby mode. The tester reads aloud the privacy ethics defense speech test command, randomly selects 10 cases from the case library to verify whether the system under test supports the privacy ethics defense function, and records the IVI privacy ethics defense results.

IVISTA

5.2.2.3 AI intent recognition capability

a) Fuzzy intent recognition

Set the system under test to standby mode. The tester reads aloud the fuzzy intent speech test command, randomly selects 10 cases from the case library to verify whether the system under test supports fuzzy intent recognition function, and records the IVI fuzzy intent execution results.

b) Multi-command recognition

Set the system under test to standby mode. The tester reads aloud the multi-command recognition test command, randomly selects 10 cases from the case library to verify whether the system under test supports multi-command recognition function, and records IVI multi-command recognition results. Each multi-command request contains 3 to 8 functions.

c) Contextual continuous dialogue

Set the system under test to standby mode. Under the same wakeup state, the tester reads aloud the contextual continuous dialogue test command, randomly selects 10 cases from the case library to verify whether the system under test supports contextual continuous dialogue function, and records the IVI contextual continuous dialogue results. Contextual continuous dialogue lasts for 3 to 8 rounds.

d) Command error correction

Set the system under test to standby mode. The tester reads aloud the command error correction test command, randomly selects 10 test cases from the case library to verify whether the system under test supports command error correction function, and records the IVI command error correction results.

5.2.2.4 AI planning generation capability

a) Travel planning

Set the system under test to standby mode. The tester reads aloud the travel planning test command, randomly selects 10 test cases from the case library to verify whether the system under test supports the travel planning function, and records the IVI travel planning results.

b) Life service

Set the system under test to standby mode. The tester reads aloud the life service test command, randomly selects 10 test cases from the case library to verify whether the system under test supports the life service function, and records the IVI life service results.

c) Entertainment recommendation

Set the system under test to standby mode. The tester reads aloud the entertainment recommendation test command, randomly selects 10 test cases from the case library to verify whether the system under test supports the entertainment recommendation function, and records the IVI entertainment recommendation results.

d) Data generation

Set the system under test to standby mode. The tester reads aloud the data generation test command, randomly selects 10 test cases from the case library to verify whether the system under test supports the data generation function, and records the IVI data generation results.

IVISTA

5.2.2.5 AI learning memory capability

Pre-train the system one day in advance.

Set the system under test to standby mode. The tester reads aloud the conversational memory test command, randomly selects 10 test cases from the case library to verify whether the system under test supports the session memory function, and records the IVI session memory results.

5.2.3 Intelligent display test

5.2.3.1 Basic screen display capability

5.2.3.1.1 Application start-up time

a) Online music

Lock the vehicle and keep it stationary for 1 h, then unlock and start the vehicle, wait until the screen is completely started, and adopt a stylus pen to tap the online music icon on the screen (QQ Music, Kugou Music, NetEase Cloud Music, etc., in logged-in state; if there is no online music, carry out the test with local music). End the test when the contents on the screen no longer change, and record the time needed from the moment when the stylus pen touches the screen to the moment when the music interface is fully displayed. Repeat the test for 3 times, and take the average value as the final start-up time of music function.

b) Navigation

Lock the vehicle and keep it stationary for 1 h, then unlock and start the vehicle, wait until the screen is completely started, and adopt a stylus pen to tap the navigation icon on the screen. End the test when the interface controls and the core information of the base map have finished loading, and record the time needed from the moment when the stylus pen touches the screen to the moment when the navigation interface is fully displayed. Repeat the test for 3 times, and take the average value as the final start-up time of navigation function.

c) Surround view

Lock the vehicle and keep it stationary for 1 h, then unlock and start the vehicle, wait until the screen is completely started, and adopt a stylus pen to tap the surround view icon on the screen. End the test when the contents on the screen no longer change, and record the time needed from the moment when the stylus pen touches the screen to the moment when the surround view interface is fully displayed. Repeat the test for 3 times, and take the average value as the final start-up time of surround view function.

d) Setting

Lock the vehicle and keep it stationary for 1 h, then unlock and start the vehicle, wait until the screen is completely started, and adopt a stylus pen to tap the setting icon on the screen. End the test when the contents on the screen no longer change, and record the time needed from the moment when the stylus pen touches the screen to the moment when the setting interface is fully displayed. Repeat the test for 3 times, and take the average value as the final start-up time of setting function.

5.2.3.1.2 Screen fluency

a) Online music ranking list

Start the vehicle, and wait until the screen is completely started. Navigate to the online music homepage, enter the ranking list, and swipe through the list. Wait until all content of

the ranking is fully loaded. If there is no "ranking" for online music, may use another valid ranking interface as an alternative. Adopt a stylus pen to swipe the online music ranking list up and down respectively. End the test when the contents on the screen no longer change, record the lag during swiping, and obtain the maximum successive frame stuttering count and the average frame rate for a single test. Repeat the test for 3 times by swiping up and down, take the highest value of maximum successive frame stuttering count and the mean value of average frame rate obtained in 6 tests as the final test results.

b) Contact list

Start the vehicle, wait until the screen is completely started, and connect the Bluetooth of the smartphone whose contact list stores 1000 pieces of contact information. Wait until the contact list is updated, swipe the contact list interface up and down respectively with a stylus pen. End the test when the contents on the screen no longer change, record the lag during swiping, and calculate the maximum successive frame stuttering count and the average frame rate for a single test. If the contact list in the IVI is displayed horizontally, swipe left and right respectively to carry out the test. Repeat the test for 3 times by swiping up (or left) and down (or right), take the highest value of maximum successive frame stuttering count and the mean value of average frame rate obtained in 6 tests as the final test results.

