ERA, 31(3): 1439–1465.
DOI: 10.3934/era.2023073
Received: 27 October 2022
Revised: 26 December 2022
Accepted: 02 January 2023
Published: 11 January 2023

Electronic Research Archive

http://www.aimspress.com/journal/ERA

Research article

Correlation model between mesostructure and gradation of asphalt mixture based on statistical method

Chao Xing[1], Bo Liu [1], Kai Zhang[2], Dawei Wang [1], Huining Xu [1] and Yiqiu Tan[1,*]

[1] School of Transportation Science and Engineering, Harbin Institute of Technology, No. 73, Huanghe Road, Nangang District, Harbin, Heilongjiang, China
[2] China State Construction International Holdings Limited, No.5 Qimin Street, Lang Shan No.2 Road, North of High Tech IndustrialPark, Nanshan District, Shenzhen, China

* **Correspondence:** Email: tanyiqiu@hit.edu.cn; Tel: +86045186283090; Fax: +86045186283090.

Abstract: Asphalt mixture has complex gradation and mesostructure. Accurate prediction of the relationship between gradation and mesostructure is of great significance for the establishment of mesostructure numerical simulation model and image-based gradation detection. In this paper, featurization, stepwise regression, econometric hypothesis test are utilized for establishing the predicting models. Firstly, asphalt mixtures with 64 kinds of gradation are scanned by Computed Tomography (CT) to obtain the mesostructure images; Then a series of mesostructure parameters of voids and aggregates are put forward. On this basis, the relationship model between gradation and mesostructure is established and verified by featurization and statistical modeling method. The results show that for predicting the passing percentage of the 4.75 mm sieve and the mean value of average distance between aggregate centroids for 9.5–4.75 mm aggregates, the prediction error of passing percentage is acceptable. It illustrates that the relationship model between gradation and mesostructure established by statistical method is effective, and it is significance for material design and testing under the condition of big data in the future.

Keywords: asphalt mixture; gradation; mesostructured; image analysis; statistical modeling

1. Introduction

Asphalt mixture is a multiphase composite material composed of aggregate, asphalt and mineral

powder. There are wide variations for the physical and mechanical properties in different kinds of asphalt mixtures, due to the composite gradation. The complexity and randomness of materials lead to the diversity of mesostructure and the complexity of mechanical behavior. As a bridge between gradation and macro mechanical properties, mesostructure is of great significance to reveal the influence of gradation on mechanical properties. In general, the macro performance refers to the overall performance of asphalt mixture, and the characteristics in molecular scale is defined as the micro performance. The meso scale belongs to the middle scale, and the mesostructure mainly includes aggregate, asphalt mortar and void. On the one hand, the mesostructure model of asphalt mixture used for numerical simulation needs to have a good correlation with gradation information[1,2]; On the other hand, in pavement detection, predicting gradation information through the images of core samples can reduce the amount of test and improve the detection efficiency [3,4]. However, at present, the correlation model between mesostructure and gradation is relatively simple, less influencing factors are considered, and there is a lack of statistical model, which limits the application of mesostructure research results.

With the development of CT technology, it provides an important approach to study the mesostructure of asphalt mixture. The mesostructure CT images can be accurately identified and segmented to obtain the aggregate, asphalt mortar and void by digital image processing technology [5–8]. For the study of void characteristics, the plane and spatial distribution of voids are mainly studied. Wang Linbing [9] obtained the internal images of asphalt mixture from the Westrack test road by the industrial CT, and extracted the void spatial distribution and size information to evaluate the mesostructure property of asphalt mixture. Masad [10] concluded that the spatial distribution of void size in asphalt mixture showed a trend of larger at both ends and smaller in the middle. Compaction mode and aggregate size had a great impact on void size distribution. Xu Huining [11] put forward the parameters such as void fraction, tortuosity and average void diameter, and revealed the change of void distribution characteristics during freeze-thaw cycles. Further study [12] obtained the influence of void structure on water seepage in asphalt mixture. Arambula [13] obtained the void distribution in the depth direction of asphalt mixture with different gradation and compaction method and established the relationship between void and water stability. Jiang Wei [14] put forward the evaluation indexes of void quantity and area. The research showed that the meso void distribution characteristics were greatly affected by gradation and maximum nominal particle size, and the equivalent void diameter had a good correlation with the permeability, sound absorption and shear strength of asphalt mixture.

As the main load bearing structure of asphalt mixture, the spatial distribution of aggregate directly affects the service performance [15–17]. Industrial CT can extract the information of aggregates in asphalt mixture, and then obtain its position and volume. Yue [18] analyzed the quantitative measurement results of coarse aggregate distribution and shape in dense asphalt mixture, and found that the main axis direction has a trend of horizontal distribution. Gopalakrishnan [19] extracted the nearest distance between aggregates and measured the change of distance during compaction. Coenen [20] defined the contact distance threshold through the relationship between the contact distance and the number of contact points, and then determined the number of contact points as the evaluation index of the skeleton structure. Sefidmazgi [21] proposed aggregate contact evaluation indexes such as the number of contact points, contact line length and contact plane direction, and proved that aggregate contact have a good correlation with the rutting resistance.

It can be seen from the above that at present, the research on the mesostructure of asphalt mixture is carried out from the aspects of morphology and distribution of void and aggregate. However, each study only focuses on a small number of gradation, and there is a lack of statistical relationship model between mesostructure and gradation. Therefore, this study obtains a large number of mesostructure images of asphalt mixture with different gradation through industrial CT, extracts mesostructure parameters, and then establishes the relationship between gradation and mesostructure through the statistical methods. The research objective of this paper is to accurately predict the meso structure and gradation, mainly including two aspects:

1) In the road detection process, it is necessary to obtain the gradation information of a section of the road. The gradation can be predicted by the images of the core sample of the road.

