

·2024 中国汽车工程学会年会优秀论文专题·

Research on the Mechanism of Multi-Sensor Fusion Configuration Based on the Optimal Principle of the Vehicle

Zhao Binggen, Zeng Dong, Lin Haoyu, Qiu Xubo, Hu Pijie

(Automotive Research Institute, BYD Auto Industry Company Limited, Shenzhen 518118)

【Abstract】In order to address the issue of sensor configuration redundancy in intelligent driving, this paper constructs a multi-objective optimization model that considers cost, coverage ability, and perception performance. And then, combining a specific set of parameters, the NSGA-II algorithm is used to solve the multi-objective model established in this paper, and a Pareto front containing 24 typical configuration schemes is extracted after considering empirical constraints. Finally, using the decision preference method proposed in this paper that combines subjective and objective factors, decision scores are calculated and ranked for various configuration schemes from both cost and performance preferences. The research results indicate that the multi-objective optimization model established in this paper can screen and optimize various configuration schemes from the optimal principle of the vehicle, and the optimized configuration schemes can be quantitatively ranked to obtain the decision results for the vehicle under different preference tendencies.

Key words: Multi-sensor fusion, Intelligent driving, Multi-objective optimization, Vehicle optimization

基于整车最优原则的多传感器融合配置机理研究(英文)*

赵炳根 曾董 林昊宇 邱旭波 胡丕杰

(比亚迪汽车工业有限公司, 汽车工程研究院, 深圳 518118)

【摘要】针对智能驾驶系统传感器配置冗余问题, 构建了考虑成本、覆盖能力与感知性能三方面的多目标优化模型。然后, 结合一组具体参数, 利用 NSGA-II 算法对所建立的多目标优化模型进行求解, 并在考虑了经验约束后提取出一个包含 24 种典型配置方案的 Pareto 前沿。最后, 利用所提出的主客观结合的决策偏好方法, 从成本偏好和性能偏好两方面对各类配置方案进行决策得分计算与排序。研究表明, 所建立的多目标优化模型可以从整车最优层面对各类配置方案进行筛选和优化, 从而获得不同偏好倾向下满足整车最优的配置方案决策结果。

主题词: 多传感器融合 智能驾驶 多目标优化 整车最优

中图分类号: U463.1 **文献标志码:** A **DOI:** 10.19620/j.cnki.1000-3703.20240707

【引用格式】 赵炳根, 曾董, 林昊宇, 等. 基于整车最优原则的多传感器融合配置机理研究(英文)[J]. 汽车技术, 2024(10): 28-37.
ZHAO B G, ZENG D, LIN H Y, et al. Research on the Mechanism of Multi-Sensor Fusion Configuration Based on the Optimal Principle of the Vehicle[J]. Automobile Technology, 2024(10): 28-37.

1 Introduction

The perception technology of intelligent driving is a key way for vehicles to obtain information of surrounding

environment, and it is also the primary link of the intelligent driving^[1]. The mainstream perception routes in the industry include the multi-sensor fusion route (hereinafter referred to as the fusion route) and the pure

*基金项目: 国家重点研发计划项目(2023YFB2504500)。

通信作者: 林昊宇(1995—), 男, 中级工程师, 博士, 主要研究方向为智能驾驶感知技术与整车智驾平台化, lin.haoyu1@byd.com

visual route. Narrowly speaking, the pure visual route only includes cameras, so it has a higher technical threshold due to algorithm capabilities, hardware facilities, data, and other reasons. Currently, only Tesla has truly achieved the pure visual perception solution. The fusion route is still the perception solution adopted by the vast majority of carmakers. Although the current fusion algorithm is still dominated by the visual processing of cameras, the existence of radar enables the fusion scheme to deal with more long tail scenes, so as to solve the problems of the failure of camera (night, rainstorm, and other scenes) and misjudgment (scenes with the change of complex light)^[2]. At present, the common practice is to pre-embed hardware configurations, and later push software updates and functional upgrades through Over the Air (OTA) technology. However, a major problem brought about by this approach is that it can easily trigger an arms race. The hardware-first strategy significantly increases the cost of Bill of Material (BOM), but cannot maximize the advantages of hardware at the software and algorithm levels. Taking LiDAR as an example, as the most expensive sensor in high-order intelligent driving systems, its point cloud information is currently not used in most perception algorithms developed on actual vehicles.

From the perspective of the optimization of the vehicle, cost, coverage capability, and perceiving performance are three important aspects for measuring intelligent driving systems in the vehicle. In fact, the different types of the vehicle will affect the cost of configuration schemes, and indirectly affect the coverage capability and perceiving performance. From the perspective of the optimization, the optimization directions of cost, coverage capability, and perceiving performance are conflicting with each other. Therefore, establishing a methodology for multi-sensor fusion configuration based on the optimal principle of the vehicle is of great significance for the selection and decision-making of configuration schemes. There is insufficient academic research on this problem. Some academic teams have conducted research on the optimal configuration of individual sensors, such as LiDAR^[3-4] and ultrasonic radar^[5]. Meadows et al. proposed a placement optimization method for multi-LiDAR systems and demonstrated the

effectiveness of the proposed method through data generation, training, and evaluation^[4]. Kim et al. proposed a genetic algorithm based on layout optimization (position and orientation) method to improve the point cloud resolution and blind spot size of LiDAR^[3]. To address the issue of lane changing collisions caused by blind spot monitoring, Jamaluddin et al. analyzed the impact of the position of ultrasonic radar on the driver reminder function^[5]. Some scholars have also studied the joint optimization problem between different sensors, but there are some limitations and shortcomings. Zhou et al. established an integer programming model that considers cost, coverage, and redundancy to determine the optimal number and location of multiple sensors, and solve it using the IBM ILOG CPLEX solver^[6]. However, this model transforms multi-objective problems into single objective problems for modeling, simplifying the problem while sacrificing the model's generalization. In addition, [7-10] and other literatures have also studied the configuration optimization problem of different sensors, but most of these studies still take the sensor layout location and orientation as the optimization goal, and unable to establish systematic constraints and decisions at the level of the vehicle.

