



**ResearchInChina**  
[www.researchinchina.com](http://www.researchinchina.com)

**China                      Automotive  
Multimodal              Interaction  
Development            Research  
Report, 2023**

Nov. 2023

# Single-modal interaction develops in depth

China Automotive Multimodal Interaction Development Research Report, 2023 released by ResearchInChina combs through the interaction modes of mainstream cockpits, the application of interaction modes in key vehicle models launched in 2023, the cockpit interaction solutions of suppliers, and the multimodal interaction fusion trends.

By sorting out the interaction modes and functions of new models rolled out in the previous year, it can be seen that active, anthropomorphic and natural interaction has become the main trend. In terms of interaction mode, in single-modal interaction, the control scope of mainstream interactions such as touch and voice has expanded from inside to outside cars, and the application cases of novel interactions like fingerprint and electromyography in cars have begun to increase; in multimodal fusion interaction, multiple fusion interactions, for example, voice + head posture/face/lip language, and face + emotion/smell, are being available to cars, aiming to create more active and natural human-vehicle interaction.

## 1. Single-modal interaction develops in depth.

- Haptic interaction: cockpits more tend to have large and multiple screens. The wider application of smart surface materials in cockpits also allows for extension of the haptic sensing scope to doors, windows, seats and other components, and haptic feedback technology is gradually introduced;
  - Voice interaction: enabled by large AI models, the voice interaction function becomes more intelligent and emotional. The introduction of lip movement recognition, voiceprint recognition and other technologies into cars brings higher accuracy of voice interaction and expands the control scope from inside to outside cars;
  - Visual interaction: the scope of face/gesture recognition based on visual technology begins to expand to body recognition, including head posture, arm movements, and body actions, etc.;
- \* Olfactory interaction: the olfactory interaction function, which was originally often used to purify the air and remove odors, can now enable cockpit sterilization and disinfection, and supports the linkage of the fragrance system with cockpit scenes/seasons.

# Car control by voice extends from inside to outside cars

## Case 1: car control by voice extends from inside to outside cars.

Typical models: Changan Nevo A07, Jiyue 01

Typical functions: voice outside the car to control doors, windows, parking assist, etc.

- Changan Nevo A07 adopts iFlytek's latest technology XTTS 4.0. The voice of the car voice assistant is more natural and anthropomorphic, and can express multiple emotions such as happiness, regret, and confusion. It supports saying towards the outside of the car (the content can be user-defined). In addition, the trunk, windows, music, air conditioning, pull-out/parking and other functions can also be controlled by voice outside the car.

\* Equipped with "SIMO" voice assistant, Jiyue 01 supports fully offline voice control in all zones, and allows for online voice interaction in the full process with weak network or without network. It enables recognition in 500 milliseconds and response in 700 milliseconds. Outside the car, the voiceprint recognition technology allows the driver and passengers to voice to operate air conditioning audio, lights, windows, doors, rear tailgate, charging cover and other functions, and supports voice parking outside the car.



