



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

# Automotive Vision Algorithm Industry Research Report, 2023

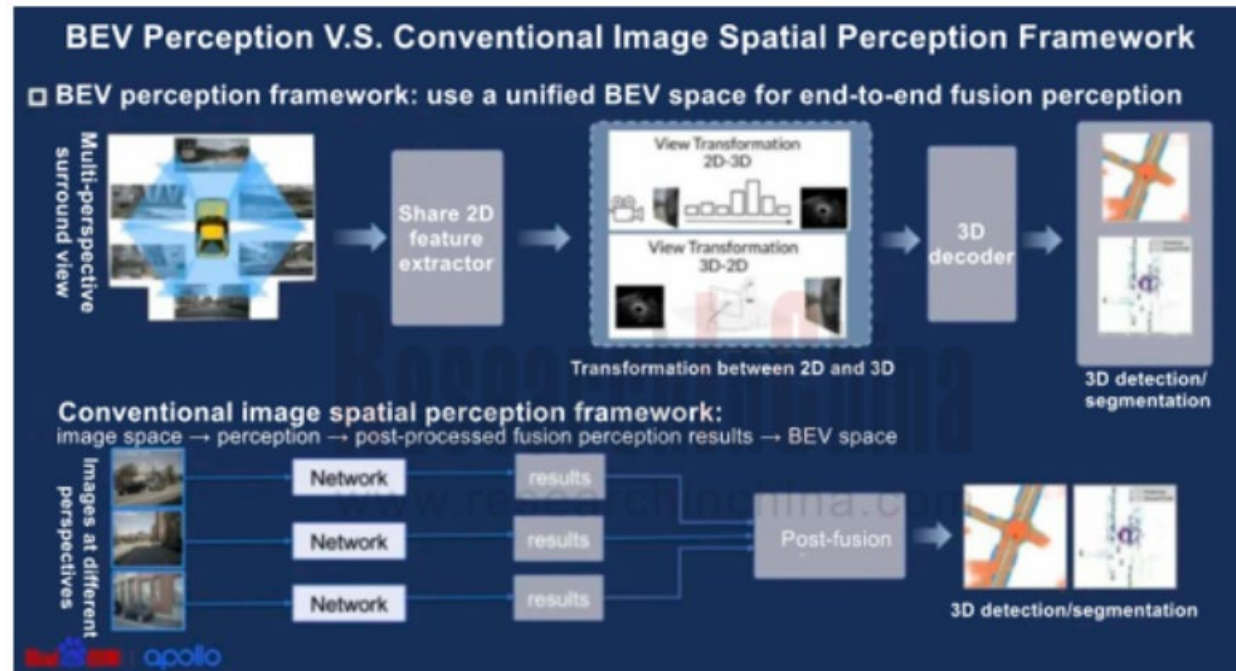
Feb. 2023

# Research on automotive vision algorithms: focusing on urban scenarios, BEV evolves into three technology routes.

## 1. What is BEV?

BEV (Bird's Eye View), also known as God's Eye View, is an end-to-end technology where the neural network converts image information from image space into BEV space.