Thus, a mechanism model for multi-sensor fusion configuration based on the optimal principle of the vehicle is proposed in this paper, and a multi-objective optimization problem with cost, coverage capability, and perceiving performance is established and solved. In addition, a decision preference method combining subjective and objective factors is proposed, which can assist decision-makers in selecting and making decisions on multi-sensor fusion configuration schemes.

2 Basic Principles

At present, camera, radar, and LiDAR are commonly used in intelligent driving perception schemes. The advantages and disadvantages of all kinds of sensors are very obvious. The camera can recognize the geometric features and shapes of objects. The visual algorithm is mature, and the cost is low. However, it is greatly affected by the change of illumination, and there is the risk of failure in harsh environments^[11]. Radar has all-day, all-weather detection ability, and it can achieve accurate ranging and speed

measurement. The algorithm of radar is mature, and the cost is low, but it does not have the ability to measure height and recognize the stationary objects^[12]. LiDAR has high precision and wide detection range, and can directly obtain high-density 3D environment information, but the cost is high and it is easy to be affected by weather^[11]. In recent years, 4D millimeter wave radar (hereinafter referred to as 4D radar) as a new type of sensor began to rise. 4D radar can solve the inherent defects of traditional radar, such as height measurement and stationary objects recognition, and it can also output high-density point cloud information for subsequent processing links^[12]. It can be predicted that 4D radar, as a cost-effective sensor between LiDAR and traditional radar, will play an important role in the autonomous driving of L2+ or L3 level driving in the future. A statistical comparison of the characteristics, advantages, disadvantages and application scenarios of the above four sensors is given in Figure 1, and each item is measured on a ten-point scale.

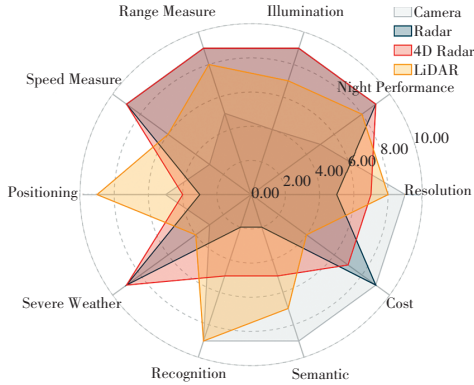


Figure 1 Statistics and comparison of characteristics of various sensors

3 Modeling Methods

A model for multi-sensor fusion configuration based on the optimal principle of the vehicle is proposed in this section. Further, a decision preference method combining subjective and objective factors is proposed to calculate the optimization results of the multi-objective model mentioned above, in order to obtain the optimal ranking result of the configuration scheme at the vehicle level.

3.1 Scene Model

In order to describe the problem of multi-sensor fusion configuration in this paper, the vehicle and surrounding coverage range are first modeled, as shown in Figure 2. The

Region of Interest (RoI) area P_z is divided into ten areas, namely $P_z = P_1 \cup P_2 \cup P_3 \cup P_4 \cup P_5 \cup P_6 \cup P_7 \cup P_8 \cup P_9 \cup P_{10}$, and each RoI area can be assigned different perception weights. For different types of sensors, the possible layout areas are considered according to their respective characteristics, there are, $c \in \{1, 2, \dots, C\} = C$, $r \in \{1, 2, \dots, R\} = R$, and $l \in \{1, 2, \dots, L\} = L$. Where C , R , and L represent the potential placement space of camera, radar/4D radar and LiDAR respectively, which are represented by boxes of different colors in Figure 2. In order to facilitate modeling, the orientation angle of the sensor arrangement is not considered in this paper, and it is assumed that the sensor is placed at an angle perpendicular to the horizontal surface of the vehicle.

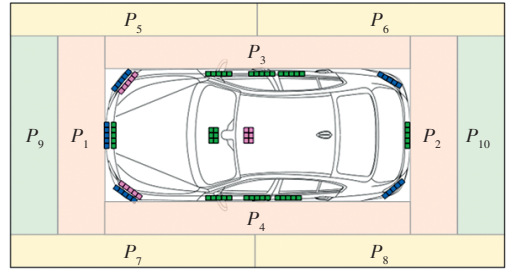


Figure 2 A bird's-eye view (BEV) of the scene model

The RoI area covered by the perception is divided, and its possible layout area is constrained according to different sensor types. The green, blue, and pink boxes represent the layout areas of the camera, radar/4D radar, and LiDAR respectively.

3.2 Multi-Objective Optimization Model

The multi-objective optimization function established in this paper can be expressed as:

$$\begin{aligned} \min F(\mathbf{n}, \boldsymbol{\omega}, \boldsymbol{\kappa}) &= [J_1(\mathbf{n}) \quad J_2(\mathbf{n}, \boldsymbol{\omega}) \quad J_3(\boldsymbol{\omega}, \boldsymbol{\kappa})] \quad (1) \\ \text{s.t. } \mathbf{n} &\in N, \boldsymbol{\omega} \in \boldsymbol{\Omega} \end{aligned}$$

where $J_1(\mathbf{n})$, $J_2(\mathbf{n}, \boldsymbol{\omega})$ and $J_3(\boldsymbol{\omega}, \boldsymbol{\kappa})$ represent 3 sub-objective functions: cost, coverage capability and perceiving performance respectively. \mathbf{n} represents the quantity vector, $\boldsymbol{\omega}$ represents the arrangement vector, and $\boldsymbol{\kappa}$ is used to represent the influence of the perception algorithm. N and $\boldsymbol{\Omega}$ represent the quantity space and layout space of the configuration scheme, respectively, which can be expressed as:

$$N = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1k} \\ n_{21} & n_{22} & \cdots & n_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ n_{m1} & n_{m2} & \cdots & n_{mk} \end{bmatrix}, \boldsymbol{\Omega} = \begin{bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1k} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{m1} & \omega_{m2} & \cdots & \omega_{mk} \end{bmatrix} \quad (2)$$

where $\mathbf{n}_i = (n_{i1}, n_{i2}, \dots, n_{ik})$ represent one of the quantity vectors, $\boldsymbol{\omega}_i = (\omega_{i1}, \omega_{i2}, \dots, \omega_{ik})$ represents one of the arrangement vectors, and $i \in [1, m]$.