2) In order to accurately establish the mesostructure discrete element model of asphalt mixture [22–24], the mesostructure parameters should be accurately predicted by gradation.

2. Materials and test method

2.1. Properties of materials

In this study, andesite from Heilongjiang province, limestone powder from Jilin Province and 90# neat asphalt are used. Materials tests are carried out according to JTG E42-2005 Test Methods of Aggregate for Highway Engineering and JTGF40-2004 Technical Specification for Construction of Highway Asphalt Pavements. The properties of materials are shown in Tables 1–3.

Table 1. Properties of aggregate.

Indexes	1–16 mm	16–13.2 mm	13.2–9.5 mm	9.5–4.75 mm	4.75–2.36 mm	2.36–1.18 mm	Sand
Bulk density (g/cm^3)	2.669	2.658	2.691	2.685	2.649	2.679	-
Apparent density (g/cm^3)	2.682	2.665	2.694	2.690	2.666	2.724	2.678
Water absorption (%)	0.24	0.32	0.35	0.36	0.47	0.62	-

Table 2. Properties of mineral powder.

Apparent density (g/cm^3)	Water content (%)	Percentage passing (%)			
		0.6 mm	0.3 mm	0.15 mm	0.075 mm
2.720	0.5	100	100	96.3	80.7

Table 3. Properties of asphalt.

Indexes	Unit	Test results	Specification requirements
Penetration (25℃, 100 g, 5 s)	0.1 mm	91	80~100
Ductility (15℃, 5 cm/min)	cm	75	≥ 50
Softening point	℃	60	≥ 45

2.2. Gradation of asphalt mixture

In order to improve the accuracy of prediction results, a large number of CT scanning specimens are needed. By adjusting the proportion of coarse and fine aggregate, 66 gradations are selected in this research. Two repetitions are manufactured for each gradation, and a total of 132 specimens are scanned by CT. The 66 gradations are shown in Table 4.

Table 4. Passing percentages of asphalt mixtures.

Number	Mixture Types	19 mm	16 mm	13.2 mm	9.5 mm	4.75 mm	2.36 mm	1.18 mm	0.6 mm	0.3 mm	0.15 mm	0.075 mm
1	AC20-1	94.0	87.0	79.7	58.3	36.5	23.5	16.1	11.3	9.0	7.9	5.5
2	AC20-2	90.0	80.0	72.0	66.0	50.0	38.0	31.0	22.0	15.0	12.0	7.0
3	AC20-3	92.0	82.0	68.0	60.0	43.0	34.0	28.0	19.0	13.0	10.5	6.0
4	AC20-4	95.0	85.0	71.0	61.0	50.5	30.0	22.5	16.0	11.0	8.5	5.0
5	AC20-5	95.0	85.0	71.0	61.0	41.0	30.0	22.5	16.0	11.0	8.5	5.0
6	AC20-6	97.7	88.2	74.8	70.9	50.0	34.1	24.1	15.1	10.2	7.0	5.2
7	AC20-7	97.1	86.0	70.5	66.2	45.0	30.0	21.4	13.6	9.5	6.7	5.2
8	AC20-8	96.5	83.7	66.1	61.6	40.1	25.9	18.7	12.2	8.7	6.4	5.1
9	SMA20-1	95.0	82.0	72.0	47.5	24.0	17.5	16.0	13.0	11.5	10.5	10.0
10	SMA20-2	100.0	90.0	79.0	53.0	28.0	18.0	15.0	11.0	10.0	8.5	8.0
11	SMA20-3	94.0	87.0	79.7	58.3	35.8	19.6	12.9	9.1	7.2	6.3	4.4
12	AC16-1	100.0	94.1	86.2	74.1	47.4	30.0	21.4	13.6	9.5	6.7	5.2
13	AC16-2	100.0	94.9	88.1	77.7	52.0	34.1	24.1	15.1	10.2	7.0	5.2
14	AC16-3	100.0	95.8	90.1	81.4	56.6	38.2	26.9	16.5	11.0	7.3	5.3
15	AC16-4	100.0	95.0	84.0	70.0	48.0	34.0	24.5	17.5	12.5	9.5	6.0
16	AC16-5	100.0	97.5	88.0	75.0	48.0	27.0	18.8	13.3	9.8	7.3	5.0
17	AC16-6	100.0	97.5	88.0	75.0	55.0	41.0	24.5	13.3	9.8	7.3	5.0
18	SMA16-1	100.0	90.5	76.2	62.5	28.8	18.2	15.5	12.9	11.5	10.4	9.7
19	SMA16-2	100.0	90.5	76.2	62.6	31.3	20.7	17.1	13.7	11.9	10.6	9.7
20	SMA16-3	100.0	90.2	75.4	61.7	31.9	21.5	17.7	14.0	12.1	10.7	9.7
21	SMA16-4	100.0	97.5	80.0	60.0	26.0	17.3	16.0	13.5	11.3	10.3	9.0
22	SMA16-5	100.0	95.0	75.0	55.0	26.0	19.5	18.0	15.0	12.5	11.5	10.0
23	SMA16-6	100.0	97.5	80.0	60.0	29.0	21.8	18.0	13.5	11.3	10.3	9.0
24	OGFC16-1	100.0	90.5	76.2	62.4	24.8	13.2	10.5	7.9	6.5	5.5	4.9
25	OGFC16-2	100.0	90.3	75.4	63.3	27.2	15.7	12.1	8.7	6.9	5.7	4.9
26	OGFC16-3	100.0	90.5	76.2	62.6	29.7	18.1	13.7	9.6	7.4	5.9	5.0
27	OGFC16-4	100.0	97.5	85.0	63.8	26.0	13.0	9.0	6.8	5.3	4.3	3.0
28	OGFC16-5	100.0	95.0	80.0	57.5	26.0	16.0	12.0	9.5	7.5	5.5	4.0
29	OGFC16-6	100.0	97.5	85.0	63.8	28.0	19.0	12.0	6.8	5.3	4.3	3.0
30	AC13-1	100.0	100.0	95.4	67.9	37.2	34.6	24.9	18.0	12.8	9.3	5.5
31	AC13-2	100.0	100.0	95.3	67.4	41.4	33.4	24.0	17.4	12.4	9.0	5.3
32	AC13-3	100.0	100.0	95.3	67.1	46.3	32.9	23.7	17.1	12.2	8.9	5.3
33	AC13-4	100.0	100.0	94.6	74.9	49.9	32.3	23.2	16.8	11.9	8.7	5.2
34	AC13-5	100.0	100.0	95.0	76.5	53.0	37.0	26.5	19.0	13.5	10.0	6.0
35	AC13-6	100.0	100.0	90.5	76.0	47.9	30.0	21.4	13.6	9.5	6.7	5.2