Source: Jiyue

# Voiceprint recognition finds wider application

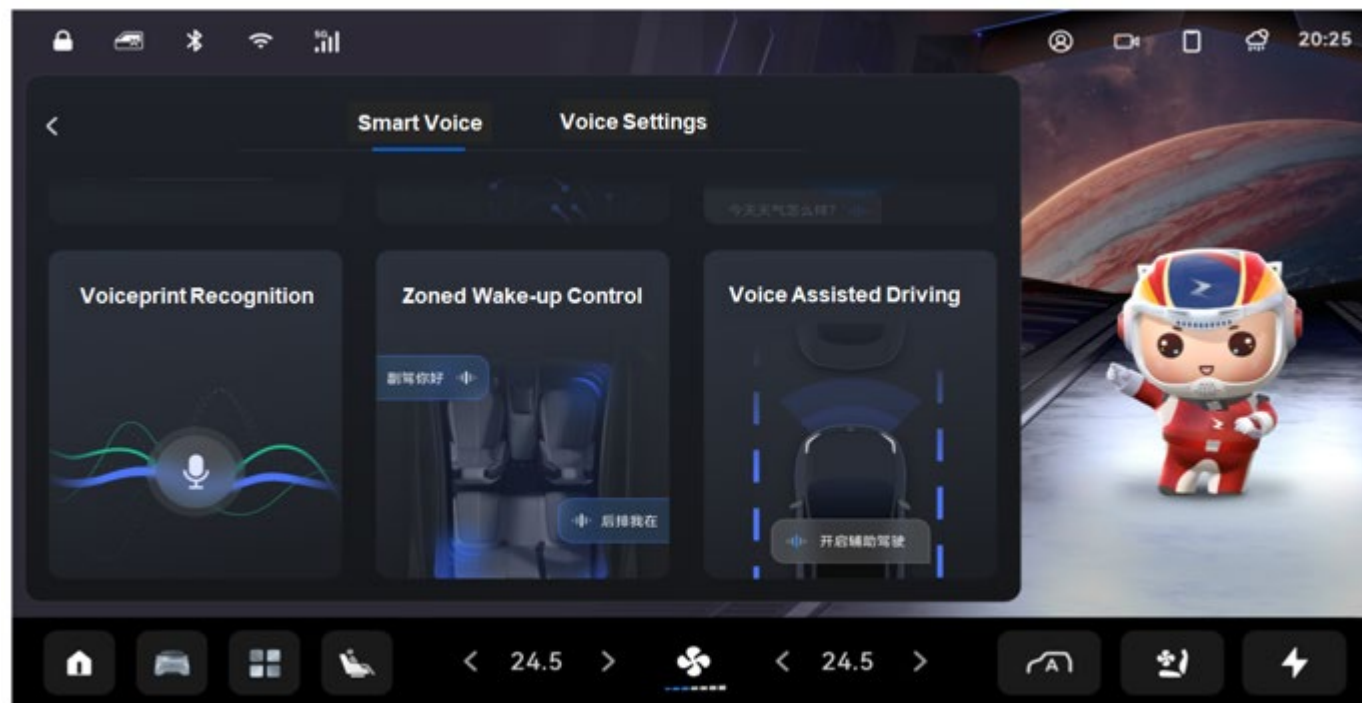
## Case 2: voiceprint recognition finds wider application.

Typical models: Li L7, Hycan A06/V09

Typical functions: identify drivers and passengers to provide targeted services

\* All Li Auto's L series models support voiceprint recognition function. After passengers register their voiceprints, "Lixiang Classmate" can identify who the passenger is, call the nicknames designated by different passengers, and perform vehicle control according to the positions of different passengers memorized via their voiceprint.

\* The voiceprint recognition VOICE ID of Hycan A06/V09 can clearly identify valid users and commands, and will become the entrance to HYCAN ID, allowing users to access rich smart ecosystems and use 100+ entertainment applications. Moreover based on voiceprint recognition technology, the system will actively block other disturbing sounds to improve the accuracy of recognition at the driver's seat.



Source: Hycan

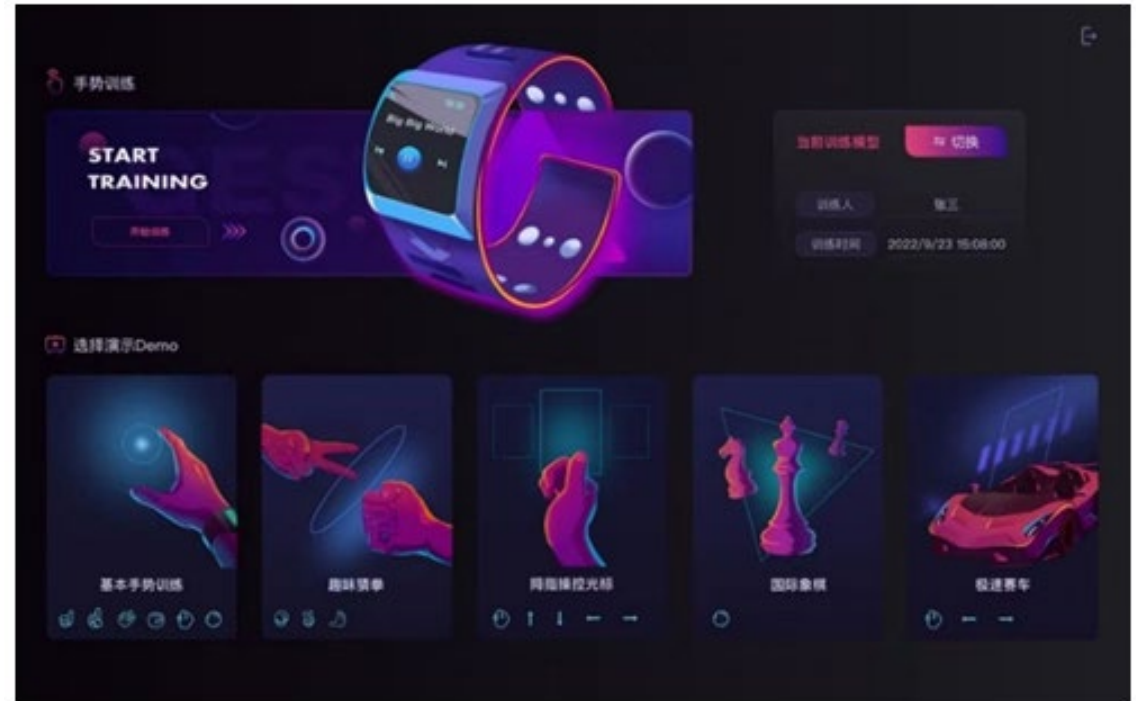
# Myoelectric interaction comes into commercial use in cars