The cost is mainly determined by the number and price of various sensors in the configuration scheme, and its objective function can be expressed as:

$$J_1(\mathbf{n}) = \sum_{j=1}^k \mathbf{p}_j \cdot n_{ij} \quad (3)$$

where $i \in [1, m]$, \mathbf{p} is the price vector, and it represents the price of each type of sensor, which satisfies $\mathbf{p}_i = (p_1, p_2, \dots, p_k)$.

The coverage capability is mainly determined by the sensing area and resolution of various types of sensors, and its objective function can be expressed as:

$$J_2(\mathbf{n}, \boldsymbol{\omega}) = \sum_{j=1}^k \left[n_{ij} \cdot \sigma(\boldsymbol{\omega}_{ij}) \cdot \left(\frac{S_{\text{hor}}^j}{\theta_{\text{hor}}^j} + \alpha(j) \cdot \frac{S_{\text{ver}}^j}{\theta_{\text{ver}}^j} \right) \right] \quad (4)$$

where S_{hor}^i and S_{ver}^i represent the horizontal perceiving area and vertical perceiving area of the j th sensor respectively. The perceiving area mainly depends on the sensor's Field of View (FOV). θ_{hor}^i and θ_{ver}^i represent the horizontal and vertical perceiving resolution of the j th sensor. $\sigma(\boldsymbol{\omega}_{ij})$ is the weight of the perceiving coverage that can be obtained by the placement. $\alpha(j)$ is a binary decision variable used to distinguish cameras from radars:

$$\alpha(j) = \begin{cases} 1, & \text{the } j\text{th sensor is radar} \\ 0, & \text{the } j\text{th sensor is camera} \end{cases} \quad (5)$$

The reason for defining the variable $\alpha(j)$ is that the image processing of the camera is in a two-dimensional coordinate, so the perceiving area of the horizontal range is mainly considered. The point cloud processing of the radar is in the three-dimensional coordinate, so the perceiving area of horizontal and vertical ranges are considered at the same time.

The perceiving performance depends on the type of sensors and algorithms. At present, the vision algorithm is still the main perception processing, but radars are used in the fusion route. Radar and vision fusion processing can improve the robustness of the perceiving performance and enhance the coverage capability of the perception scheme. In order to effectively measure the perceiving performance of different configuration schemes, this paper introduces the mean Average Precision (mAP) of various typical algorithms. In addition to visual perception, the mAP of

fusion algorithms, such as radar and vision fusion, LiDAR and vision fusion, 4D radar and vision fusion, are also considered, and sensors of different configuration schemes are combined to finally obtain the quantified evaluation results of perceiving performance. The perceiving performance can be expressed as:

$$J_3(\boldsymbol{\omega}, \boldsymbol{\kappa}) = \frac{\sum_{i=1}^o \beta(i) \cdot \boldsymbol{\kappa}_i^{3D}(\boldsymbol{\omega}) \cdot \sigma_i + \sum_{i=1}^o \beta(i) \cdot \boldsymbol{\kappa}_i^{\text{BEV}}(\boldsymbol{\omega}) \cdot \sigma_i}{2} \quad (6)$$

where o is the number of RoI, σ_i is a weight value representing the importance of each RoI, $\beta(j)$ is a binary decision variable that represents whether the corresponding RoI is within the current perceiving coverage:

$$\beta(i) = \begin{cases} 1, & \text{the } i\text{th RoI is covered} \\ 0, & \text{the } i\text{th RoI is not covered} \end{cases} \quad (7)$$

where $\boldsymbol{\kappa}_i^{3D}$ and $\boldsymbol{\kappa}_i^{\text{BEV}}$ represent the mAP of various algorithms under 2 types of 3D and BEV object detections.

Take $\boldsymbol{\kappa}_i^{3D}$ for example, the perceiving performance of a certain RoI region depends on the algorithm, and also depends on whether the region is covered by multiple sensors:

$$\boldsymbol{\kappa}_i^{3D}(\boldsymbol{\omega}) = [\boldsymbol{\kappa}_{i1}^{3D}(\boldsymbol{\omega}), \dots, \boldsymbol{\kappa}_{ij}^{3D}(\boldsymbol{\omega}), \dots, \boldsymbol{\kappa}_{io}^{3D}(\boldsymbol{\omega})] \quad (8)$$

where $\boldsymbol{\kappa}_{ij}^{3D}(\boldsymbol{\omega})$ can be repressed as:

$$\boldsymbol{\kappa}_{ij}^{3D}(\boldsymbol{\omega}) = \begin{cases} \text{mAP}_j^C \cdot \nu(\boldsymbol{\omega}) + [\text{mAP}_j^C, \text{mAP}_j^{\text{Multi}}] \cdot \boldsymbol{\xi} \cdot [1 - \nu(\boldsymbol{\omega})] \\ \text{mAP}_j^C \cdot \boldsymbol{\xi}(1) \\ \text{mAP}_j^R \end{cases} \quad (9)$$

where mAP_j^C and $\text{mAP}_j^{\text{Multi}}$ represent the mAP of visual algorithms and fusion algorithms, respectively. $\text{mAP}_j^{\text{Multi}}$ can represent the mAP of visual and radar fusion algorithms, visual and 4D radar fusion algorithms, visual and LiDAR fusion algorithms, visual and 4D radar and LiDAR fusion algorithms. $\boldsymbol{\xi}$ represents the weight distribution of vision algorithms and fusion algorithms in the perception task. $\nu(\boldsymbol{\omega})$ represents the coverage ratio of the visual perception within this RoI, which is determined by the placement vector of the configuration scheme.