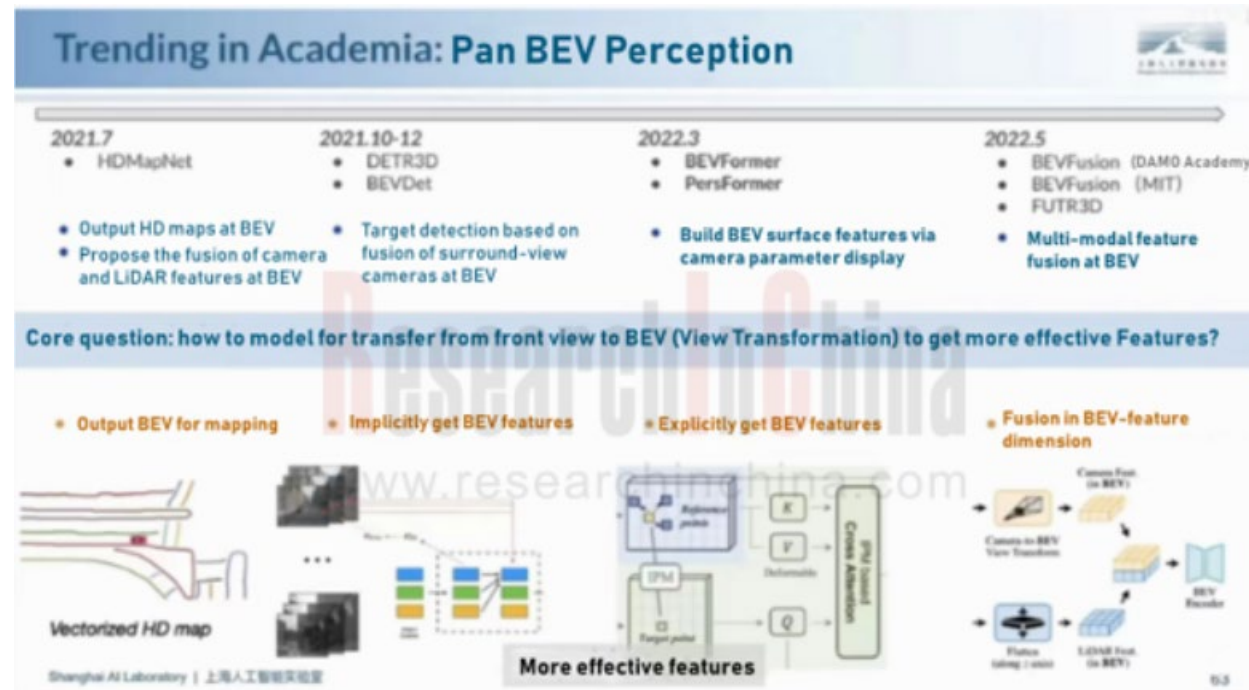
Compared with conventional image space perception, BEV perception can input data collected by multiple sensors into a unified space for processing, acting as an effective way to avoid error superposition, and also makes temporal fusion easier to form a 4D space.



Source: Baidu

# What is BEV?

BEV is not a new technology. In 2016, Baidu began to realize point cloud perception at the BEV; in 2021, Tesla's introduction of BEV draw widespread attention in the industry. There are BEV perception algorithms corresponding to different sensor input layers, basic tasks, and scenarios. Examples include BEVFormer algorithm only based on vision, and BEVFusion algorithm based on multi-modal fusion strategy.



Source: Shanghai Artificial Intelligence Laboratory



# Three technology routes of BEV perception algorithm

## 2. Three technology routes of BEV perception algorithm

In terms of implementation of BEV technology, the technology architecture of each player is roughly the same, but technical solutions they adopt are different. So far, there have been three major technology routes:

Vision-only BEV perception route in which the typical company is Tesla;

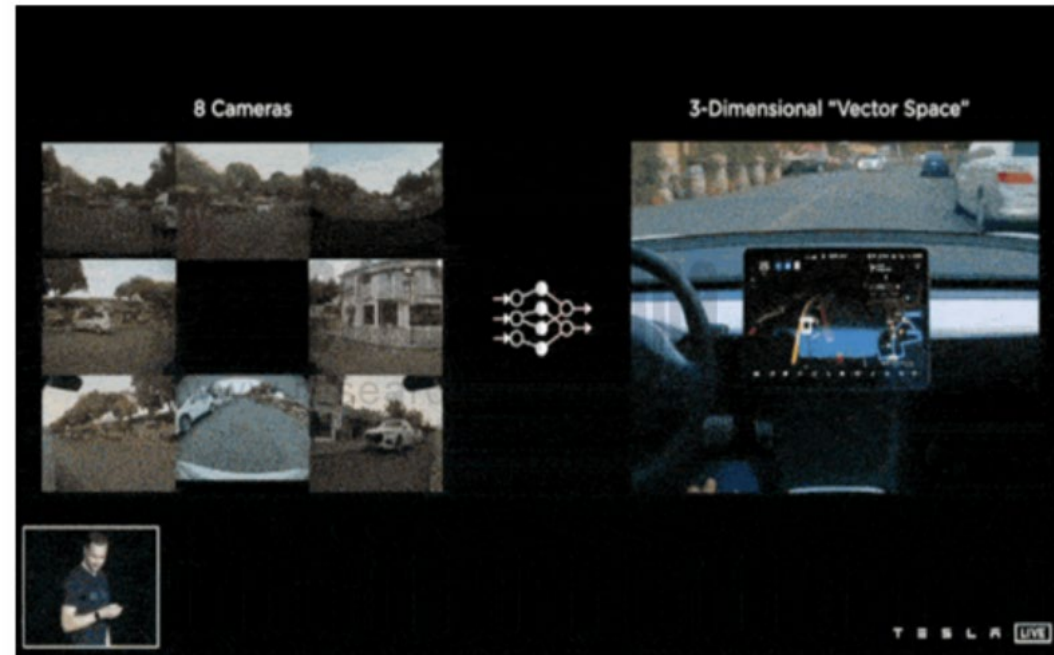
BEV fused perception route in which the typical company is Haomo.ai;