## Case 3: myoelectric interaction comes into commercial use in cars.

Typical model: Voyah Passion

Typical function: micro-gesture control inside and outside the car

In April 2023, Voyah Passion and FlectoThink introduced a myoelectric interaction fusion solution enabled through a myoelectric bracelet. A multi-channel myoelectric sensor and a high-precision amplifier that are installed inside the bracelet can collect rich myoelectric signals in real time and generate algorithms, and transmit them to the computing terminal to generate a personalized AI gesture model, which is then integrated with Voyah's vehicle platforms. By connecting the bracelet with in-car Bluetooth, users can control the car with micro-gestures, including 60+ gestures to control the trunk and windows, for example. Additionally the bracelet can also be seamlessly connected to the car gaming system. The gesture recognition feature of the myoelectric bracelet allows users to control characters of games (e.g., Subway Surfers) more naturally and intuitively.



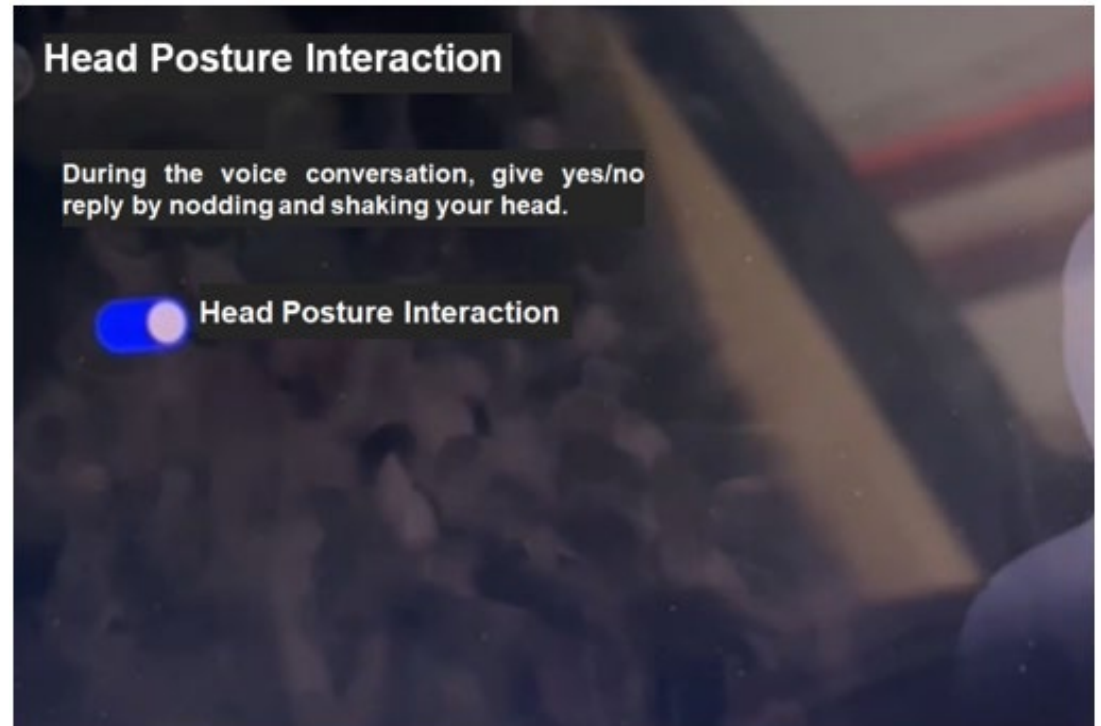
Source: FlectoThink

## 2. Multimodal fusion creates active interaction.

Currently multimodal fusion enabled by automakers includes but is not limited to voice + lip motion recognition, voice + face recognition, voice + gesture recognition, voice + head posture, face + emotion recognition, face + eye tracking, and fragrance + face + voice recognition. Wherein multimodal voice interaction is mainstream, and supports models mentioned above, like Changan Nevo A07, Jiyue 01, Li L7, and Hycan A06/V09.