3.3 Decision Preference Method Combining Subjective and Objective Factors

Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) algorithm is a common comprehensive evaluation method, which can be applied to the problem of schemes selection when there are multiple indicators^[13].

Therefore, this paper will use TOPSIS algorithm to calculate the optimized data, so as to obtain the ranking result. The type of different vehicles often has different preferences on configuration schemes. For example, there is a cost preference for medium-low models, while there is a performance preference for medium-high models. Therefore, this paper introduces the entropy weight method on the basis of TOPSIS algorithm, and calculates the score by subjectively introducing the preference weight and objectively calculating the information entropy. Finally, the ranking result of configuration schemes is obtained considering the influence of preferences.

First of all, data forward transformation is carried out. Since TOPSIS model is generally calculated on the basis that all indicators are extremely large indicators (the larger the data, the better). While in the multi-objective model in this paper, the cost is extremely small data (the smaller the data, the better), so the data of the cost in the optimization results need to be processed forward first. Generally, the following methods can be used for processing.

$$\begin{cases} \bar{x}_i = \left[\frac{1}{x_{i1}}, \frac{1}{x_{i2}}, \dots, \frac{1}{x_{im}} \right] \\ \bar{x}_i = \max(x_i) - [x_{i1}, x_{i2}, \dots, x_{im}] \end{cases} \quad (10)$$

where i represents the index to be turned forward, x_i represents the vector to be turned forward, x_{ij} represents the corresponding data, and $j \in [1, m]$. \bar{x}_i represents the data after the forward transformation.

Secondly, in order to eliminate the dimensional influence between different indicators, the normalization of the forward matrix is carried out:

$$z_{ij} = \frac{\bar{x}_{ij}}{\sqrt{\sum_{i=1}^m \bar{x}_{ij}^2}} \quad (11)$$

Then, the entropy weight method is introduced to calculate the final score. The probability matrix is calculated using the forward normalized matrix:

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^m z_{ij}} \quad (12)$$

Calculate the information entropy e_i corresponding to each index:

$$e_i = -\frac{1}{\ln(m)} \cdot \sum_{j=1}^m [p_{ij} \cdot \ln(p_{ij})] \quad (13)$$

Further, the utility value d_i of each index is calculated:

$$d_i = 1 - e_i \quad (14)$$

The entropy weight ω_i of each index can be calculated by the following formula:

$$\omega_i = \frac{d_i}{\sum_{i=1}^n d_i} \cdot v_i \quad (15)$$

where n is the number of indicators, v_i represents the subjective preference weight, and the sum of the weights of each indicator is 1.

Furthermore, the ideal best and worst values are determined for each indicator:

$$\begin{cases} Z^+ = [\max(z_{11}, \dots, z_{1m}), \max(z_{21}, \dots, z_{2m}), \dots, \max(z_{n1}, \dots, z_{nm})] \\ Z^- = [\min(z_{11}, \dots, z_{1m}), \min(z_{21}, \dots, z_{2m}), \dots, \min(z_{n1}, \dots, z_{nm})] \end{cases} \quad (16)$$

Calculate the Euclidean distance between each configuration scheme and its ideal best and worst values, taking into account the entropy weight. Define the distance of the j th scheme from the maximum and minimum values as:

$$\begin{cases} D_j^+ = \sqrt{\sum_{i=1}^m \omega_i \cdot (Z_i^+ - z_{ij})^2} \\ D_j^- = \sqrt{\sum_{i=1}^m \omega_i \cdot (Z_i^- - z_{ij})^2} \end{cases} \quad (17)$$

Finally, the relative proximity S of the ideal best and worst distances is calculated and normalized to obtain the final score result \tilde{S} of each configuration method:

$$S = \frac{D_j^-}{D_j^- + D_j^+} \quad (18)$$

$$\tilde{S} = \frac{S_j}{\sum_{j=1}^m S_j} \quad (19)$$

4 Model Solving

For the multi-objective optimization problem constructed in this paper, the optimization direction of multiple objectives is conflicting, and the optimization of the cost is at the cost of the deterioration of the coverage capability or the perceiving performance, and vice versa. Therefore, there is no single optimal solution for the configuration scheme, and the different positioning of the vehicle will inevitably lead to cost differences, which in turn lead to differences in coverage capability and perceiving performance.

Therefore, the NSGA-II algorithm is adopted in this

paper to solve the above multi-objective optimization problem, and the Pareto frontier is obtained^[14]. The Pareto frontier here is a set of optimal solutions that comprehensively consider the three objectives of cost, coverage capability and perceiving performance, and none of the solutions can be improved without reducing the value of other objectives. The steps of the algorithm are shown in Table 1, and the framework diagram of the NSGA-II algorithm is shown in Figure 3.

Table 1 Steps of NSGA-II algorithm

Algorithm NSGA-II	
Step 1 Initialization:	The random algorithm is used to generate the initial population P_0 .
Step 2 Non-dominated sorting operation:	Rationality judgment and non-dominated sorting are carried out for all individuals in the current population. a. Determine whether each individual has a reasonable scheme; b. Pareto grading is carried out to each individual; c. The crowding degree of individuals in the same Pareto level is calculated and sorted in descending order.
Step 3 Selection:	Select optimal individuals from the current population P_t , and perform crossover operations and mutation operations to generate sub-population Q_t .
Step 4 Merge:	Merge populations P_t and Q_t to produce combined population R_t .
Step 5 Instead:	Conduct rationality judgment and non-dominant sorting operation for the combined population (same as step 2). Select the optimal individual, and produce a new generation of population P_{t+1} .
Step 6 Judgment:	Judge whether the end condition is met, if not, jump to step 3 ; Otherwise exit the loop to get the optimal solution set.