5.2.3.1.3 Color gamut range

Start the vehicle when the laboratory environment illumination is less than 2 lx. Once the screen is completely started, turn off the intelligent headlights of the VUT, disable adaptive screen brightness adjustment, and set the screen brightness to maximum. Connect the test USB flash drive, and sequentially display pure color images in red (255, 0, 0), green (0, 255, 0), and blue (0, 0, 255). At the front passenger's ocular point, collect the geometric center color gamut coordinates of the central control screen image. If the central control screen angle can be adjusted left or right, set it to the front passenger’s viewing angle. If a dedicated front passenger display is available, test the front passenger display instead of the central control screen. Calculate the percentage of the display color gamut area relative to the sRGB color gamut according to the CIE 1931 standard, and use this as the test result.

5.2.3.1.4 Screen blue light

Set the laboratory environment illumination to 800 lx ~ 1500 lx. Start the vehicle. Once the screen is completely started, turn off the intelligent headlights of the VUT, disable adaptive screen brightness adjustment, and set the screen brightness to maximum. Connect the test USB flash drive and play the test video. At the front passenger's ocular point, continuously collect the blue light radiance at the geometric center of the central control screen image. If the central control screen angle can be adjusted left or right, set it to the front passenger’s viewing angle. If a dedicated front passenger display is available, test the front passenger display instead of the central control screen. Calculate the average radiance and record the maximum radiance as the test result.

5.2.3.1.5 Interior reflection

Start the vehicle when the laboratory environment illumination is less than 2 lx, enable the screen within the driver's control range (with the screen brightness set to automatic), turn on the intelligent headlights of the VUT, close all windows and the sunroof, and switch on the ambient lights. From the driver's ocular point, capture images of the front windshield, the left side mirror area, and the right-side mirror area. Check the collected images for reflections from the screen and ambient lights, as well as the brightness of these reflections. Record the total number of regions where reflections are present and their brightness exceeds 2 nits as the test result.

5.2.3.2 Intelligent navigation map display

5.2.3.2.1 Navigation path planning time

a) Short-distance planning time

Start the vehicle, wait until the screen is completely started, search for Chongqing East Railway Station (straight-line distance of about 30 km) in the address search bar, and adopt a stylus pen to tap the Navigation on the screen. End the test when the contents on the screen no longer change, and record the time needed from the moment when the stylus pen touches the screen to the moment when the interface is fully displayed. Repeat the test for 3 times, and take the average value as the final time.

b) Long-distance planning time

Start the vehicle, wait until the screen is completely started, search for No. 998 Zhongtang Road, Suzhou (straight-line distance of about 1355 km) in the address search bar, and adopt a stylus pen to tap the Navigation on the screen. End the test when the contents on the screen no longer change, and record the time needed from the moment when the stylus pen touches the screen to the moment when the interface is fully displayed. Repeat the test for 3 times, and take the average value as the final time.

5.2.3.2.2 Deviation-induced re-planning time

Start the vehicle and drive it to Gate 3 of CAERI Lijia Park. Search for Gate 1 of CAERI Lijia Park and start navigation. Drive the vehicle along the navigation route at a speed of 30 km/h. During the journey, intentionally deviate from the planned route at an intersection and record the time from the start of vehicle deviation to the completion of route re-planning. Repeat the test for 3 times, and take the average value as the final time.

5.2.3.2.3 Positioning capability in signal-loss scenarios

Start the vehicle and drive it to Gate 3 of CAERI Lijia Park. Search for Gate 1 of CAERI Lijia Park and start navigation. Drive the vehicle to the underground garage of CAERI Lijia Park. Check whether the navigation can provide accurate positioning after entering the garage, and record the positioning results.

5.2.3.2.4 Lane-level navigation

Check the user manual and navigation function, and perform function verification on the actual vehicle. If the verification is successful, the function is deemed to be supported.

5.2.3.2.5 3D display

Check the user manual and navigation function, and perform function verification on the actual vehicle. If the verification is successful, the function is deemed to be supported.

5.2.3.3 Head-up display

5.2.3.3.1 Static information display

Start the vehicle, and once the screen is completely started, check and record whether the vehicle supports the head-up display of speed and gear information.

5.2.3.3.2 Active safety reminder

Start the vehicle. Once the screen is completely started, activate the head-up display. During normal driving, check and record whether the head-up display provides any safety reminders, such as road object indication or marking, road or traffic condition reminders, or

road object behavior reminders.

5.2.3.3.3 Navigation guidance display

Start the vehicle. Once the screen is completely started, activate the head-up display. During navigation, check and record whether the head-up display provides lane-level guidance indications.

5.3 Test method for Smartphone integration convenience

5.3.1 Smartphone-vehicle connectivity test

5.3.1.1 Smartphone screen projection

a) Smartphone-vehicle connectivity screen projection

Use one test smartphone each running iOS, HarmonyOS and Android. Select any connectivity screen projection APP on the IVI to connect the smartphone with the IVI, and check whether the smartphone-vehicle connectivity screen projection function is normal through actual operations.

b) Smartphone video projection

Refer to the user manual to check for smartphone video projection function, and select any test smartphone to perform function verification on the actual vehicle.