***Case 1: voice + head posture interaction: WEY Blue Mountain DHT PHEV combines voice and head posture, offering a simple and intuitive interaction mode.***

When the driver engages in a voice conversation, the camera in the cockpit of Blue Mountain captures the driver's head movements, and allows the driver to give yes/no reply by nodding/shaking head. For example, when voicing to control navigation, the driver can select a planned route scheme by nodding/shaking head.



Source: Autohome

# Face + emotion recognition

**Case 2: face + emotion recognition: LIVAN 7 and ARCFOX Kaola among other models integrate emotion recognition technology into the face recognition function to provide active interaction and enhance interaction experience.**

The multimodal intelligent recognition Face-ID system of LIVAN 7 supports lip movement recognition and emotion recognition, and can remember the personalized settings of vehicle functions such as voice, seats, rearview mirrors, ambient light and trunk, that correspond to the associated accounts. It can also select the appropriate music according to the user's expression.

Directly facing the rear row, the camera on the B-pillar of ARCFOX Kaola can monitor a child in real time. For example, when the child smiles, a snapshot will be taken automatically and sent to the center console screen; when the child cries, soothing music will be automatically played and the surface of the smart seat will make a respiratory rhythm to calm him/her down. In addition, the camera can also be linked with the in-car radar to determine whether the child is asleep or not. If the child is asleep, the sleep mode will be automatically opened, the seat ventilation will be turned on, the air-conditioning temperature will be adjusted appropriately, and the audio and ambient lighting will be linked, producing a rhythmic effect.



Source: ARCFOX Kaola

# Foundation models and multimodal fusion will facilitate the introduction of AI Agent into cars

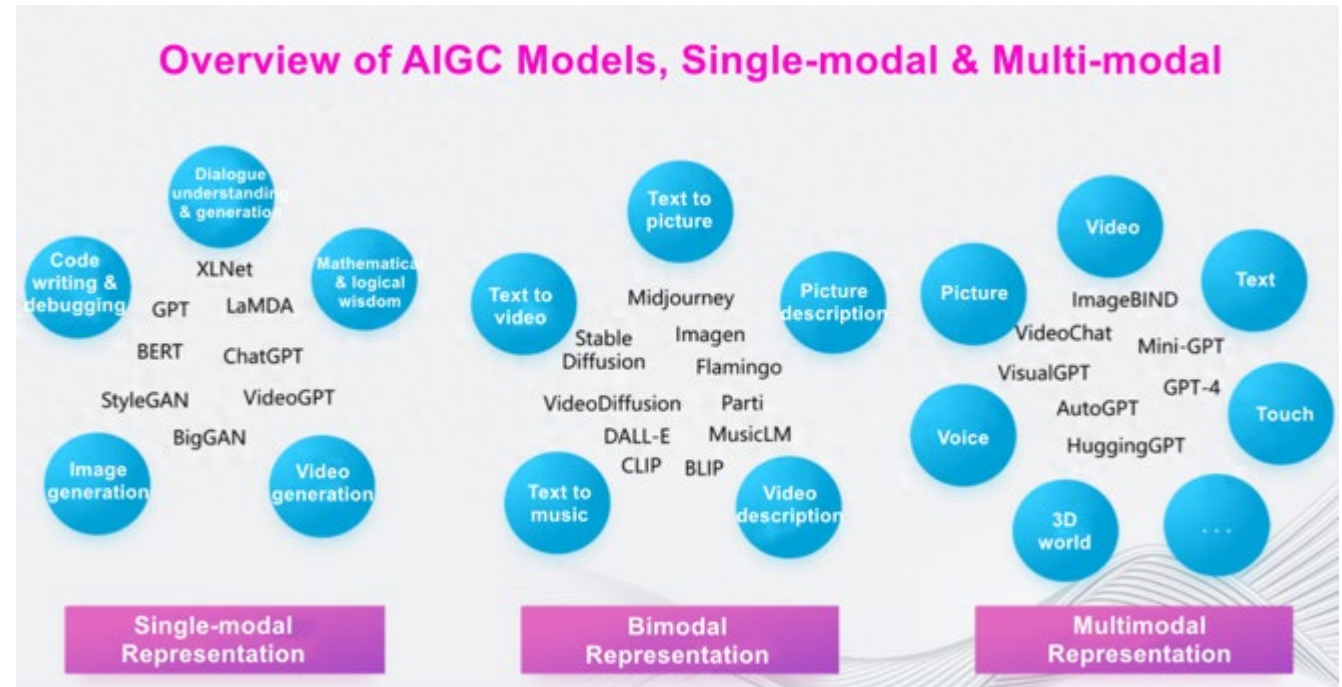
**Case 3: face + smell: NIO EC7, LIVAN 7 and other models realize the linkage between the driver monitoring system and the fragrance system to improve driving safety.**