36	AC13-7	100.0	100.0	91.8	79.4	52.5	34.1	24.1	15.1	10.2	7.0	5.2
37	AC13-8	100.0	100.0	93.2	82.7	57.0	38.2	26.9	16.5	11.0	7.3	5.3
38	AC13-9	100.0	100.0	97.1	78.8	53.0	32.5	19.9	15.3	9.4	7.2	5.5
39	SMA13-1	100.0	100.0	90.2	74.8	29.9	16.6	14.4	12.3	11.2	10.3	9.6
40	SMA13-2	100.0	100.0	90.2	74.8	31.5	18.2	15.5	12.9	11.5	10.4	9.7
41	SMA13-3	100.0	100.0	90.2	74.8	33.2	19.9	16.6	13.5	11.8	10.5	9.7
42	SMA13-4	100.0	100.0	95.0	62.5	27.0	20.5	19.0	16.0	13.0	12.0	10.0
43	OGFC13-1	100.0	100.0	90.1	61.5	28.3	20.6	17.4	14.3	10.3	8.2	5.3
44	OGFC13-2	100.0	100.0	90.1	61.3	23.4	14.8	12.7	10.8	8.3	7.0	5.0
45	OGFC13-3	100.0	100.0	90.1	61.1	18.5	9.0	8.1	7.5	6.3	5.7	4.8
46	OGFC13-4	100.0	100.0	100.0	80.0	30.0	22.0	18.0	15.0	12.0	8.0	6.0
47	OGFC13-5	100.0	100.0	90.0	60.0	12.0	10.0	6.0	4.0	3.0	3.0	2.0
48	OGFC13-6	100.0	100.0	97.5	75.0	25.5	19.0	15.0	12.3	9.8	6.8	5.0
49	OGFC13-7	100.0	100.0	92.5	65.0	16.5	13.0	9.0	6.8	5.3	4.3	3.0
50	OGFC13-8	100.0	100.0	95.0	70.0	21.0	16.0	12.0	9.5	7.5	5.5	4.0
51	OGFC13-9	100.0	100.0	90.7	76.0	26.9	12.4	9.9	7.6	6.3	5.5	4.9
52	OGFC13-10	100.0	100.0	90.7	76.0	28.5	14.0	11.0	8.2	6.6	5.6	4.9
53	OGFC13-11	100.0	100.0	90.2	74.8	29.9	15.7	12.1	8.7	6.9	5.7	4.9
54	AC10-1	100.0	100.0	100.0	99.0	59.5	38.2	26.9	16.5	11.0	7.3	5.3
55	AC10-2	100.0	100.0	100.0	99.1	63.6	42.3	29.6	18.0	11.7	7.6	5.4
56	AC10-3	100.0	100.0	100.0	99.2	67.6	46.4	32.3	19.4	12.5	7.9	5.4
57	AC10-4	100.0	100.0	100.0	95.0	60.0	44.0	32.0	22.0	16.0	11.0	5.0
58	SMA10-1	100.0	100.0	100.0	98.3	40.8	24.0	19.3	14.9	12.5	10.8	9.8
59	SMA10-2	100.0	100.0	100.0	98.4	44.0	27.2	21.5	16.1	13.1	11.1	9.8
60	SMA10-3	100.0	100.0	100.0	98.5	47.3	30.5	23.7	17.2	13.7	11.3	9.9
61	SMA10-4	100.0	100.0	100.0	95.0	44.0	26.0	20.0	17.0	14.0	12.0	10.0
62	OGFC10-1	100.0	100.0	100.0	99.2	51.6	15.0	10.5	7.9	6.5	5.5	4.9
63	OGFC10-2	100.0	100.0	100.0	99.2	53.6	18.9	13.2	9.3	7.2	5.8	5.0
64	OGFC10-3	100.0	100.0	100.0	99.0	51.9	21.0	14.8	10.2	7.7	6.0	5.0
65	OGFC10-4	100.0	100.0	100.0	92.0	31.0	27.0	19.0	14.0	10.0	8.0	5.0
66	OGFC10-5	100.0	100.0	100.0	95.0	60.0	16.0	12.0	9.0	7.0	5.0	5.0

It can be seen from Table 4 that the gradations cover different types of asphalt mixture and maximum nominal aggregate size, and there are big differences between gradations. By Marshall design method, the optimal asphalt contents are obtained according to JTGF40-2004 Technical Specification for Construction of Highway Asphalt Pavements. The gradation coverage is relatively wide, which can also ensure the accuracy of mesostructure parameters prediction.