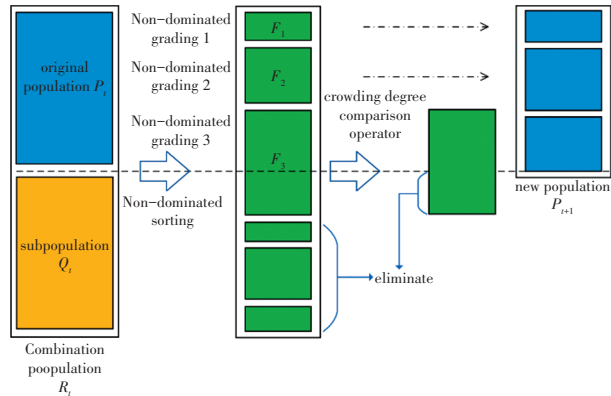


Figure 3 Basic framework diagram of the NSGA-II algorithm

In order to obtain objective and quantified optimization results for perceiving performance, this paper fully surveys various typical algorithms in 3D and BEV object detections, and uses the mAP of the corresponding algorithms to the objective function of perceiving

performance shown in equation (6). The quantized values of perceiving performance are shown in Table 2.

Table 2 Quantified perceiving performance of various typical algorithms

Types	Algorithms	Performance[mAP]	
		3D	BEV
Camera	FCOS3D ^[15]	32.60	
	DD3D ^[17]	41.80	
	FUTR3D ^[15]	41.20	
Radar/Lidar	PointRCNN ^[17]	11.40	18.74
	SECOND ^[18]	18.02	31.01
	RV-RCNN ^[19]	22.08	39.88
	PointPillars ^[20]	20.49	38.21
	Part-A ^[21]	13.76	21.47
	Voxel R-CNN ^[22]	18.71	31.26
Camera+Radar	MVX-Net ^[23]	11.69	20.36
Camera+LiDAR	FUTR3D ^[15]	43.40	51.20
Camera+4D Radar	SMURF ^[24]	32.99	40.98
Camera+4D Radar+LiDAR	M ² -Fusion ^[25]	49.85	61.24

5 Result Analysis

In this paper, a multi-objective optimization model is constructed to study the configuration mechanism of the fusion route. The 3 objective functions of cost, coverage ability and perceiving performance are considered, and use NSGA-II algorithm to generate the optimal configuration scheme of the vehicle. In addition, this paper puts forward a preference decision method combining subjective and objective factors, and uses TOPSIS algorithm and entropy weight method to determine the optimal solution under different preferences.

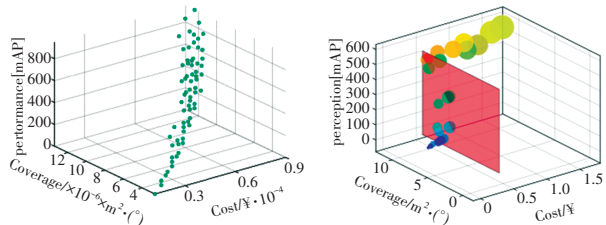
Table 3 shows parameters of sensors and algorithm parameters used for simulation analysis in this section. The parameters here are only an example and can be changed according to the actual situation.

Table 3 Simulation parameters

Sensor types	Parameters	Values
Front radar	Price/¥	300
	Horizontal FOV(°)/resolution(°)	30/1.6
	Vertical FOV(°)/resolution(°)	10/1.5
	Detection distance/m	250
Corner radar	Price/¥	200
	Horizontal FOV(°)/resolution(°)	150/8
	Vertical FOV(°)/resolution(°)	20/10
	Detection distance/m	20

Table 3. Continued

Sensor types	Parameters	Values
4D radar	Price/¥	2 000
	Horizontal FOV(°)/resolution(°)	30/0.3
	Vertical FOV(°)/resolution(°)	10/0.2
	Detection distance/m	250
Monocular camera	Price/¥	300
	Horizontal FOV(°)/resolution(°)	120
	Detection distance/m	250
Binocular camera	Price/¥	600
	Horizontal FOV(°)/resolution(°)	60/120
	Detection distance/m	250/50
Three-lens camera	Price/¥	900
	Horizontal FOV(°)/resolution(°)	28/150/52
	Detection distance/m	250/60/150
Rear camera	Price/¥	300
	Horizontal FOV(°)/resolution(°)	50
	Detection distance/m	50
360-degree camera	Price/¥	300
	Horizontal FOV(°)/resolution(°)	190
	Detection distance/m	20
Side-view camera	Price/¥	300
	Horizontal FOV(°)/resolution(°)	90
	Detection distance/m	100
Front LiDAR	Price/¥	6 000
	Horizontal FOV(°)/resolution(°)	120/0.2
	Vertical FOV(°)/resolution(°)	25/0.2
	Detection distance/m	200
Lateral LiDAR	Price/¥	6000
	Horizontal FOV(°)/resolution(°)	120/0.2
	Vertical FOV(°)/resolution(°)	90°/0.2
	Detection distance/m	30
Algorithm parameters	$\sigma(\omega_{ij})/\sigma_i$	[1,0.6,0.6,0.6,0.8,0.5,0.8,0.5,1,0.5]
	ξ	[0.7, 0.3]
	$\nu(\vec{\omega})$	0.6