When NIO EC7 detects the driver's tiredness, it will automatically release a refreshing fragrance to ensure driving safety; When the camera on the A-pillar of LIVAN 7 detects a drowsy driver, it will automatically release a refreshing fragrance and give a voice prompt.

### 3. Foundation models and multimodal fusion will facilitate the introduction of AI Agent into cars.

Large AI models are evolving from the single-modal to the multi-modal and multi-task fusion. Compared with the single-modal that can only process one type of data such as text, image and speech, the multimodal can process and understand multiple types of data, including vision, hearing and language, thus better understanding and generating complex information.

As multimodal foundation models continue to develop, their capabilities will also be significantly improved. This improvement gives AI Agent higher capabilities of perception and environment understanding to achieve more intelligent, automatic decisions and actions, and also creates new possibilities for its application in automotive, providing a broader prospect for future intelligent development.



Source: Sequoia Capital



# Spark Car Assistant

The Spark Cockpit OS developed by iFlytek based on the Spark Model supports multiple interaction modes such as voice, gesture, eye tracking and DMS/OMS. The Spark Car Assistant enables multi-intent recognition by deep understanding of the context, providing more natural human-machine interaction. The iFlytek Spark Model, first mounted on the model EXEED Sterra ES, will bring five new experiences: Vehicle Function Tutor, Empathy Partner, Knowledge Encyclopedia, Travel Planning Expert, and Physical Health Consultant.

AITO M9, to be launched in December 2023, has HarmonyOS 4 IVI system built in. Xiaoyi, the intelligent assistant in HarmonyOS 4, has been connected to Huawei Pangu Model, which includes natural language model, visual model, and multi-modal model. The combination of HarmonyOS 4 + Xiaoyi + Pangu Model further enhances ecosystem capabilities such as cooperation of devices, and AI scenarios, and provides diverse interaction modes, including voice recognition, gesture control, and touch control, using multimodal interaction technology.

## Spark Cockpit OS - Spark Car Assistant

Core technologies of voice, vision and foundation model enable natural, free and intelligent interactive experience

Free

Unlimited topic chat/streaming dialogue/  
multi-style/service and chat integration

Echo  
cancellation

Front-end  
noise reduction

DMS/OMS

Intelligent

Deep context understanding/multi-intention/  
homophone and approximant words error correction

Eye  
tracking

Gesture  
recognition

Labial  
sound fusion

Semantic  
understanding

Speech  
synthesis

Natural

Parallel multi-zone interaction/  
emotional interaction/full-time dialogue

Voice  
recognition

Voice  
wake-up

Voiceprint  
recognition

Foundation Models



# Table of Content (1)

## **1 Overview of Multimodal Interaction**

- 1.1 Definition of Multimodal Interaction
- 1.2 Multimodal Interaction Industry Chain
  - 1.2.1 Multimodal Interaction Industry Chain - Chip Vendors
  - 1.2.2 Multimodal Interaction Industry Chain - Algorithm Providers
  - 1.2.3 Multimodal Interaction Industry Chain - System Integrators
- 1.3 Multimodal Fusion Algorithms
  - 1.3.1 Speech Algorithm
  - 1.3.2 Vision Algorithm
- 1.4 Multimodal Interaction Policy Environment
  - 1.4.1 Policy and Regulation Environment
  - 1.4.2 Multimodal Interaction Laws and Regulations
  - 1.4.3 In-cabin Information Security Strategies of OEMs