2.3. Computed tomography scan and digital image processing method

The CT scanning equipment used for asphalt mixture samples is the micro focus industrial CT owned by Harbin Institute of Technology. Both cone beam scanning system and cone beam filter back projection reconstruction algorithm are applied to reconstruct the images. The scanning voltage/current is 195 kv/95μA.

Through CT scanning, the cross section images of asphalt mixture specimens can be obtained, and 400 images are extracted from each specimen. These images are the basis for extracting the

mesostructure parameters of asphalt mixture. At present, the most used method to extract the mesostructure of asphalt mixture is digital image processing technology. The research team also proposed an image processing method [4] used in this research, as shown in the Figure 1.

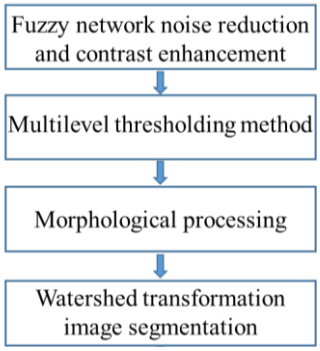
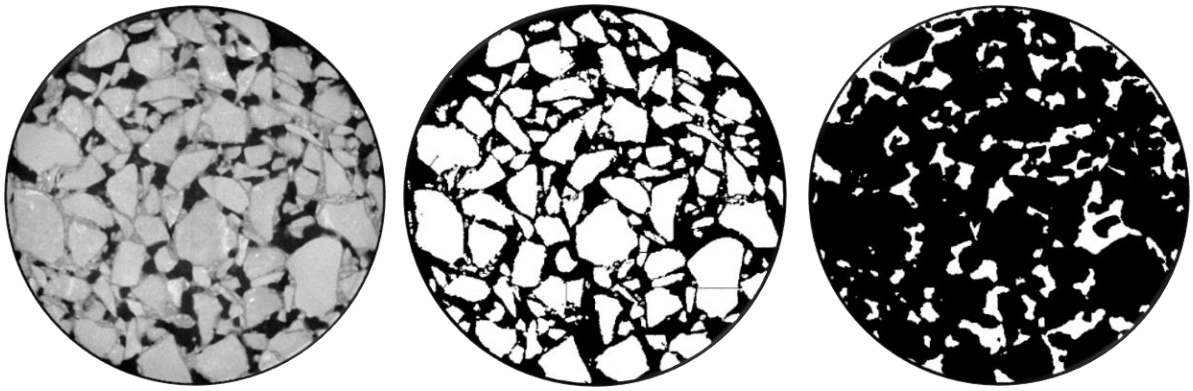


Figure 1. Image processing for CT images.

3. Methodology of mesostructure parameters extraction and modelling

Typical asphalt mixture images are shown in the Figure 2(a). The size of the image is 1000 pixels × 1000 pixels. It can be seen that asphalt mixture mainly includes aggregates and voids on the meso scale as shown in Figure 2(b),(c). The simple mesostructure parameters include the area of voids and aggregates, and the complex parameters include the shape of voids and aggregate contact. Based on the summary of existing literature and previous research results of our research group, 14 void parameters and 12 aggregate parameters in meso scale are list in this paper. These mesostructure parameters are directly related to the gradation of asphalt mixture, and the void and aggregate parameters will be introduced in detail below.



(a) CT image (b) Aggregate (c) Void

Figure 2. CT image and mesostructured.

3.1. Mesostructure parameters of void

1) Number of voids

Void area

The number of all pixels in the void position in the image, and the unit need to be converted to metric system as mm^2.

2) Ratio of void area to cross-sectional area of specimen

The sum of the areas of all voids in the image can be divided by the cross-sectional area of the specimen to obtain the parameter.

$$Ratio_{void} = \frac{\sum_{i=1}^{p} Area_i}{S} \tag{1}$$

where $Area_i$ is area of the i^{th} void area, p is the number of voids, S is the cross-sectional area of the specimen.

3) Centroid coordinates of void

The abscissa of the centroid coordinate of the void is the average value of the abscissas of all pixels constituting the void area, and the ordinate calculation method is the same as above. The calculation methods are shown as Eqs (2) and (3).

$$X_i = \sum_{j=1}^{n_i} x_{i,j} \Big/ n_i \tag{2}$$

$$Y_i = \sum_{j=1}^{n_i} y_{i,j} \Big/ n_i \tag{3}$$

where X_i and Y_i are the centroid coordinates of the $NO.i$ void area, n_i is the number of pixels in the i^{th} void area, $x_{i,j}$ and $y_{i,j}$ are the abscissa and ordinate of the j^{th} pixel in the i^{th} void area.

4) Equivalent circle radius of void

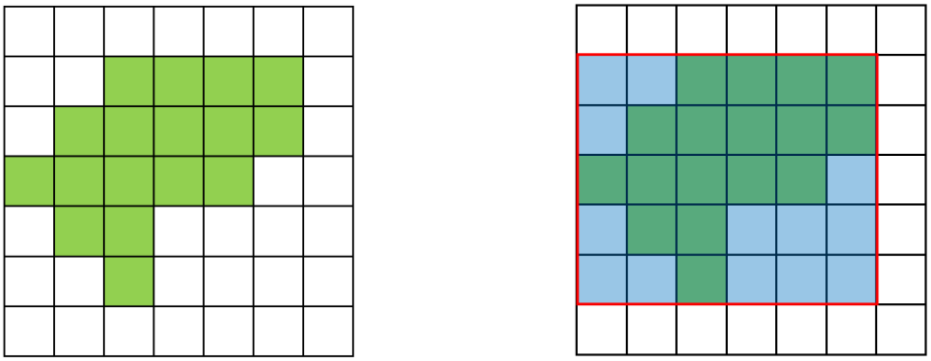
First, the area of the void should be determined, and then the area is used to deduce the radius of the circle with the same area as the void, as shown in Eq (4).

$$R_i = \sqrt{Area_i / \pi} \tag{4}$$

where Ri is the equivalent circle radius of the i^{th} void area, $Area_i$ is the area of the i^{th} void.