Vehicle-road integrated BEV perception route in which the typical company is Baidu.

**Vision-only BEV perception technology route:** Tesla is a representative company of this technology route. In 2021, it was the first one to use the pre-fusion BEV algorithm for directly transmitting the image perceived by cameras into the AI algorithm to generate a 3D space at a bird's-eye view, and output perception results in the space. This space incorporates dynamic information such as vehicles and pedestrians, and static information like lane lines, traffic signs, traffic lights and buildings, as well as the coordinate position, direction angle, distance, speed, and acceleration of each element.

Tesla uses the backbone network to extract features of each camera. It adopts the Transformer technology to convert multi-camera data from image space into BEV space. Transformer, a deep learning model based on the Attention mechanism, can deal with massive data-level learning tasks and accurately perceive and predict the depth of objects.

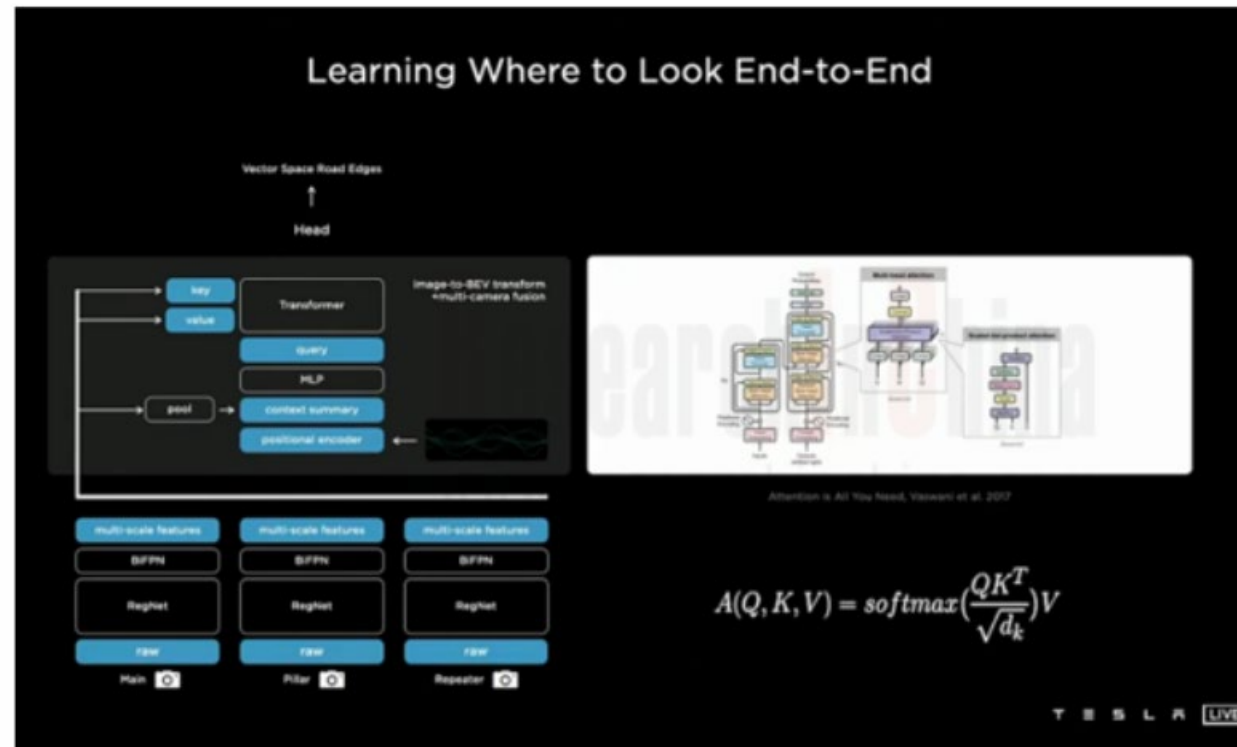
Tesla uses 8 cameras to collect data and output 3D vector space via neural network