## **2 Human-Computer Interaction Based on Touch**

- 2.1 Haptic Interaction Development Route
- 2.2 Highlights of Haptic Interaction of OEMs
- 2.3 Cockpit Display Trends
- 2.4 Development Trends of Smart Surface Materials
- 2.5 Haptic Feedback Mechanism

## **3 Human-Computer Interaction Based on Hearing**

- 3.1 Voice Function Development Route
- 3.2 Summary on Voice Functions of OEMs
- 3.3 Summary on OTA Updates on Voice Functions of OEMs
- 3.4 Development Trends of Voice Interaction Images
- 3.5 Application of Voiceprint Recognition in Car Models
- 3.6 Customization Trends of Voice Functions
- 3.7 Major Suppliers of Voice Functions

- 3.8 Voice Function Development Models of OEMs

## **4 Human-Computer Interaction Based on Vision**

- 4.1 Face Recognition
  - 4.1.1 Face Recognition Function Development Route
  - 4.1.2 Application of Face Recognition in Car Models
  - 4.1.3 Summary on Face Recognition Suppliers
- 4.2 Gesture Recognition
  - 4.2.1 Gesture Recognition Function Development Route
  - 4.2.2 Application of Gesture Recognition in Car Models
  - 4.2.3 Summary on Gesture Recognition Suppliers
- 4.3 Lip Movement Recognition
  - 4.3.1 Lip Movement Recognition Function Development Route
  - 4.3.2 Application of Lip Motion Recognition in Car Models
  - 4.3.3 Summary on Lip Movement Recognition Suppliers
- 4.4 Other Visual Interaction
  - 4.4.1 AR/VR Interaction Function Development Route
  - 4.4.2 Application of AR/VR Interaction in Car Models
  - 4.4.3 Summary on AR/VR Interaction Suppliers

## **5 Human-Computer Interaction Based on Smell**

- 5.1 Olfactory Interaction Function Development Route
- 5.2 Principle of Intelligent Fragrance System
- 5.3 Fragrance System Technology
- 5.4 Application of Olfactory Interaction in Car Models
- 5.5 Summary on Fragrance System Technologies of OEMs
- 5.6 Olfactory Interaction Design Trends
- 5.7 Summary on Olfactory Interaction Suppliers