5) Circumscribed rectangular area and ratio of the rectangular area to void area

For getting the circumscribed rectangle of the closed area, the abscissa and ordinate values of all pixels in the void area should be extracted, and then the circumscribed rectangle can be determined by the maximum and minimum values of abscissa and ordinate, as shown in Figure 3.



(a) Void area (b) Circumscribed rectangular area

Figure 3. Circumscribed rectangular of the closed area.

After the circumscribed rectangular is determined, the calculation formulas of circumscribed rectangular area and ratio of the rectangular area to void area are shown in Eq (5) and (6).

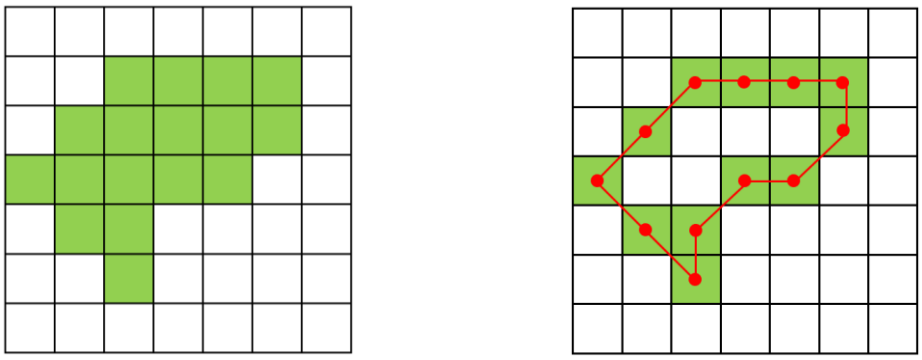
$$Area_{i,rectangle} = (x_{i,\max} - x_{i,\min}) \cdot (y_{i,\max} - y_{i,\min}) \tag{5}$$

$$Ratio_{i,rectangle} = Area_{i,rectangle} / Area_i \tag{6}$$

where $Area_{i,rectangle}$ is the circumscribed rectangular area of the i^{th} void area, $x_{i,max}$, $x_{i,min}$ are the maximum and minimum abscissa values of pixels in the i^{th} void area, $y_{i,max}$, $y_{i,min}$ are the maximum and minimum ordinate values of pixels in the i^{th} void area, $Ratio_{i,rectangle}$ is the ratio of the circumscribed rectangular area to the i^{th} void area.

6) Perimeter of void

The calculation of void perimeter is based on the identification of the edge of the closed area. After identifying the edge of the closed area, the connecting line length of the pixel center of the edge pixel is the perimeter of the void, as shown in Figure 4. The calculation method is shown as Eq (7).



(a) Void area (b) Edge of closed area

Figure 4. Perimeter calculation method.

$$C_i = \sum_{k=1}^{m_i} \sqrt{(x_{i,k} - x_{i,k+1})^2 + (y_{i,k} - y_{i,k+1})} \tag{7}$$

where Ci is the perimeter of the i^{th} void area, mi is the number of edge pixels of the i^{th} void area, $x_{i,k}$, $y_{i,k}$, $x_{i,k+1}$, $y_{i,k+1}$ are the abscissa and ordinate values of kth and k+1 th pixels on the edge of the i^{th} void area.

7) Center of void edge

The edge is divided into mi segments. The abscissa and ordinate of the starting point of each segment are weighted by its segment length. Finally, and then it is divided by the edge perimeter to calculate the center of void edge. The calculation formula is shown in Eqs (8) and (9).

$$X_{i,edge} = \sum_{k=1}^{m_i} x_{i,k} \cdot \sqrt{(x_{i,k} - x_{i,k+1})^2 + (y_{i,k} - y_{i,k+1})^2} \Big/ C_i \tag{8}$$

$$Y_{i,edge} = \sum_{k=1}^{m_i} y_{i,k} \cdot \sqrt{(x_{i,k} - x_{i,k+1})^2 + (y_{i,k} - y_{i,k+1})^2} \Big/ C_i \tag{9}$$

where $X_{i,edge}$ and $Y_{i,edge}$ are the abscissa and ordinate of the edge center of the i^{th} void area, mi is the number of edge pixels of the i^{th} void area, $x_{i,k}$ and $y_{i,k}$ are the abscissa and ordinate of the kth pixel on the edge of the i^{th} void area, $x_{i,k+1}$ and $y_{i,k+1}$ are the abscissa and ordinate of the $k+1^{th}$ pixel on the edge of the i^{th} void area.

8) Average distance from void edge to center of void edge

Based on the center information of void edge, the average distance from the void edge to the center can be further calculated, as shown in Eq (10).

$$D_{i,edge} = \left\{ \sum_{k=1}^{m_i} \left(\sqrt{(x_{i,k} - X_{i,edge})^2 + (y_{i,k} - Y_{i,edge})^2} \cdot \Delta s_{i,k} \right) \right\} \Big/ C_i \tag{10}$$

where $D_{i,edge}$ is average distance from void edge to center, $\Delta s_{i,k}$ is the distance between the kth edge point and the k+1st edge point of the i^{th} void area.