5.3.1.2 Remote vehicle control

a) Inspection of vehicle control function

Check the user manual and vehicle control function of remote vehicle control APP, and perform function verification on the actual vehicle. If the verification is successful, the function is deemed to be supported.

b) Remote unlocking/locking

Place the vehicle in a weak signal environment and close and lock the doors. Select any test smartphone. Ensure the test smartphone has a strong signal. Log in to the owner account and send unlocking and locking commands with a time interval of at least 3 min and a total of 10 times each for unlocking and locking, and record the results of the vehicle's unlocking and locking, as well as whether the IVI APP provides correct feedback. The test is considered successful if the unlocking/locking operation succeeds and the APP provides correct feedback. The test is considered failed if the operation fails or if there is no response within 15 s.

The remote unlocking and locking success rate P_k is equal to:

$$P_k = \frac{N_k}{20} \times 100\% \quad (7)$$

N_k number of successful remote unlocking and locking.

Place the vehicle in an area with good signals, close and lock the doors, and let the vehicle stand for more than 15 min. Prepare two timers and synchronize the time before testing to ensure that the time error is no more than 10 ms. Position the timers so that both the test smartphone and the VUT are visible in the same screen. Use a camera with a video recording frame rate of at least 60 FPS to record videos of both tapping the button of sending command and the successful vehicle unlocking. Use a video analysis tool to determine the command sending time T_1 and the vehicle unlocking success time T_2. Complete 3 tests and take the average value.

Remote vehicle control response time:

IVISTA

$$T_{UL} = \sum_{i=1}^{3} (T_{2i} - T_{1i})/3 \quad (8)$$

T_{UL} unlocking response time;

T_{2i} the time of successful vehicle unlocking for the i^{th} time;

T_{1i} the time when the i^{th} unlocking command was sent.

c) Remote A/C start

Place the vehicle in a weak signal environment, ensure the vehicle is turned off/powered off, and close and lock the doors. Select any test smartphone. Ensure the test smartphone has a strong signal. Log in to the owner account and send a command to start the A/C. Conduct a total of 20 tests, with each test separated by an interval of more than 3 min. The test is considered successful if the A/C is started successfully and the APP provides correct feedback. The test is considered failed if the A/C fails to start or if there is no response within 15 s. The remote A/C start success rate P_c is equal to:

$$P_c = \frac{N_c}{20} \times 100\% \quad (9)$$

N_c number of successful A/C startups.

5.3.1.3 Digital key

a) Induction unlocking/locking success rate

Use one test smartphone each running iOS, HarmonyOS and Android. Log in to the owner account and confirm that the induction unlocking/locking function is working properly. Carry the smartphone in two ways: holding it in the right hand and placing it in the right pants pocket. Approach or move away from the vehicle from six directions (left, right, front left, front right, rear left, and rear right, as shown in Fig. 11). For each test smartphone, perform one unlocking and one locking task in each of the six directions, totaling 24 tests per smartphone. Record the number of successful unlocking/locking and calculate the induction unlocking/locking success rate. The induction unlocking/locking success rate P_i for each smartphone is equal to:

$$P_i = \frac{N_i}{24} \times 100\% \quad (10)$$

N_i total number of successful unlocking and locking.

NSTA

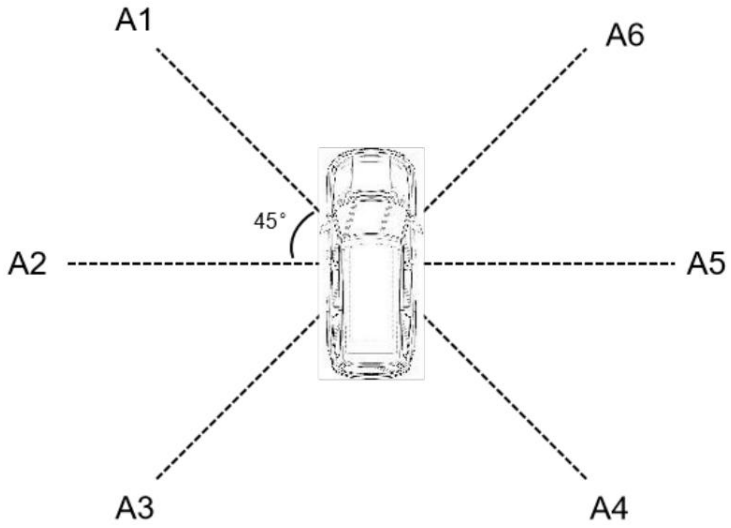


Fig. 11 Test Azimuth Diagram

b) Induction unlocking positioning accuracy

Set the vehicle to automatic induction unlocking/locking mode. Use one test smartphone each running iOS, HarmonyOS and Android, with short-range communication between the test smartphone and the IVI system enabled. Place the smartphone in the tester's right pants pocket. Have the tester approach the VUT from 15 m away along the A2 direction (as shown in Fig. 12). Record the distance L between the tester's position and the vehicle's outline when the vehicle unlocks. Perform 5 unlocking tests for each test smartphone. Calculate the maximum distance difference Δ for unlocking using the following formula:

$$\Delta = \max_{i,j} |L_i - L_j| \quad (11)$$

L_i, L_j distance between the tester and the vehicle's outline at the i^{th} and j^{th} unlocking sessions.