Source: Tesla

# Vision-only BEV perception technology route

## Tesla introduces BEV 3D space transformation layer in the perception network architecture



Source: Tesla

# BEV fused perception technology route

**BEV fused perception technology route:** Haomo.ai is an autonomous driving company under Great Wall Motor. In 2022, it announced an urban NOH solution that underlines perception and neglects maps. The core technology comes from MANA (Snow Lake).

In the MANA perception architecture, Haomo.ai adopts BEV fused perception (visual Camera + LiDAR) technology. Using the self-developed Transformer algorithm, MANA not only completes the transformation of vision-only information into BEV, but also finishes the fusion of Camera and LiDAR feature data, that is, the fusion of cross-modal raw data.







Since its launch in late 2021, MANA has kept evolving. With Transformer-based perception algorithms, it has solved multiple road perception problems, such as lane line detection, obstacle detection, drivable area segmentation, traffic light detection & recognition, and traffic sign recognition.

In January 2023, MANA got further upgraded by introducing five major models to enable the transgenerational upgrade of the vehicle perception architecture and complete such tasks as common obstacle recognition, local road network and behavior prediction. The five models are: visual self-supervision model (automatic annotation of 4D Clip), 3D reconstruction model (low-cost solution to data distribution problems), multi-modal mutual supervision model (common obstacle recognition), dynamic environment model (using perception-focused technology for lower dependence on HD maps), and human-driving self-supervised cognition model (driving policy is more humane, safe and smooth).



Source: Haomo.ai

Iteration History of Perception Algorithms of Haomo MANA System

Time	2021	2022	2023
Algorithm improvement	<ul style="list-style-type: none"><li>Traffic light recognition</li><li>Transformer lane line recognition</li></ul>	<ul style="list-style-type: none"><li>Virtual real-time mapping on the BEV space via the transformer model that uses timing</li></ul>	<ul style="list-style-type: none"><li>End-to-end perception architecture to complete such tasks as common obstacle recognition, local road network, and behavior prediction</li></ul>
			
			

Source: Haomo.ai

# Vehicle-road integrated BEV perception technology route

**Vehicle-road integrated BEV perception technology route:** in January 2023, Baidu introduced UniBEV, a vehicle-road integrated solution which is the industry's first end-to-end vehicle-road integrated perception solution.