9) The variance of distance between void edge to center of void edge

The parameter can be calculated by Eq (11), and it can be used to characterize the dispersion of edge pixels.

$$Variance_{i,edge} = \left\{ \sum_{k=1}^{m_i} \left(\left((x_{i,k} - X_{i,edge})^2 + (y_{i,k} - Y_{i,edge})^2 \right) \cdot \Delta s_{i,k} \right) \right\} \Big/ C_i \tag{11}$$

10) Gaussian descriptors of voids

Gaussian descriptors are invariant to translation, transformation and scaling. The calculation formulas of Gaussian descriptors are shown as Eqs (12) and (13).

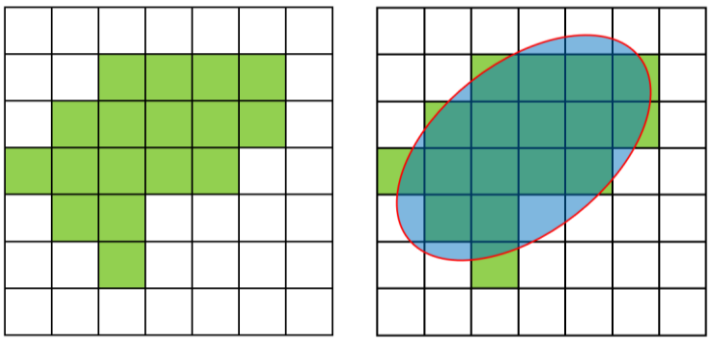
$$f(\theta, \lambda)_i = \frac{\sum_{k=1}^{m_i} e^{\frac{(x_{i,k} - X_{i,edge} - \lambda \cos\theta \cdot D_{i,edge})^2 + (y_{i,k} - Y_{i,edge} - \lambda \sin\theta \cdot D_{i,edge})^2}{2 \times Variance_{i,edge}}} \cdot \Delta s_{i,k}}{C_i} \tag{12}$$

$$I(\lambda)=\frac{\sum_{l=1}^{K} f(\frac{2\pi}{K}\cdot l,\lambda)}{K} \tag{13}$$

where $I(\lambda)$ is Gaussian descriptors, $f(\theta,\lambda)$ is Gaussian potential function. K is the number of copies divided equally at 360 degrees, and K is generally taken as 3–8. In order to extract the difference of details as much as possible, $K = 8$ is taken in this study. In order to extract enough information to characterize shape features, λ takes 8 different values, namely: $0.25D_{i,edge}$, $0.5\ D_{i,edge}$, $0.75\ D_{i,edge}$, $D_{i,edge}$, $1.25\ D_{i,edge}$, $1.5\ D_{i,edge}$, $1.75D_{i,edge}$, $2\ D_{i,edge}$.

11) Length of major and minor axes of ellipse with the same standard second-order central moment as the void

The relation between void area and ellipse is shown in Figure 5. The calculation formula for the parameter is shown in Eqs (14)–(17).



(a) Void area (b) Ellipse

Figure 5. Ellipse determination method.

$$u_{xx}=\frac{\sum_{j=1}^{n_i}(x_{i,j}-X_i)^2}{n_i} \tag{14}$$

$$u_{yy}=\frac{\sum_{j=1}^{n_i}(y_{i,j}-Y_i)^2}{n_i} \tag{15}$$

$$u_{xy}=\frac{\sum_{j=1}^{n_i}(x_{i,j}-X_i)\cdot(y_{i,j}-Y_i)}{n_i} \tag{16}$$

$$\begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}=\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}=\begin{bmatrix} u_{yy} & u_{xy} \\ u_{xy} & u_{xx} \end{bmatrix} \tag{17}$$

where u_{xx}, u_{yy} are the variances of abscissa and ordinate of void area, u_{xy} is the covariance of abscissa and ordinate of void area. λ_1 and λ_2 are eigenvalues. a and b are the lengths of major and minor axes of ellipse.

12) Eccentricity of ellipse with the same standard second-order central moment as the void

To calculate the eccentricity of the ellipse, the half focal length of the ellipse should be calculated firstly. The calculation method is shown as Eq (18) and (19).

$$c^2 = a^2 - b^2 \tag{18}$$

$$e = c/a \tag{19}$$

where c is the half focal length of ellipse, e is the eccentricity of ellipse.

13) Area of convex polygon surrounding void

Firstly, the minimum convex polygon surrounding the void area should be determined, and then the number of pixels in the convex polygon can be obtained. The calculation method is shown in Figure 6.

(a) Void area (b) Minimum convex polygon surrounding void area

Figure 6. Area of convex polygon surrounding void.

14) The ratio of void area to the number of pixels at the edge of the void

$$Ratio_{i,edge} = n_i / m_i \tag{20}$$

where $Ratio_{i,edge}$ is the ratio of the i^{th} void area to the number of pixels at the edge of the i^{th} void.

3.2. Mesostructure parameters of aggregate

Before calculating the aggregate parameters, it is necessary to classify the aggregate size. The width of the smallest circumscribed rectangle of the aggregate is taken as the size of the aggregate. In the CT image of asphalt mixture, it is obvious that the voids are black, and the aggregate and asphalt are gray. Therefore, it is easy to extract voids by digital image processing method. However, for the separation of aggregate and asphalt, it is difficult to distinguish the aggregates smaller than 0.6 mm

from asphalt. For the current CT scanning accuracy and digital image processing algorithm, only the aggregate larger than 0.6 mm can be accurately extracted. Since it is difficult to identify aggregates smaller than 0.6 mm by image processing, the aggregates sizes are classified as: greater than 19 mm, 19–16 mm, 16–13.2 mm, 13.2–9.5 mm, 9.5–4.75 mm, 4.75–2.36 mm, 2.36–1.18 mm and 1.18–0.6 mm.