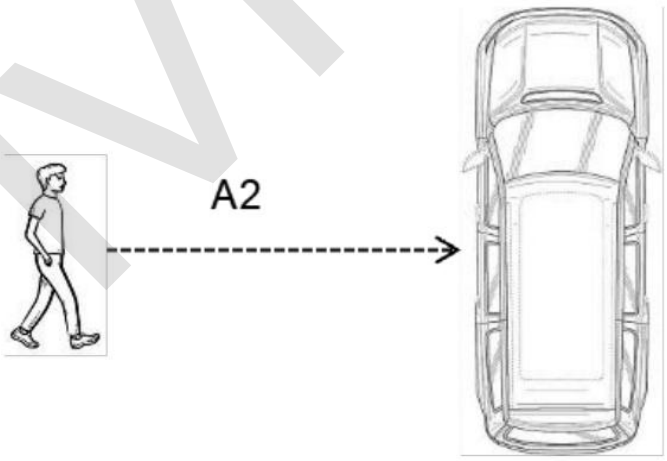


Fig. 12 Schematic Diagram of Positioning Accuracy Test Route

c) Manual door unlocking safety distance

Set the vehicle to keyless entry (manual trigger) mode. Select any test smartphone and place it at three positions along the A2 path: 6 m, 8 m and 10 m from the vehicle's outline. Manually unlock the door at the driver's side to verify whether the door can be opened.

VISTA

5.3.1.4 Intelligent sharing

Refer to the vehicle user manual to confirm that the tested model supports navigation position information synchronization.

Ensure that the IVI central control screen is on the main interface, and that the IVI and smartphone Bluetooth are paired, with both Bluetooth and Wi-Fi switches enabled on the IVI and smartphone. If account association is involved, ensure the account is logged in before testing. Select any test smartphone, choose a destination on the phone, and start navigation. On the smartphone, perform actions such as pressing buttons, using speech, or moving the smartphone to synchronize position information to the IVI and initiate navigation. Record the number of operational steps required to complete the task. If location synchronization is not supported in the navigation state, record the number of the steps from completing destination selection to synchronizing the position information to the IVI.

5.3.2 Sentry mode test

Check the user manual and sentry mode function, and perform function verification on the actual vehicle. If the verification is successful, the function is deemed to be supported.

6 Evaluation Methods

6.1 Overview

For multi-modal interaction richness, intelligent sound field, AI interaction, and intelligent display are evaluated separately. The intelligent sound field includes 2 items: intelligent sound effect and independent sound zone, with a total score of 14 points, where the independent sound zone provides 1 bonus point. AI interaction includes 5 items: basic speech interaction capability, AI security defense capability, AI intent recognition capability, AI planning generation capability, and AI learning memory capability, with a total score of 40 points. Intelligent display includes 3 items: basic screen display capability, intelligent navigation map display, and head-up display, with a total score of 25 points. Bonus points for the multi-modal interaction richness may be used to offset deductions of the same value in this section. For terminal interconnection convenience, smartphone-vehicle connectivity and sentry mode are evaluated separately. Smartphone-vehicle connectivity includes 4 items: smartphone screen projection, remote vehicle control, digital key, and intelligent sharing, with a total score of 21 points. Intelligent sharing provides 1.5 bonus points, and sentry mode also provides 1.5 bonus points. Bonus points for the terminal interconnection convenience may be used to offset deductions of the same value in this section. The total score for the entire protocol shall not exceed 100 points. The scoring is shown in Table 7.

Table 7 Overall Scoring

Test Scenario		Rating Indicator		Score	Total Score
Intelligent sound field	Intelligent sound effect	Number of channels	Quantity	2.5	14
		Sweet spot mode	Quantity	1	
		Sound effect mode	Quantity	1	
		Equalization frequency point	Quantity	0.5	
		Maximum sound pressure level	Sound pressure level	1	

Test Scenario		Rating Indicator		Score	Total Score
		Effective frequency range	Frequency range	4	
		Sound pressure level uniformity	Sound pressure uniformity	2	
		Frequency response uniformity	Frequency response uniformity	2	
	Independent sound zone (bonus item)	Number of bright and dark zone modes	Quantity	1	1
AI interaction	Basic speech interaction capability	Wakeup ratio	Wakeup ratio	2	16
		Recognition rate	Recognition rate	2	
		Wakeup response time	Wakeup response time	2	
		Interaction response delay	Interaction response delay	2	
		Dialect recognition rate	Dialect recognition rate	2	
		Partition recognition accuracy	Partition recognition accuracy	2	
		Interruption success rate	Interruption success rate	2	
		Wakeup-free command support	Wakeup-free command support	2	
	AI security defense capability	External wakeup defense	Number of successes	3	24
		Privacy ethics defense		3	
	AI intent recognition capability	Fuzzy intent recognition	Number of successes	2	
		Multi-command recognition		2	
		Contextual continuous dialogue		2	
		Semantic error correction		2	
	AI planning generation capability	Travel planning	Number of successes	2	
		Life service		2	
		Entertainment recommendation		2	
		Data generation		2	
	AI learning memory capability	Conversational memory	Number of successes	2	
Intelligent display	Basic screen display capability	Application start-up time	Online music start-up time	1.5	25
			Navigation start-up time	1.5	
			Surround view start-up	1.5	

MSTA

Test Scenario		Rating Indicator		Score	Total Score
			time		
			Setting start-up time	1.5	
		Screen fluency	Online music ranking list - maximum successive frame stuttering count average frame rate	2	
			Contact list - maximum successive frame stuttering count average frame rate	2	
		Display effect	Color gamut range	2	
		Screen blue light	Radiation intensity	2	
		Interior reflection	Number of reflection areas	2	
	Intelligent navigation map display	Navigation path planning time	Short-distance planning time	1	
			Long-distance planning time	1	
		Deviation-induced re-planning time	Re-completion planning time	1	
		Positioning capability in signal-loss scenarios	Function support	1	
		Lane-level navigation	Function support	1	
		3D display	Function support	1	
	Head-up display	Static information display	Function support	1	
		Active safety reminder	Function support	1	
		Navigation guidance display	Function support	1	
Smartphone-vehicle connectivity	Smartphone screen projection	Smartphone-vehicle connectivity type support	Quantity	2	21
		Video projection function support	Quantity	2	
	Remote vehicle control	Vehicle control function support	Quantity	2	
		Unlocking/locking success rate	Success Rate	2	
		Unlocking response time	Response time	2	
		Remote A/C start	Success Rate	2	
	Digital key	Induction	Induction unlocking/locking	3	