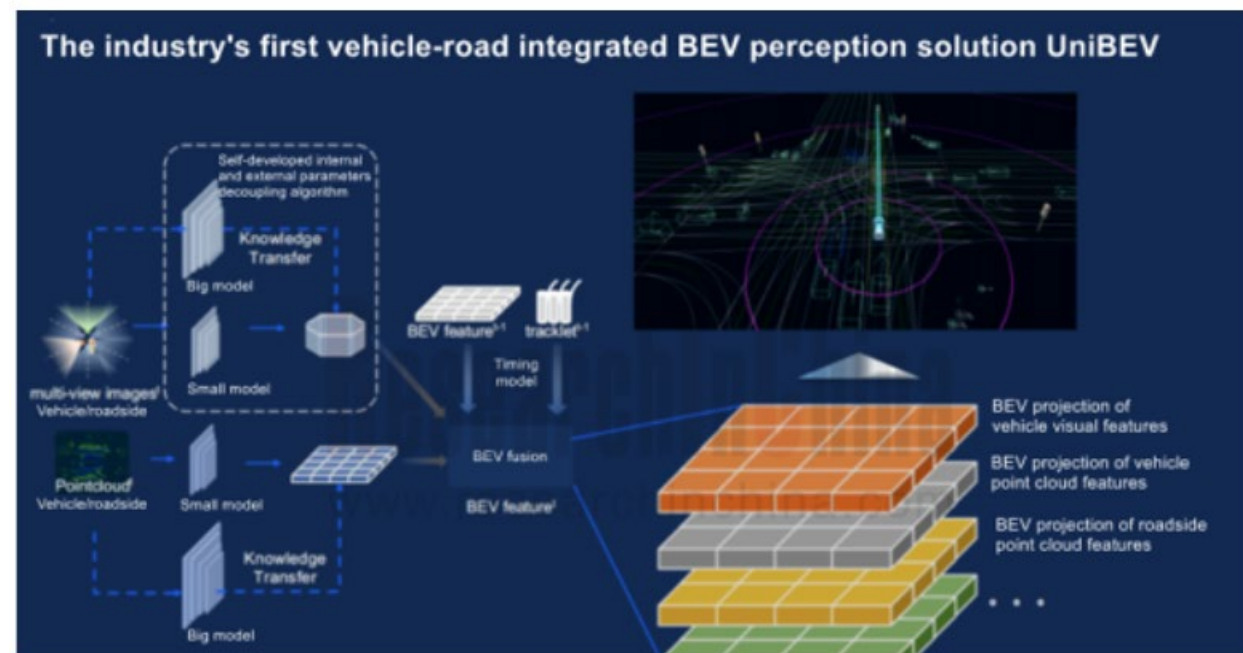
## Features:

Fusion of all vehicle and roadside data, covering online mapping with multiple vehicle cameras and sensors, dynamic obstacle perception, and multi-intersection multi-sensor fusion from the roadside perspective;

- Self-developed internal and external parameters decoupling algorithm, enabling UniBEV to project the sensors into a unified BEV space regardless of how they are positioned on the vehicle and at the roadside
- In the unified BEV space, it is easier for UniBEV to realize multi-modal, multi-view, and multi-temporal fusion of spatial-temporal features;
- The big data + big model + miniaturization technology closed-loop remains superior in dynamic and static perception tasks at the vehicle side and roadside.

Baidu's UniBEV solution will be applied to ANP3.0, its advanced intelligent driving product planned to be mass-produced and delivered in 2023. Currently, Baidu has started ANP3.0 generalization tests in Beijing, Shanghai, Guangzhou and Shenzhen.

Baidu ANP3.0 adopts the "vision-only + LiDAR" dual redundancy solution. In the R&D and testing phase, with the "BEV Surround View 3D Perception" technology, ANP3.0 has become an intelligent driving solution that enables multiple urban scenarios solely relying on vision. In the mass production stage, ANP3.0 will introduce LiDAR to realize multi-sensor fused perception to deal with more complex urban scenarios.



Source: Baidu



# BEV perception algorithm favors application of urban NOA

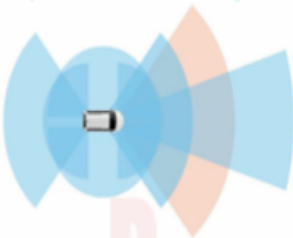
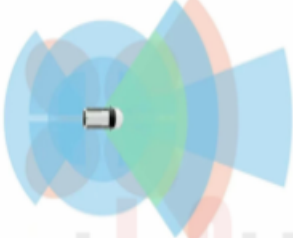
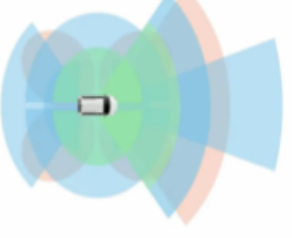
### 3. BEV perception algorithm favors application of urban NOA.

As vision algorithms evolve, BEV perception algorithms become the core technology for OEMs and autonomous driving companies such as Tesla, Xpeng, Great Wall Motor, ARCFOX, QCraft and Pony.ai, to develop urban scenarios.

**Xpeng Motors:** the new-generation perception architecture XNet can fuse the data collected by cameras before multi-frame timing, and output 4D dynamic information (e.g., vehicle speed and motion prediction) and 3D static information (e.g., lane line position) at the BEV.