Based on the aggregate classification above, the following aggregate parameters are calculated for each grade of aggregate.

1) Number of each grade aggregate

2) Proportion of quantity of each grade aggregate in the total

3) Total area of aggregate

4) Area of each grade aggregate

5) Proportion of area of each grade aggregate in the total

6) Mean value of average distance between aggregate centroids for each grade aggregates

The parameters can be calculated by Eq (21).

$$\begin{cases} D_{mean,i} = \dfrac{\sum_{j=1}^{n-1}\left[\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\right]}{n-1} \\ D_{mean} = \dfrac{\sum_{i=1}^{n} D_{mean,i}}{n} \end{cases} \tag{21}$$

where $D_{mean,i}$ is the average value of the distance between the centroid of i^{th} aggregate and the other aggregates for the same grade. D_{mean} is mean value of average distance between aggregate centroids for each grade aggregates. n is the number of aggregates for a certain grade.

7) Mean value of minimum distance between aggregate centroids for each grade

This parameter can evaluate the local distribution of aggregates in the mixture, and the parameters can be calculated by Eq (22).

$$\begin{cases} D_{\min,i} = \sum_{j=1}^{n-1} \min\left[\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\right] \\ D_{\min} = \dfrac{\sum_{i=1}^{n} D_{\min,i}}{n} \end{cases} \tag{22}$$

where $D_{min,i}$ is the minimum distance between the centroid of i^{th} aggregate and the other aggregates for the same grade. D_{mean} is mean value of minimum distance between aggregate centroids for each grade aggregates. n is the number of aggregates for a certain grade.

8) Mean value of the minimum distance between the edges of aggregates for the same grade

The parameter is calculated by subtract the sum of the equivalent circle radius of the two aggregates from the distance between the two aggregate centroids.

$$\begin{cases} D_{edge,i} = \sum_{j=1}^{n-1} \min\left[\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} - r_i - r_j\right] \\ D_{edge} = \dfrac{\sum_{i=1}^{n} D_{edge,i}}{n} \end{cases} \tag{23}$$

where $D_{edge,i}$ is the minimum distance between the edge of i^{th} aggregate and that of the other aggregate for the same grade. D_{edge} is the average of $Dedge,i$ of all the aggregate. r_i is the equivalent circle radius of the i^{th} aggregate. r_j is the equivalent circle radius of the j^{th} aggregate. n is the number of aggregates for a certain grade.

9) Area ratio of different grade aggregates

This parameter is mainly used to characterize the distribution state of small aggregates around a large aggregate. A circle is draw with the centroid of aggregate *i* and radius *Dmin*. m aggregates smaller than the grade which aggregate *i* belongs to is counted. The calculation method is shown in Eq (24).

$$\begin{cases} S_{small} = \sum_{j=1}^{m} S_j \\ S_i = \dfrac{S_{small}}{Area_i} \\ S_{ratio} = \dfrac{\sum_{i=1}^{n} S_i}{n} \end{cases} \tag{24}$$

where S_j is the area of the j^{th} small aggregate within D_{min} around aggregate *i*, S_{small} is the total area of the small aggregates, $Area_i$ is the area of aggregate *i*, Si is the area ratio of aggregate *i*, S_{ratio} is the average of Si of all the n aggregate in a certain grade. n is the number of aggregates for a certain grade.

10) Number of small aggregate in the range of minimum distance between centroids

The parameter is the number of aggregates in the circle drawn in the previous step, and it can characterize the distribution of small aggregates around a large aggregate.

11) Ratio of aggregate area to cross-sectional area of specimen

The sum of the areas of all aggregates in the image can be divided by the cross-sectional area of the specimen to obtain the ratio.

$$Ratio_{aggregate} = \frac{\sum_{i=1}^{n} Area_i}{S} \tag{25}$$

where $Ratio_{aggregate}$ is the ratio of aggregate area to cross-sectional area of specimen. n is the number of aggregates in the image. S is the cross-sectional area of specimen.

12) Ratio of asphalt mortar area to cross-sectional area of specimen

The mortar area is equal to the cross section area minus the void area and aggregate area, and then the parameter can be calculated.

3.3. Sampling and featurization

For each gradation, 800 images can be obtained from the two specimens, and the average value and standard deviation of the meso structure parameters of 800 image samples can be calculated. Before modeling, featurization needs to be carried out. The featurization is to find and screen better features to join the model to improve the performance of the model. The process is shown as follow:

1) Explained variable selection

2) Feature filtering based on standard deviation

The data of the feature pool is transformed to [-1,1], and the standard deviation of each feature in all samples is calculated. After repeated trials, it is found that when the standard deviation threshold is 0.4, the remaining features in the final screening can be roughly about 20. In order to simplify the calculation, the threshold is set to 0.4, and the features with standard deviation less than 0.4 are removed.

3) Feature transformation

Data transformation is carried out for the remaining features, that is, x is transformed into five forms: x, x^2, ln (x + 1), $x^{0.5}$ and 1/(x + 1). Because the relationship between features and explained variables is not necessarily linear, for establishing a linear model, the relationship between features and explained variables needs to be modified into a linear relationship. The absolute value of the correlation coefficient between the transformed data and the explained variable is used to judge the data transformation of the feature finally. The transformation form with the highest correlation coefficient will be retained. This step does not increase or reduce the features, but only changes the data form of the feature.