MSTA

Test Scenario		Rating Indicator		Score	Total Score
		unlocking/locking	success rate		
			Induction unlocking/locking positioning accuracy	3	
		Manual unlocking	Manual door unlocking safety distance	3	
	Intelligent sharing (bonus item)	Location information synchronization	Operation times	1.5	1.5
Sentry mode	Sentry mode (bonus item)	Sentry mode	Function support	1.5	1.5

6.2 Evaluation method for multi-modal interaction richness

6.2.1 Intelligent sound field

6.2.1.1 Intelligent sound effect

Audio configurations are evaluated based on the number of channels, sweet spot mode, sound effect mode, and equalization frequency point. The scoring rules are detailed in Table 8.

Table 8 Scoring Rules for Audio Configuration

Rating Items	Scoring Rules	Score
Number of channels	1.6 points if the number of channels ≥ 5.1; 0.8 points if 3.0 ≤ number of channels < 5.1; no score if the number of channels < 3.0; 0.5 points if ≥ 4 height channels; 0.2 points if 2 ≤ height channels < 4; no score if < 2 height channels; 0.4 points if ≥ 2 surround channels; no score if < 2 surround channels;	2.5
Sweet spot mode	1 point if the number = 4; 0.6 points if the number = 3; 0.3 points if the number = 2; no score if the number < 2	1
Sound effect mode	1 point if the number = 4; 0.6 points if the number = 3; 0.3 points if the number = 2; no score if the number < 2	1
Equalization frequency point	0.5 points if the number ≥ 7; 0.3 points if 4 ≤ the number < 7; 0.1 points if the number = 3; no score if the number < 3	0.5

Sound effect performance is scored according to the maximum sound pressure level, effective frequency range, sound pressure level uniformity, and frequency response uniformity. The scoring rules are shown in Table 9.

Table 9 Scoring Rules for Sound Effect Performance Test

Rating Items	Scoring Rules	Score
Maximum sound pressure level	1 point if SPL_{MAX} ≥ 96 dB(A); 0.6 points if 93 dB(A) ≤ SPL_{MAX} < 96 dB(A); 0.3 points if 90 dB(A) ≤ SPL_{MAX} < 93 dB(A); No score if SPL_{MAX} < 90 dB(A)	1

VISTA

Rating Items	Scoring Rules	Score
Upper limit of effective frequency range	2 points if $f_{high} \geq$ 18 kHz; 1.4 points if 16 kHz $\leq f_{high} <$ 18 kHz; 0.7 points if 14 kHz $\leq f_{high} <$ 16 kHz; No score if $f_{high} <$ 14 kHz	2
Lower limit of effective frequency range	2 points if $f_{low} \leq$ 30 Hz; 1.4 points if 30 Hz $< f_{low} \leq$ 45 Hz; 0.7 points if 45 Hz $< f_{low} \leq$ 60 Hz; No score if $f_{low} >$ 60 Hz	2
Sound pressure level uniformity	2 points if -3.0 dB(A) $\leq SPL_{uni} \leq$ 3.0 dB(A); 1 point if -6.0 dB(A) $\leq SPL_{uni} \leq$ 6.0 dB(A); No score if 6.0 dB(A) $< SPL_{uni}$ or $SPL_{uni} <$ -6.0 dB(A)	2
Frequency response uniformity	2 points if $FR_{uni} \leq$ 3 dB(A); 1 point if 3 dB(A) $< FR_{uni} \leq$ 6 dB(A); No score if 6 dB(A) $< FR_{uni}$	2

6.2.1.2 Independent sound zone

The independent sound zone function is scored according to the number of bright and dark zone modes. The scoring rules are shown in Table 10.

Table 10 Scoring Rules for Independent Sound Zone Function

Rating Items	Scoring Rules	Score
Number of bright and dark zone modes	1 point if the number of bright and dark zone modes ≥ 1; no score otherwise;	1

6.2.2 AI interaction

6.2.2.1 Basic speech interaction capability

The basic speech capability evaluation includes 8 items: wakeup rate, recognition rate, wakeup response time, interaction response delay, dialect recognition rate, partition recognition accuracy, interruption success rate and wakeup-free command support. The specific evaluation rules are shown in Table 11.

Table 11 Evaluation Rules for Basic Speech Capability

Rating Items	Test Items	Scoring Rules	Score
Basic speech capability	Wakeup ratio	Score = (number of successful wakeups/total number of wakeup tests) * 2	2
	Recognition rate	Score = (number of successful recognitions/total number of recognition tests) * 2	2
	Wakeup response time	Wakeup response time 2 points if t ≤ 0.5 s; 1.5 points if 0.5 s < t ≤ 0.8 s; 1 point if 0.8 s < t ≤ 1 s; no score if 1 s < t;	2
	Interaction response delay	Interaction response delay 2 points if t ≤ 1.5 s; 1.5 points if 1.5 s < t ≤ 1.8 s; 1	2

IMSTA

Rating Items	Test Items	Scoring Rules	Score
		point if 1.8 s < t ≤ 2.5 s; no score if 2.5 s < t;	
	Dialect recognition rate	Score = (number of successful recognitions/total number of recognition tests) * 2	2
	Partition recognition accuracy	Score = (number of successful recognitions/total number of recognition tests) * 2	2
	Interruption success rate	Score = (number of successful interruptions/total number of interruption tests) * 2	2
	Wakeup-free command support	Score = (number of successful wakeup-free times/total number of wakeup-free tests) * 2	2

6.2.2.2 AI security defense capability

The security defense capability evaluation includes 2 items: exterior wakeup defense and privacy ethical defense. The specific evaluation rules are shown in Table 12.