**Pony.ai:** In January 2023, it announced the intelligent driving solution - Pony Shitu. The self-developed BEV perception algorithm, the key feature of the solution, can recognize various types of obstacles, lane lines and passable areas, minimize computing power requirements, and enable highway and urban NOA only using navigation maps.

Pony Shitu Intelligent Driving Solution

Solution	PonyClassic	PonyPro	PonyUltra
	Highway NOA + home AVP #超高性价比 #传感器数量精简 #畅行城市高架 Ultrahigh cost performance Minimum number of sensors Traveling freely on city viaducts	Urban NOA + AVP #仅用一颗激光雷达 #优于业界主流产品 Only need one LiDAR Superior mainstream products	Super Urban NOA + AVP #媲美L4 Robotaxi #极致舒适 #顶级体验 Comparable to L4 Robotaxi Extreme comfort Top configuration experience
			
	Camera x6 Radar x1	Camera x11 Radar x5 LiDAR x1	Camera x11 Radar x5 LiDAR x3 (Optional blind spot LiDARs)
Sensor Configuration	<ul style="list-style-type: none"><li>• 6 cameras</li><li>• 1 radar</li></ul>	<ul style="list-style-type: none"><li>• 11 cameras</li><li>• 5 radars</li><li>• 1 LiDAR</li></ul>	<ul style="list-style-type: none"><li>• 11 cameras</li><li>• 5 radars</li><li>• 3 LiDARs (adding 2 blind-spot LiDARs)</li></ul>
Computing Power	<ul style="list-style-type: none"><li>• ~50-100TOPS</li></ul>	<ul style="list-style-type: none"><li>• ~200TOPS</li></ul>	<ul style="list-style-type: none"><li>• ~500TOPS</li></ul>
SOP Plan	<ul style="list-style-type: none"><li>• 2023Q2</li></ul>	<ul style="list-style-type: none"><li>• 2023</li></ul>	—

Source: Pony.ai

# Table of Content (1)

## 1 Overview of Vision Algorithm

- 1.1 Vehicle Perception System Architecture
- 1.2 Vehicle Visual Sensors and Solutions
- 1.3 Vehicle Visual Perception Tasks
- 1.4 Computing Architecture and Algorithms of Exterior Visual Perception Systems
  - 1.4.1 Mono Camera Algorithm
  - 1.4.2 Stereo Camera Algorithm
  - 1.4.3 Surround View Camera Algorithm
- 1.5 Architecture and Algorithms of In-vehicle Visual DMS
  - 1.5.1 Visual DMS Solution
  - 1.5.2 Visual OMS Solution
- 1.6 BEV Perception Algorithm

## 2 Foreign Vision Algorithm Companies

- 2.1 Mobileye
  - 2.1.1 Profile
  - 2.1.2 Main Technologies
  - 2.1.3 Visual Solutions
  - 2.1.4 Major Customers
- 2.2 Continental
  - 2.2.1 Profile
  - 2.2.2 Vision Algorithm Layout
  - 2.2.3 DMS Vision and Algorithm
  - 2.2.4 In-cabin Vision and Algorithm
  - 2.2.5 Surround View Camera and Algorithm

- 2.3 Bosch
  - 2.3.1 Profile
  - 2.3.2 Front View Camera and Algorithm
  - 2.3.3 Surround View Camera and Algorithm
  - 2.3.4 In-cabin Vision and Algorithm
  - 2.3.5 Fused Perception Algorithm

- 2.4 StradVision
  - 2.4.1 Profile
  - 2.4.2 Products
  - 2.4.3 Vision Algorithm
  - 2.4.4 Dynamics

- 2.5 NVIDIA
  - 2.5.1 Profile
  - 2.5.2 Core Algorithms for Autonomous Driving
  - 2.5.3 Autonomous Vehicle Software Stack
  - 2.5.4 DRIVE Perception
  - 2.5.5 Perception Algorithm for Driving Scenario
  - 2.5.6 Perception Algorithm for Parking Scenario
  - 2.5.7 In-cabin Perception Algorithm
  - 2.5.8 Cooperation Dynamics and Partners

- 2.6 Qualcomm
  - 2.6.1 Snapdragon Ride Platform
  - 2.6.2 Snapdragon Ride Vision System
  - 2.6.3 Vision Algorithm Layout
  - 2.6.4 Partners