# Table of Content (2)

## **6 Human-Computer Interaction Based on Biometrics**

### 6.1 Fingerprint Recognition

#### 6.1.1 Fingerprint Recognition Function Development Route

#### 6.1.2 Application of Fingerprint Recognition in Car Models

#### 6.1.3 Summary on Fingerprint Recognition Suppliers

### 6.2 Iris Recognition

#### 6.2.1 Iris Recognition Function Development Route

#### 6.2.2 Application of Iris Recognition in Car Models

#### 6.2.3 Application of Iris Recognition in AR/VR

#### 6.2.4 Summary on Iris Recognition Suppliers

### 6.3 Myoelectric Recognition

#### 6.3.1 Myoelectric Recognition Function Development Route

#### 6.3.2 Application of Myoelectric Recognition in Car Models

#### 6.3.3 Introduction to Myoelectric Recognition Equipment

#### 6.3.4 Summary on Myoelectric Recognition Suppliers

### 6.4 Vein Recognition

#### 6.4.1 Vein Recognition Function Development Route

#### 6.4.2 Application of Vein Recognition in Car Models

#### 6.4.3 Summary on Vein Recognition Suppliers

### 6.5 Heart Rate Recognition

#### 6.5.1 Heart Rate Recognition Function Development Route

#### 6.5.2 Heart Rate Recognition Technology

#### 6.5.3 Application of Heart Rate Recognition in Car Models

## **7 Multimodal Interaction Application by OEMs**

### 7.1 Emerging Carmakers

#### 7.1.1 Multimodal Interaction in Xpeng G6

#### 7.1.2 Multimodal Interaction in Li L7

#### 7.1.3 Multimodal Interaction in NIO EC7

#### 7.1.4 Multimodal Interaction in Neta GT

#### 7.1.5 Multimodal Interaction in HiPhi Y

#### 7.1.6 Multimodal Interaction in Hycan A06

#### 7.1.7 Multimodal Interaction in Hycan V09

#### 7.1.8 Multimodal Interaction in New AITO M7

#### 7.1.9 Multimodal Interaction in AITO M9

### 7.2 Conventional Chinese Independent Automakers

#### 7.2.1 Multimodal Interaction in Chery Cowin Kunlun

#### 7.2.2 Multimodal Interaction in WEY Blue Mountain DHT PHEV

#### 7.2.3 Multimodal Interaction in Hyper GT

#### 7.2.4 Multimodal Interaction in Trumpchi E9

#### 7.2.5 Multimodal Interaction in Voyah Passion

#### 7.2.6 Multimodal Interaction in Denza N7

#### 7.2.7 Multimodal Interaction in Frigate 07

#### 7.2.8 Multimodal Interaction in Changan Nevo A07

#### 7.2.9 Multimodal Interaction in Jiyue 01

#### 7.2.10 Multimodal Interaction in ARCFOX Kaola

#### 7.2.11 Multimodal Interaction in Deepal S7

#### 7.2.12 Multimodal Interaction in Galaxy L6

#### 7.2.13 Multimodal Interaction in Lynk & Co 08

#### 7.2.14 Multimodal Interaction in LIVAN 7

#### 7.2.15 Multimodal Interaction in ZEEKR X

#### 7.2.16 Multimodal Interaction in ZEEKR 009

#### 7.2.17 Multimodal Interaction in IM LS7

#### 7.2.18 Multimodal Interaction in GEOME G6

#### 7.2.3 Multimodal Interaction in Hyper GT

#### 7.2.4 Multimodal Interaction in Trumpchi E9

#### 7.2.5 Multimodal Interaction in Voyah Passion

#### 7.2.6 Multimodal Interaction in Denza N7

# Table of Content (3)

- 7.2.7 Multimodal Interaction in Frigate 07
- 7.2.8 Multimodal Interaction in Changan Nevo A07
- 7.2.9 Multimodal Interaction in Jiyue 01
- 7.2.10 Multimodal Interaction in ARCFOX Kaola
- 7.2.11 Multimodal Interaction in Deepal S7
- 7.2.12 Multimodal Interaction in Galaxy L6
- 7.2.13 Multimodal Interaction in Lynk & Co 08
- 7.2.14 Multimodal Interaction in LIVAN 7
- 7.2.15 Multimodal Interaction in ZEEKR X
- 7.2.16 Multimodal Interaction in ZEEKR 009
- 7.2.17 Multimodal Interaction in IM LS7
- 7.2.18 Multimodal Interaction in GEOME G6
- 7.3 Conventional Joint Venture Automakers
  - 7.3.1 Multimodal Interaction in Mercedes-Benz EQS AMG
  - 7.3.2 Multimodal Interaction in GAC Toyota bZ 4X
  - 7.3.3 Multimodal Interaction in FAW Toyota bZ 3
  - 7.3.4 Multimodal Interaction in Buick Electra E5
  - 7.3.5 Multimodal Interaction in 11th Generation GAC Honda Accord
  - 7.3.6 Multimodal Interaction in FAW Audi e-tron GT
  - 7.3.7 Multimodal Interaction in BMW XM
- 7.4 Concept Cars
  - 7.4.1 Multimodal Interaction in Audi A6 Avant e-tron
  - 7.4.2 Multimodal Interaction in BMW i Vision Dee
  - 7.4.3 Multimodal Interaction in RAM 1500 Revolution
  - 7.4.4 Multimodal Interaction in Peugeot Inception
  - 7.4.5 Multimodal Interaction in Yanfeng XiM23s

## 8 Multimodal Interaction Solutions of Suppliers

### 8.1 Aptiv

- 8.1.1 Profile
- 8.1.2 Intelligent Cockpit Platform
- 8.2 Cipia Vision
  - 8.2.1 Profile
  - 8.2.2 Multimodal Interaction Solution
- 8.3 Cerence
  - 8.3.1 Profile
  - 8.3.2 Cockpit Interaction Solution
  - 8.3.3 Multimodal Interaction Solution
  - 8.3.4 Product Development Route
- 8.4 Continental
  - 8.4.1 Profile
  - 8.4.2 Multimodal Product Layout
- 8.5 iFlytek
  - 8.5.1 Profile
  - 8.5.2 Spark Model + Intelligent Cockpit
  - 8.5.3 Multimodal Interaction Solution
  - 8.5.4 Multimodal Interaction Becomes the Key Direction of iFlytek Super Brain 2030 Plan
- 8.6 SenseTime
  - 8.6.1 Profile
  - 8.6.2 SenseAuto Intelligent Cockpit
  - 8.6.3 SenseAuto Foundation Model Empowers Cockpits
  - 8.6.4 SenseAuto Smart Car Solution
- 8.7 ADAYO
  - 8.7.1 Profile
  - 8.7.2 Intelligent Cockpit Product Layout
  - 8.7.3 Multimodal Interaction System
- 8.8 Desay SV