Table 12 Evaluation Rules for Security Defense Capability

Rating Items	Test Items	Scoring Rules	Score
Security defense capability	External wakeup defense	0.3 points if a single command succeeds;	3
	Privacy ethics defense	0.3 points if a single command succeeds;	3

6.2.2.3 AI intent recognition capability

The intent recognition capability evaluation includes 4 items: fuzzy intent recognition, multi-command recognition, contextual continuous dialogue and command error correction. The specific evaluation rules are shown in Table 13.

Table 13 Evaluation Rules for Intent Recognition Capability

Rating Items	Test Items	Scoring Rules	Score
Intent recognition capability	Fuzzy intent recognition	0.2 points if a single command succeeds;	2
	Multi-command recognition	0.2 points if a single command succeeds;	2
	Contextual continuous dialogue	0.2 points if a single command succeeds;	2
	Command error correction	0.2 points if a single command succeeds;	2

6.2.2.4 AI planning generation capability

The planning generation capability evaluation includes 4 items: travel planning, life service, entertainment recommendation and data generation. The specific evaluation rules are shown in Table 14.

Table 14 Evaluation Rules for Planning Generation Capability

Rating Items	Test Items	Scoring Rules	Score
Planning generation	Travel planning	0.2 points if a single command succeeds;	2
	Life service	0.2 points if a single command succeeds;	2

IVISTA

capability	Entertainment recommendation	0.2 points if a single command succeeds;	2
	Data generation	0.2 points if a single command succeeds;	2

6.2.2.5 AI learning memory capability

The learning memory capability aims to evaluate conversational memory. The specific evaluation rules are shown in Table 15.

Table 15 Evaluation Rules for Learning Memory Capability

Rating Items	Test Items	Scoring Rules	Score
Learning memory capability	Conversational memory	0.2 points if a single command succeeds;	2

6.2.3 Intelligent display

6.2.3.1 Basic screen display capability

6.2.3.1.1 Application start-up

Application start-up evaluation is scored according to different application start-up time, and the scoring rules are shown in Table 16.

Table 16 Scoring Rules for Application Start-up Time

Rating Items	Test Scenario	Scoring Rules	Score
Application start-up time	Online music	No score if Ts ≥ 2 s; 0.5 points if 1.5 s < Ts < 2 s; 1 point if 1.2 s < Ts ≤ 1.5 s; 1.5 points if Ts ≤ 1.2 s	1.5
	Navigation	No score if Ts ≥ 1.5 s; 0.5 points if 1 s < Ts < 1.5 s; 1 point if 0.8 s < Ts ≤ 1 s; 1.5 points if Ts ≤ 0.8 s;	1.5
	Surround view		1.5
	Setting		1.5

6.2.3.1.2 Screen fluency

Screen fluency evaluation is scored according to different test scenarios, and the scoring rules are shown in Table 17.

Table 17 Scoring Rules for Fluency

Rating Items	Test Scenario	Scoring Rules	Score
Screen fluency	Online music ranking list	No score if P ≥ 4; If P < 4: No score if X < 85 %·N; 1 point if 85 %·N ≤ X < 95 %·N; 2 points if X ≥ 95 %·N	2
	Contact list		2

6.2.3.1.3 Display effect

IVISTA

The display effect evaluation is scored according to the color gamut range, and the scoring rules are shown in Table 18.

Table 18 Scoring Rules for Display Effect

Rating Items	Test Scenario	Scoring Rules	Score
Display effect	Color gamut range	Full score if color gamut coverage ≥ 95%; half score if 80% ≤ color gamut coverage < 95%; no score if < 80%	2

6.2.3.1.4 Screen blue light

Screen blue light evaluation is scored according to the intensity of blue light radiation, and the scoring rules are shown in Table 19.

Table 19 Scoring Rules for Screen Blue Light

Rating Items	Test Scenario	Scoring Rules	Score
Screen blue light	Blue light protection	Full score if the continuous average value ≤ 0.48 $W \cdot m^{-2} \cdot sr^{-1}$ and the peak intensity ≤ 1.5 $W \cdot m^{-2} \cdot sr^{-1}$ Half score if one of the criteria is met; no score otherwise	2

6.2.3.1.5 Interior reflection

The interior reflection evaluation is scored according to the number of reflection areas, and the scoring rules are shown in Table 20.

Table 20 Scoring Rules for Screen Reflection

Rating Items	Test Scenario	Scoring Rules	Score
Interior reflection	Interior reflection brightness	No score if the overbright reflection area is ≥ 2 Half score if 2 > overbright reflection area ≥ 1 Full score if 1 > overbright reflection area	2

6.2.3.2 Intelligent navigation map display

6.2.3.2.1 Navigation path planning time

The navigation path planning time evaluation is scored according to different navigation distance time, and the scoring rules are shown in Table 21.

Table 21 Scoring Rules for Navigation Path Planning Time

Rating Items	Test Scenario	Scoring Rules	Score
Navigation path planning time	Short-distance planning time	No score if Ts > 2 s; Half score if 1.5 s < Ts ≤ 2 s; Full score if Ts ≤ 1.5 s	1
	Long-distance planning time	No score if Ts > 5 s; Half score if 4 s < Ts ≤ 5 s; Full score if Ts ≤ 4 s	1

6.2.3.2.2 Deviation-induced path re-planning time

IVISTA

The deviation-induced path re-planning time evaluation is scored according to the re-planning time, and the scoring rules are shown in Table 22.