- 2.7 Valeo
  - 2.7.1 Profile
  - 2.7.2 Core Algorithm Layout
  - 2.7.3 Drive4U Fully Autonomous Driving Solution
  - 2.7.4 Remote Park4U Fully Automated Parking System
  - 2.7.5 Major Customers

- 2.8 Seeing Machines
  - 2.8.1 Profile
  - 2.8.2 DMS Product Roadmap
  - 2.8.3 DMS Technology
  - 2.8.4 DMS Algorithm and Solution
  - 2.8.5 OMS Algorithm and Solution
  - 2.8.6 Cooperation Dynamics

- 2.9 Smart Eyes
  - 2.9.1 Profile
  - 2.9.2 DMS General Development Platform
  - 2.9.3 Eye Tracking Technology and System Solutions
  - 2.9.4 DMS Algorithm
  - 2.9.5 IMS Perception Algorithm
  - 2.9.6 Software and Hardware Integrated Driver Monitoring System (AIS)
  - 2.9.7 Cooperation Dynamics

# Table of Content (2)

- 2.10 Cippa
- 2.10.1 Profile
- 2.10.2 DMS Solution
- 2.10.3 In-cabin Solution
- 2.10.4 Fleet Solution
- 2.10.5 Cooperation Dynamics

- 2.11 XPERI
- 2.11.1 Profile
- 2.11.2 DMS Solution
- 2.11.3 New Generation DMS Solution
- 2.11.4 OMS Solution
- 2.11.5 Partners and Dynamics

- 2.12 Tesla
- 2.12.1 Overview of AI Algorithms for Autopilot Systems
- 2.12.2 Occupancy Networks Algorithm
- 2.12.3 New Lane Detection Algorithm
- 2.12.4 HydarNet Algorithm
- 2.12.5 Autopilot Solutions

## 3 Chinese Vision Algorithm Companies

- 3.1 Momena
- 3.1.1 Profile
- 3.1.2 Visual Perception Algorithm
- 3.1.3 Mass-produced Autonomous Driving Solutions
- 3.1.4 Fully Intelligent Driving Solution
- 3.1.5 Dynamics in Autonomous Driving

- 3.2 Haomo.ai
- 3.2.1 Profile
- 3.2.2 Development Strategy
- 3.2.3 Core Business
- 3.2.4 Intelligent Data System MANA
- 3.2.5 Intelligent Data System MANA - Perception Algorithm
- 3.2.6 Intelligent Data System MANA - Cognition Algorithm
- 3.2.7 Urban Scenario Solutions
- 3.2.8 Service Model and Implemented Projects

- 3.3 Nullmax
- 3.3.1 Profile
- 3.3.2 Core Technologies
- 3.3.3 MaxView Perception Technology System
- 3.3.4 Multi-camera BEV Solution
- 3.3.5 MaxFlow Data Closed Loop
- 3.3.6 Autonomous Driving Solutions
- 3.3.7 Competitive Edges and Major Partners

- 3.4 Motovis
- 3.4.1 Profile
- 3.4.2 Main Products and Solutions
- 3.4.3 Core Algorithm Team and Technologies
- 3.4.4 Visual Perception Based on Deep Learning
- 3.4.5 BEV-based Fused Perception Algorithm

- 3.5 MINIEYE
- 3.5.1 Profile
- 3.5.2 Autonomous Driving Solutions
- 3.5.3 Out-cabin Perception Solution
- 3.5.4 Out-cabin Algorithm and Capabilities
- 3.5.5 Improvements in Out-cabin Algorithm
- 3.5.6 In-cabin Perception Solution
- 3.5.7 In-cabin Perception Algorithm and Capabilities
- 3.5.8 Partners and Dynamics

- 3.6 JIMU Intelligent
- 3.6.1 Profile
- 3.6.2 Out-cabin Perception Algorithm
- 3.6.3 The Work Done by JIMU to Improve Algorithm Accuracy
- 3.6.4 Application of Out-cabin Detection Algorithm
- 3.6.5 In-cabin Driver Monitoring Technology
- 3.6.6 Cooperation Dynamics and Future Development