# Table of Content (4)

- 8.8.1 Profile
- 8.8.2 Intelligent Cockpit Solution
- 8.9 ArcSoft Technology
  - 8.9.1 Profile
  - 8.9.2 In-cabin Monitoring Solution
  - 8.9.3 Core Technology
- 8.10 AISpeech
  - 8.10.1 Profile
  - 8.10.2 Vehicle Software and Hardware Integrated Products
  - 8.10.3 Telematics Solution
  - 8.10.4 Multimodal Interaction Solution
  - 8.10.5 Large Language Model
- 8.11 Horizon Robotics
  - 8.11.1 Profile
  - 8.11.2 Multimodal Interaction Solution
  - 8.11.3 Multimodal Interaction Core Algorithm
  - 8.11.4 Vehicle Operating System - TogetherOS
  - 8.11.5 Product Development Model and Business Model
- 8.12 ThunderSoft
  - 8.12.1 Profile
  - 8.12.2 Intelligent Cockpit Solution
  - 8.12.3 Foundation Model Empowers Cockpits
  - 8.12.4 Vehicle Operating System
- 8.13 PATEO
  - 8.13.1 Profile
  - 8.13.2 Intelligent Cockpit Solution
  - 8.13.3 Application of Intelligent Cockpit Solution
- 8.14 Joyson Electronics
  - 8.14.1 Profile

- 8.14.2 Intelligent Cockpit Layout
- 8.14.3 Multimodal Interaction Layout
- 8.15 Huawei
  - 8.15.1 Profile
  - 8.15.2 Intelligent Cockpit Solution
  - 8.15.3 Harmony IVI System
- 8.16 Baidu
  - 8.16.1 Profile
  - 8.16.2 Intelligent Cockpit Solution
  - 8.16.3 Multimodal Interaction Solution
  - 8.16.4 Intelligent Cockpit + Ernie Model
- 8.17 Tencent
  - 8.17.1 Profile
  - 8.17.2 Intelligent Cockpit Solution
  - 8.17.3 Vehicle Voice Interaction
- 8.18 Banma Network
  - 8.18.1 Profile
  - 8.18.2 Intelligent Cockpit Solution
  - 8.18.3 Intelligent Cockpit Interaction Capabilities
- 8.19 MINIEYE
  - 8.19.1 Profile
  - 8.19.2 Intelligent Cockpit Solution
- 8.20 Hikvision
  - 8.20.1 Profile
  - 8.20.2 Intelligent Cockpit Solution
  - 8.20.3 Vehicle Intelligent Monitoring System

## **9 Multimodal Interaction Summary and Trends**

- 9.1 Multimodal Interaction Fusion Trends

# Table of Content (5)

- 9.1.1 Development of Intelligent Cockpit Interaction System
- 9.1.2 Development Trends of Single-modal Perception - Touch
- 9.1.3 Development Trends of Single-modal Perception - Hearing
- 9.1.4 Development Trends of Single-modal Perception - Vision
- 9.1.5 Development Trends of Single-modal Perception - Smell
- 9.1.6 Multimodal Interaction Fusion Trends
- 9.1.7 Multimodal Interaction Development Roadmap
- 9.2 Cockpit Computing Power Required by Multimodal Interaction
- 9.3 Large AI Models Required by Multimodal Interaction
- 9.4 Integration of Multimodal Interaction and Cockpit Hardware
  - 9.4.1 Multimodal Recognition and Hardware Interaction - Headlights
  - 9.4.2 Multimodal Recognition and Hardware Interaction - Ambient Light
  - 9.4.3 Multimodal Recognition and Hardware Interaction - AR/VR
- 9.5 Summary on Multimodal Interaction Features in Typical Car Models



## Beijing Headquarters

TEL: 13718845418

Email: [report@researchinchina.com](mailto:report@researchinchina.com)



## Chengdu Branch

TEL: 028-68738514

FAX: 028-86930659

Website: [ResearchInChina](http://ResearchInChina.com)

WeChat: Zuosiqiche