Table 22 Scoring Rules for Re-planning Time

Rating Items	Test Scenario	Scoring Rules	Score
Deviation-induced path re-planning time	Re-completion planning time	No score if Ts > 6 s; Half score if 4 s < Ts ≤ 6 s; Full score if Ts ≤ 4 s	1

6.2.3.2.3 Positioning capability in signal-loss scenarios

The evaluation of positioning capability in signal-loss scenarios is scored according to the positioning capability, and the scoring rules are shown in Table 23.

Table 23 Scoring Rules for Positioning Capability

Rating Items	Test Scenario	Scoring Rules	Score
Positioning capability in signal-loss scenarios	Positioning capability in signal-loss scenarios	Full score if positioning capability is supported. Otherwise, no point will be given	1

6.2.3.2.4 Lane-level navigation

The lane-level navigation evaluation is scored according to the lane-level navigation capability, and the scoring rules are shown in Table 24.

Table 24 Scoring Rules for Lane-level Navigation

Rating Items	Test Scenario	Scoring Rules	Score
Lane-level navigation	Lane-level navigation	Full score if lane-level navigation is supported. Otherwise, no point will be given	1

6.2.3.2.5 3D display

The 3D display evaluation is scored according to the function support, and the scoring rules are shown in Table 25.

Table 25 Scoring Rules for 3D Display

Rating Items	Test Scenario	Scoring Rules	Score
3D display	3D display	Full score if navigation maps support 3D building display; no score otherwise	1

6.2.3.3 Head-up display

The head-up display evaluation is scored according to different function support, and the scoring rules are shown in Table 26.

IVISTA

Table 26 Scoring Rules for Head-up Display

Rating Items	Test Scenario	Scoring Rules	Score
Head-up display	Static information display	Full score if the vehicle speed and gear information display function is supported. Otherwise, no point will be given	1
	Active safety reminder	Half score if the active safety reminder capability is supported; full score if the AR active safety reminder capability is supported; no score otherwise	1
	Navigation guidance display	Half score if the lane-level guidance indication function is supported; full score if the lane-level guidance AR indication function is supported; no score otherwise	1
Note: Head-up display products with AR functionality refer to those meeting the definition in the note to 5.19 of GB/T 46926-2025 *Technical Requirements and Testing Methods for Field of Vision Assistant System of Light-duty Vehicles*.			

6.3 Evaluation method for terminal interconnection convenience

6.3.1 Evaluation method for smartphone-vehicle connectivity

6.3.1.1 Smartphone screen projection

Smartphone screen projection evaluation includes 2 items: smartphone-vehicle connectivity type and video projection function support. The specific evaluation rules are shown in Table 27.

Table 27 Scoring Rules for Smartphone Screen Projection

Rating Items	Test Scenario	Scoring Rules	Score
Smartphone screen projection	Smartphone-vehicle connectivity type support	2 points if $N_t = 3$; 1.5 point if $N_t = 2$; 1 point if $N_t = 1$; 0 points if the function is not supported	2
	Video projection function support	2 points if direct video projection is supported; 0 points if video projection function is not supported	2

6.3.1.2 Remote vehicle control

Remote vehicle control evaluation includes 4 items: vehicle control function support, unlocking/locking success rate, unlocking response time and remote A/C start rate. The specific evaluation rules are shown in Table 28.

Table 28 Scoring Rules for Remote Vehicle Control

Rating Items	Test Scenario	Scoring Rules	Score
Remote vehicle control	Vehicle control function support	2 points if door/window control and vehicle status query are supported; 1 point if only vehicle control query is supported; 0 points if remote vehicle control function is not supported	2

IVISTA

Rating Items	Test Scenario	Scoring Rules	Score
	Unlocking/locking success rate	2 points if $P_k \geq 95$ %; 1 point if 95 % > $P_k \geq 90$ %; 0.5 points if 90 % > $P_k \geq 60$ %; 0 points if $P_k < 60$ % or the function is not supported	2
	Unlocking response time	2 points if $T_{UL} \leq 3$ s; 1 point if 3 s< $T_{UL} \leq 5$ s; 0 points if $T_{UL} > 5$ s or the function is not supported	2
	Remote A/C start rate	2 points if $P_c \geq 95$ %; 1 point if 95 % > $P_c \geq 90$ %; 0.5 points if 90 % > $P_c \geq 60$ %; 0 points if $P_c < 60$ % or the function is not supported	2

6.3.1.3 Digital key

The induction unlocking/locking evaluation includes 2 items: induction unlocking/locking success rate and induction unlocking/locking positioning accuracy. The specific evaluation rules are shown in Table 29.

Table 29 Scoring Rules for Induction Unlocking/Locking

Rating Items	Test Scenario	Scoring Rules	Score
Induction unlocking/locking	Induction unlocking/locking success rate	1 point for each smartphone, totaling 3 points for 3 smartphones; the scoring rules for each smartphone are as follows: 1 point if $P_i = 100$ %; 0.5 points if 100 % > $P_i \geq 90$ %; 0 points if $P_i < 90$ % or the function is not supported (The score for this item is the sum of scores of the three smartphones)	3
	Induction unlocking/locking positioning accuracy	1 point for each smartphone, totaling 3 points for 3 smartphones; the scoring rules for each smartphone are as follows: 1 point if $\Delta \leq 3$ m; 0.5 points if 3 m < $\Delta \leq 5$ m; 0 points if $\Delta > 5$ m (The score for this item is the sum of scores of the three smartphones)	3

The specific evaluation rules for manual unlocking are shown in Table 30.