(a) Complete results

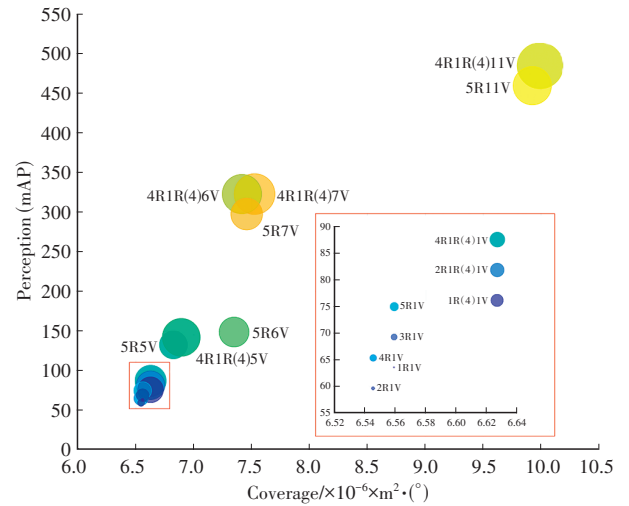
(b) Typical results extracted after considering empirical constraints

Figure 4 Pareto frontier results obtained by the NSGA-II algorithm to solve multi-objective optimization model

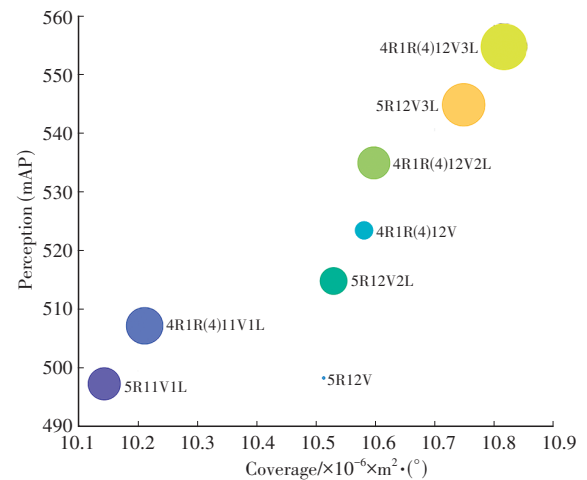
Figure 4a shows the result of Pareto frontier obtained by the NSGA-II algorithm, where the population size is set

to 100, the number of iterations is 200, and the crossover probability and mutation probability are 0.8 and 0.05, respectively. The Pareto frontier here includes 72 kinds of optimized results. Further, considering the empirical constraints such as layout habits and quantity upper limit of the current mainstream configuration schemes, 24 major configuration schemes are extracted from the optimization results, as shown in Figure 4b. Each circle in Figure 4 represents a configuration scheme. X axis, Y axis and Z axis represent the cost, coverage capability, and perceiving performance, respectively. In Figure 4b, the 5R11V scheme is used as the boundary, and 24 optimization schemes are divided into medium-low configuration and medium-high configuration.

The two-dimensional view is shown in Figure 5. The size of the circle in the figure is positively correlated with the relative cost in each category.



(a) Medium-low configuration results



(b) Medium-high configuration results

Figure 5 Two-dimensional view of the results shown in Figure 4b

The perception schematic of 24 typical configuration schemes and the numerical results of the three optimization objectives are given in Figure 6 and Table 4. Further, based on 24 typical configuration schemes obtained by the optimization algorithm, the decision preference method combining subjective and objective factors proposed in this paper is used for decision calculation and ranking. Two kinds of decision preferences are analyzed in this paper, they are cost preference and performance preference. The analysis results are shown in Figure 7, and the specific scores and ranking results are given in Table 5 and Table 6. It can be seen that when considering different decision preferences of cost preference and performance preference, the final ranking of configuration scheme also shows an obvious tendency. When the cost preference is given priority, the top five options are 1R1V (scheme 1), 2R1V (scheme 3), 3R1V (scheme 4), 4R1V (scheme 6) and 5R1V (scheme 7). This indicates that in the medium-low models with cost as the first, these schemes are the best schemes for the vehicle. When performance preference is given

priority, the top five options are 5R12V (scheme 19), 4R1R⁴12V (scheme 20), 4R1R⁴11V (scheme 16), 5R11V (scheme 15), and 5R11V1L (scheme 17). This indicates that these schemes are the best schemes for the vehicle in the medium-high models with performance as the first.

Another important conclusion about the LiDAR configuration can be drawn from Figure 7. Under the performance preference decision, the multi-LiDAR configuration scheme (scheme 21, scheme 22, scheme 23, scheme 24) does not rank high. This indicates that after considering the three important factors of cost, coverage ability and perceiving performance comprehensively, the multi-LiDAR configuration scheme is still not the optimal scheme at the vehicle level even under the performance preference decision. The main reason is that the point cloud information of LiDAR is not fully utilized in most of the current fusion algorithms. Currently, the degree of improvement brought by LiDAR to the perceiving performance is not enough to make up for the cost brought by it to the BOM of the vehicle.