- 3.7 Smarter Eye
- 3.7.1 Profile
- 3.7.2 Core Technologies
- 3.7.3 Developments and Cooperation

- 3.8 SenseTime
- 3.8.1 Profile
- 3.8.2 Intelligent Vehicle Business Layout
- 3.8.3 SenseAuto Pilot Solution
- 3.8.4 SenseAuto Cabin Solution
- 3.8.5 Core Technologies

# Table of Content (3)

- 3.9 ArcSoft
  - 3.9.1 Profile
  - 3.9.2 Strategic Layout
  - 3.9.3 Vehicle Visual Perception Algorithm
  - 3.9.4 VisDrive Vehicle Vision Solution
  - 3.9.5 Software and Hardware Integrated Vehicle Vision Solution for OEMs: Tahoe
  - 3.9.6 Customers and Partners
- 3.10 Baidu Apollo
  - 3.10.1 Profile
  - 3.10.2 Development History of Baidu Autonomous Driving Perception
  - 3.10.3 Baidu's Main Algorithms in Perception 1.0 Stage
  - 3.10.4 Baidu's Main Algorithms in Perception 2.0 Stage
  - 3.10.5 Baidu's Autonomous Driving System Solutions
  - 3.10.6 Baidu's Vision-only Solution - Apollo Lite
  - 3.10.7 Baidu's Fused Perception Solution - Apollo Lite++
  - 3.10.8 Baidu's End-to-end 3D Perception Development Kit - Paddle3D
  - 3.10.9 Major Clients and Partners of Baidu Apollo
- 3.11 UISEE
  - 3.11.1 Profile
  - 3.11.2 U-Drive Intelligent Driving Platform
  - 3.11.3 U-Pilot Solution for Mass Production
  - 3.11.4 Visual Positioning Technology
  - 3.11.5 R&D Plan and Partners

- 3.12 Horizon Robotics
  - 3.12.1 Profile
  - 3.12.2 Technologies and Solutions
  - 3.12.3 Chip Iteration History
  - 3.12.4 AI Algorithm Layout
  - 3.12.5 BEV Perception Solution
  - 3.12.6 AIDI Development Platform
  - 3.12.7 Intelligent Driving Solutions
  - 3.12.8 Intelligent Driving Solution: Front View Mono
  - 3.12.9 Intelligent Driving Solution: Driving and Parking Integrated Solution
  - 3.12.10 Intelligent Driving Solution: SuperDrive
  - 3.12.11 Partners
- 3.13 Juefx
  - 3.13.1 Profile
  - 3.13.2 Products
  - 3.13.3 Fused Location Production Solution
  - 3.13.4 Fused Location Solution with Visual Features
  - 3.13.5 Development History of BEV Perception Technology
  - 3.13.6 Cooperation Ecosystem
- 3.14 ZongMu Technology
  - 3.14.1 Profile
  - 3.14.2 Visual Products and Systems
  - 3.14.3 Vision Algorithm
  - 3.14.4 Major Customers

- 3.15 ThunderSoft
  - 3.15.1 Profile
  - 3.15.2 Intelligent Vision Products and Core Technologies
  - 3.15.3 Surround View Camera + DMS Vision Algorithms
- 3.16 iVICAR
  - 3.16.1 Profile
  - 3.16.2 Surround View Camera Algorithm Layout
- 4 Summary and Trends**
  - 4.1 Summary on Companies
    - 4.1.1 List of Foreign Vision Algorithm Companies
    - 4.1.2 List of Chinese Vision Algorithm Companies
  - 4.2 Development Trends
    - 4.2.1 Trend 1
    - 4.2.2 Trend 2
    - 4.2.3 Trend 3
    - 4.2.4 Trend 4
    - 4.2.5 Trend 5
    - 4.2.6 Trend 6
    - 4.2.7 Trend 7
    - 4.2.8 Trend 8





## Beijing Headquarters

TEL: 010-82601561, 82863481

Mobile: 137 1884 5418

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

Website:  
[www.researchinchina.com](http://www.researchinchina.com)

WeChat: [zuosiqiche](#)



## Chengdu Branch

TEL: 028-68738514

FAX: 028-86930659