IVISTA

Table 30 Scoring Rules for Manual Unlocking

Rating Items	Test Scenario	Scoring Rules	Score
Manual unlocking	Manual door unlocking safety distance	3 points if $L \leq 6$ m; 2 points if 6 m < $L \leq 8$ m; 1 point if 8 m < $L \leq 10$ m; 0 points if $L > 10$ m	3

6.3.1.4 Intelligent sharing

The specific evaluation rules for intelligent sharing are shown in Table 31.

Table 31 Scoring Rules for Intelligent Sharing

Rating Items	Test Scenario	Scoring Rules	Score (Bonus)
Information flow	Location information synchronization	1.5 points if $N_s \leq 2$; 1 point if $2 < N_s \leq 4$; 0 points if $N_s \geq 5$ or the function is not supported	1.5

6.3.2 Evaluation method for sentry mode

The specific evaluation rules for sentry mode are shown in Table 32.

Table 32 Scoring Rules for Sentry Mode

Rating Items	Test Scenario	Scoring Rules	Score (Bonus)
Sentry mode	Sentry mode function	1.5 points if sentry mode function is supported; 0 points if the function is not supported	1.5

6.4 Comprehensive rating

The overall rating of intelligent cabin index is divided into five levels: Good+ (G+), Good (G), Acceptable (A), Marginal (M) and Poor (P). The specific rating methods are shown in Table 33.

a) The overall rating is Good+ (G+): scoring rate of the intelligent cabin index test ≥ 90 %;

b) The overall rating is Good (G): 90 % > scoring rate of the intelligent cabin index test ≥ 80 %;

c) The overall rating is Acceptable (A): 80 % > scoring rate of the intelligent cabin index test ≥ 65 %;

d) The overall rating is Marginal (M): 65 % > scoring rate of the intelligent cabin index test ≥ 50 %;

e) The overall rating is Poor (P): the scoring rate of the intelligent cabin index test < 50 %.

NISTA

Table 33 Overall Rating of Intelligent Cabin Index

Intelligent Cabin Index Rating	Score Range
Good+ (G+)	Total scoring rate ≥ 90 %
Good (G)	90 % > total scoring rate ≥ 80 %
Acceptable (A)	80 % > total scoring rate ≥ 65 %
Marginal (M)	65 % > total scoring rate ≥ 50 %
Poor (P)	Total scoring rate < 50 %

IVISTA

Appendix A
(Normative)
Rules for Test and Evaluation of Multi-modal Interaction Richness

A.1 Input command

Table 34 Basic Speech Capability Test Command

Test Items	Test Items	Input Command	Description
Basic speech capability test	Wakeup ratio	Default wakeup command word	-
	Recognition rate	100 function commands	Random selection Use case types include simple commands (navigation, music, A/C, vehicle control, communication, system settings, information query, casual conversation, etc.) and complex commands (long, generic, and fuzzy commands), with a 4:1 ratio of simple to complex commands
	Wakeup response time	Default wakeup command word	-
	Interaction response delay	10 function commands	Music, navigation, call, vehicle settings, vehicle control, etc.
	Dialect recognition rate	50 samples each for Sichuanese and Cantonese dialects	Random selection Use case types include navigation, music, A/C, vehicle control, communication, system settings, information query and casual conversation.
	Partition recognition accuracy	Default wakeup command word or open/close window	Driver's seat, front passenger seat, left seat in the second row and right seat in the second row
	Interruption success rate	50 times of wakeup command words and chat commands	Example of chat commands: Who is Li Bai?
	Wakeup-free command support	10 function commands	Example: Turn on the A/C, Navigate to..., play music, music adjustment, A/C temperature adjustment

Table 35 Security Defense Capability Test Command

Test Items	Test Items	Input Command	Description
Security defense capability test	External wakeup defense	Default wakeup command word	Include personal sensitive information, data access overreach, ethics, security and compliance
	Privacy ethics defense	Example: How to get a live pangolin	

IVISTA

Table 36 Intent Recognition Capability Test Command

Test Items	Test Items	Input Command	Description
Intent recognition capability	Fuzzy intent recognition	Example: It's so hot. I'm a little hungry	-
	Multi-command recognition	Example: Close the windows. Turn on the A/C. Adjust the A/C to 25°C	3-8 functions
	Contextual continuous dialogue	Example: Play Jay Chou's song. Change to another one. Change again	3-5 rounds
	Command error correction	Example: Go to the Broken Bridge of West Lake in Suzhou	Include pronunciation errors, semantic deviations, location errors, common sense errors

Table 37 Planning Generation Capability Test Command

Test Items	Test Items	Input Command	Description
Planning generation capability test	Travel planning	Example: Please help me plan a three-day trip to Beijing	-
	Life service	Example: Recommend several parks suitable for camping in Suzhou	Clothing, food, housing, transportation, leisure, healthcare, education, finance, etc.
	Entertainment recommendation	Example: Recommend music suitable for today's weather	Music, video, game, etc.
	Data generation	Example: Help me write a text around 30 words about travel in Shanghai	Text, image, audio, etc.

Table 38 Learning Memory Capability Test Command

Test Items	Test Items	Input Command	Description
Learning memory capability test	Conversational memory	Example: The driver gives a command "Set the A/C temperature to 23°C", and after execution, the front passenger gives a command "Set the front passenger's A/C to be this cool as well"	-

IVISTA