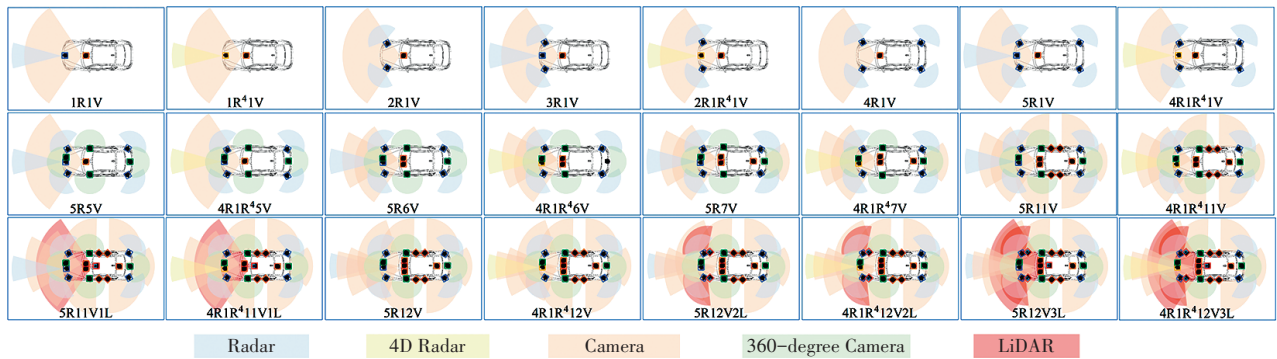


Figure 6 Perception diagram of 24 typical configuration schemes

Table 4 Optimization results of 24 typical configuration schemes

Schemes	Configurations	Cost/¥	Coverage ability / $\times 10^{-6} \times m^2 \cdot (^{\circ})$	Perceiving performance (mAP)	Schemes	Configurations	Cost/¥	Coverage ability / $\times 10^{-6} \times m^2 \cdot (^{\circ})$	Perceiving performance (mAP)
1	1R1V	600	6.558 8	63.56	13	5R7V	3 200	7.457 1	297.89
2	1R ⁴ 1V	2 300	6.626 8	76.14	14	4R1R ⁴ 7V	4 900	7.525 1	323.05
3	2R1V	700	6.545 1	59.74	15	5R11V	4 100	9.924 1	459.73
4	3R1V	1 000	6.559 0	69.36	16	4R1R ⁴ 11V	5 800	9.992 1	484.89
5	2R1R ⁴ 1V	2 700	6.626 9	81.93	17	5R11V1L	10 100	10.143 4	497.26
6	4R1V	1 100	6.545 3	65.54	18	4R1R ⁴ 11V1L	11 800	10.211 3	507.16
7	5R1V	1 400	6.559 1	75.15	19	5R12V	4 700	10.513 2	498.27
8	4R1R ⁴ 1V	3 100	6.627 1	87.73	20	4R1R ⁴ 12V	6 400	10.581 1	523.42
9	5R5V	2 600	6.824 4	132.70	21	5R12V2L	16 700	10.529 7	514.78
10	4R1R ⁴ 5V	4 300	6.892 4	142.26	22	4R1R ⁴ 12V2L	18 400	10.597 6	534.90
11	5R6V	2 900	6.933 5	148.95	23	5R12V3L	22 700	10.748 9	544.80
12	4R1R ⁴ 6V	4 600	7.001 5	323.05	24	4R1R ⁴ 12V3L	24 400	10.816 9	554.70

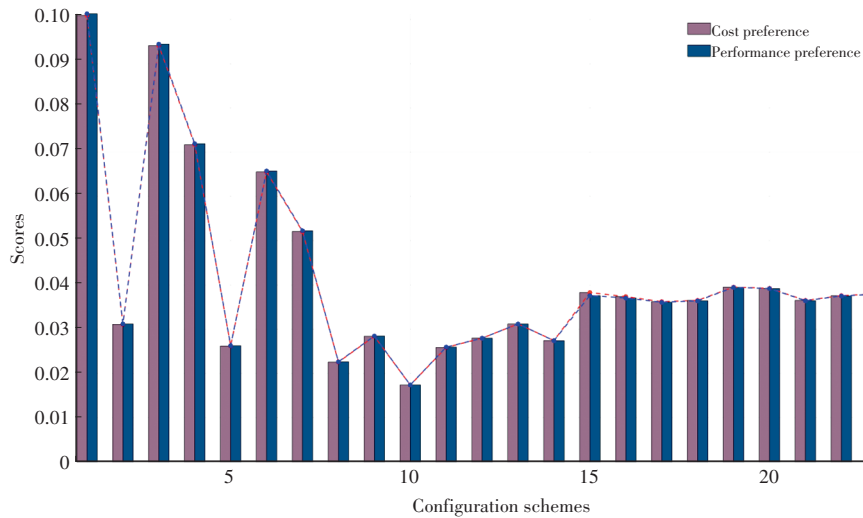


Figure 7 Ranking results of scores for 24 typical configuration schemes under two different decision preferences

Table 5 Calculation and ranking results of decision preferences combining subjective and objective factors (Scheme 1 to 12)

Schemes	1	2	3	4	5	6	7	8	9	10	11	12
Cost preference score	0.100 9	0.031 1	0.094 1	0.071 7	0.026 2	0.065 6	0.052 1	0.022 6	0.028 4	0.017 4	0.025 9	0.028 0
Cost preference ranking	1	17	2	3	21	4	5	23	18	24	22	19
Performance preference score	0.027 7	0.026 8	0.027 5	0.027 6	0.026 8	0.027 4	0.027 5	0.026 7	0.029 3	0.028 5	0.030 1	0.044 2
Performance preference ranking	17	22	20	18	23	21	19	24	15	16	14	11

Table 6 Calculation and ranking results of decision preferences combining subjective and objective factors (Scheme 13 to 24)

Schemes	13	14	15	16	17	18	19	20	21	22	23	24
Cost preference score	0.031 2	0.027 4	0.038 3	0.037 4	0.036 3	0.036 5	0.039 5	0.039 2	0.036 5	0.037 6	0.038 0	0.038 5
Cost preference ranking	16	20	9	12	15	13	6	7	14	11	10	8
Performance preference score	0.042 5	0.044 1	0.060 7	0.062 0	0.058 1	0.056 7	0.064 5	0.064 4	0.051 7	0.050 9	0.047 6	0.046 7
Performance preference ranking	13	12	4	3	5	6	1	2	7	8	9	10

6 Conclusions

Multi-sensor fusion sensing technology is the most widely used scheme for intelligent driving, which is composed of various types of sensors. From the perspective of the vehicle, how to select the appropriate configuration to obtain the optimal scheme at the vehicle level is a problem worthy of in-depth study.

To solve this problem, this paper establishes a multi-objective optimization model considering three important objectives: cost, coverage ability and perceiving performance. Considering the influence of different types of the vehicle on decision-makers' preferences, a subjective and objective decision preference method based on the TOPSIS algorithm and entropy weight method is proposed. Further, the NSGA-II algorithm is used in this paper to solve the multi-objective optimization model, and a Pareto

frontier containing 24 typical configuration schemes is extracted after considering the empirical constraints. Finally, the calculation and ranking of the various schemes is carried out. The cost preference and the performance preference are taken for example to calculate the decision of 24 typical configuration schemes.

The proposed method can be used to screen and optimize all kinds of configuration schemes, and the optimized schemes can be quantitatively ranked, so as to obtain the decision results under different preference tendencies. The research results of this paper are of great significance for decision-makers to select and make decisions on multi-sensor fusion configuration schemes.

References

- [1] ZHAO J Y, ZHAO W Y, DENG B, et al. Autonomous Driving System: A Comprehensive Survey[J]. Expert Systems with Applications. 2024, 242.

- [2] QIAN R, LAI X, LI X R, et al. 3D Object Detection for Autonomous Driving: A Survey[J]. *Pattern Recognition*, 2022, 138(8): 1–24.
- [3] KIM T H, PARK T H. Placement Optimization of Multiple Lidar Sensors for Autonomous Vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 21(5): 2139–2145.
- [4] MEADOWS W, HUDSON C R, GOODIN C, et al. Multi-Lidar Placement, Calibration, Co-Registration, and Processing on A Subaru Forester for Offroad Autonomous Vehicles Operations[J]. *Autonomous Systems: Sensors, Processing, and Security for Vehicles and Infrastructure*, 2019, 11009: 99–116.
- [5] JAMALUDDIN M H, SHUKOR A Z, MISKON M F, et al. An Analysis of Sensor Placement for Vehicle's Blind Spot Detection and Warning System[J]. *Journal of Telecommunication, Electronic and Computer Engineering*, 2016, 8(7): 101–106.
- [6] ZHOU M, CHEN Q F, CAO Y G, et al. Optimal Configuration Scheme of Multi-Sensor Perception for Autonomous Vehicles Based on Solid-State Lidar[C]// 35th Chinese Control and Decision Conference (CCDC). 2023: 188–193.
- [7] CHEN L, LI Q, LI M, et al. Design of A Multi-Sensor Cooperation Travel Environment Perception System for Autonomous Vehicle[J]. *Sensor*, 2012, 12(9): 12386–12404.
- [8] DEY J, TAYLOR W, PASRICHA S. VESPA: A Framework for Optimizing Heterogeneous Sensor Placement and Orientation for Autonomous Vehicles[J]. *IEEE Consumer Electronics Magazine*, 2020, 10(2): 16–26.
- [9] HARTSTERN M, RACK V, STORK W. Conceptual Design of Automotive Sensor Systems: Analyzing the Impact of Different Sensor Positions on Surround-View Coverage[J]. *IEEE Sensors*, 2020: 1–4.
- [10] PRAMANIK S, VAIDYA V, MALVIYA G, et al. Optimization of Sensor-Placement on Vehicles Using Quantum-Classical Hybrid Methods[J]. 2022 IEEE International Conference on Quantum Computing and Engineering (QCE), 2022: 820–823.
- [11] CUI Y D, CHEN R, CHU W B, et al. Deep Learning for Image and Point Cloud Fusion in Autonomous Driving: A Review[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(2): 722–739.
- [12] HAN Z Y, WANG J H, XU Z, et al. 4D Millimeter-Wave Radar in Autonomous Driving: A Survey[EB/OL]. (2023–06–07)[2024–08–12]. <https://arxiv.org/abs/2306.04242>.
- [13] MITRA M, HAMED T. A Comprehensive Guide to the TOPSIS Method for Multi-Criteria Decision Making[J]. *Sustainable Social Development*, 2023, 1(1): 1–6.
- [14] DEB K, PRATAP A, AGARWAL S, et al. A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II[J]. *IEEE Transactions on Evolutionary Computation*, 2002, 6(2): 182–197.
- [15] CHEN X, ZHANG T, WANG Y, et al. Futr3D: A Unified Sensor Fusion Framework For 3D Detection[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023.
- [16] PARK D, AMBRUS R, GUIZILINI V, et al. Is Pseudo-Lidar Needed for Monocular 3D Object Detection? [C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Nashville: IEEE, 2021.
- [17] SHI S, WANG X, LI H. PointRCNN: 3D Object Proposal Generation and Detection from Point Cloud[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019.
- [18] YAN Y, MAO Y, LI B. SECOND: Sparsely Embedded Convolutional Detection[J]. *Sensors*, 2018, 18(10): 3337.
- [19] SHI S S, GUO C X, JIANG L, et al. PV-RCNN: Point-Voxel Feature Set Abstraction for 3D Object Detection[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020.
- [20] LANG A H, VORA S, CAESAR H, et al. Point-Pillars: Fast Encoders for Object Detection from Point Clouds[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019.
- [21] SHI S S, WANG Z, SHI J P, et al. From Points to Parts: 3D Object Detection from Point Cloud with Part-Aware and Part-Aggregation Network[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(8): 2647–2664.
- [22] DENG J J, SHI S S, LI P W, et al. Voxel R-CNN: Towards High Performance Voxel-Based 3D Object Detection[C]// *Proceedings of the AAAI Conference on Artificial Intelligence*. 2021: 1201–1209.
- [23] SINDAGI V A, ZHOU Y, TUZEL O. MVX-Net: Multimodal Voxelnets for 3D Object Detection[C]// *IEEE International Conference on Robotics and Automation (ICRA)*, 2019, 35(2): 1201–1209.
- [24] LIU J N, ZHAO Q C, XIONG W Y, et al. SMURF: Spatial Multi-Representation Fusion for 3D Object Detection with 4D Imaging Radar[J]. *IEEE Transactions on Intelligent Vehicles*. 2023, 9(1): 1–14.
- [25] WANG L, ZHANG X Y, LI J, et al. Multi-Modal and Multi-Scale Fusion 3D Object Detection of 4D Radar and LiDAR for Autonomous Driving[J]. *IEEE Transactions on Vehicular Technology*. 2023, 72(5): 5628–5641.

(责任编辑 王 一)

修改稿收到日期为2024年8月12日。