4) Screening based on correlation coefficient between features

To ensure the independence between features, the correlation coefficient matrix between features is calculated, and the absolute value of the matrix is obtained. The frequency of the absolute value of correlation coefficient between each feature and other features is greater than 0.6 is counted. The features with the highest frequency are removed and iterated successively until the highest frequency is less than or equal to 3.

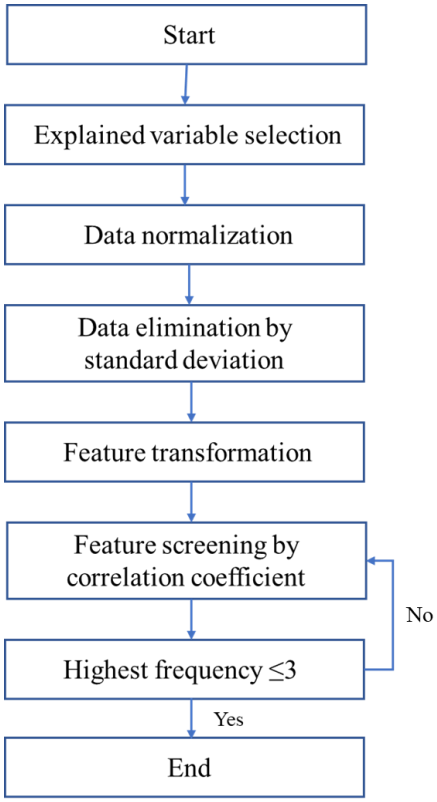


Figure 7. Flow chart of featurization.

Through the above featurization method, on the one hand, it can greatly reduce the interference of features with low correlation, on the other hand, it can improve the quality of features and lay a good foundation for the correlation models.

3.4. Gradation prediction model by mesostructure parameters

Forward stepwise regression method are utilized for predicting gradation through mesostructure with the P-value of 0.05. The reason why stepwise regression rather than machine learning method is adopted here is that if the neural network model is constructed, its input and output scale is in double digits, and the constructed network scale is too large compared with 64 groups of samples. It is not suitable to use machine learning method for modeling.

The essence of forward stepwise regression is OLS (Ordinary Least Square Method), but stepwise regression is a method combining multiple OLS. Its purpose is to find the most appropriate explanatory variable and add it to the equation. The specific steps are as follows:

1) Unary linear regression is performed between explanatory variable (X_1, X_2...) and explained variable (Y), and the F-statistics of each regression equation are calculated as (F^1_1, F^1_2...). For the maximum value F^1_i, the critical value is F^1 at the significance level α. If $F^1_i \geq F^1$, Xi is introduced into the equation.

2) Binary linear regression is performed between explanatory variable ((X_1, X_i), (X_2, X_i)...) and explained variable (Y), and the F-statistics of each regression equation are calculated as (F^2_1, F^2_2...). For the maximum value F^2_j, the critical value is F^2 at the significance level α. If $F^2_j \geq F^2$, X_j is

introduced into the equation.

3) More explanatory variables are introduced until the maximum F-statistic is less than the critical value, that is, the end of stepwise regression.

The explained and explanatory variables are shown in Tables 5 and 6.

In the modeling process, 8 grades of aggregates, 102 mesostructure parameters, 64 gradation (128 specimens) are utilized. The modeling process is as follows:

1) Firstly, based on the explained variables, the featurization is performed to screen the explanatory variables.

2) The filtered explanatory variables are imported into the stepwise regression model for modeling.

3) Breusch-Pagan-Godfrey (BPG) test is carried out on the established model to test whether there is heteroscedasticity problem.

4) If there is heteroscedasticity problem, weight should be calculated to modeling by the weighted least squares regression. Otherwise, the modeling is completed.

The modeling process are shown in Figure 8.

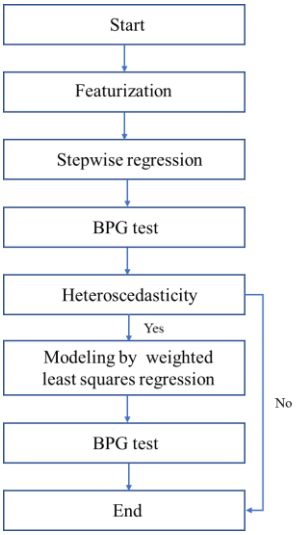


Figure 8. Flow chart of modeling.

Table 5. Explained variable.

Variable abbreviation	Variable name
$P_{19\ mm}$	Passing percentage on 19 mm sieve (%)
$P_{16\ mm}$	Passing percentage on 16 mm sieve (%)
$P_{13.2\ mm}$	Passing percentage on 13.2 mm sieve (%)
$P_{9.5\ mm}$	Passing percentage on 9.5 mm sieve (%)
$P_{4.75\ mm}$	Passing percentage on 4.75 mm sieve (%)
$P_{2.36\ mm}$	Passing percentage on 2.36 mm sieve (%)
$P_{1.18\ mm}$	Passing percentage on 1.18 mm sieve (%)
$P_{0.6\ mm}$	Passing percentage on 0.6 mm sieve (%)
CA	CA ratio (Bailey method)
FA	FA ratio (Bailey method)

Table 6. Explanatory variable.

Variable abbreviation	Variable name
AQ	Aggregates quantity for each grade
PAQ	Proportion of aggregate quantity of each grade in the total
TAA	Total aggregate area (mm^2)
AA	Aggregate area for each grade (mm^2)
PA	Proportion of area of each grade aggregate in the total
AD_{mean}	Mean value of average distance between aggregate centroids for each grade aggregates (mm)
AD_{min}	Minimum distance between aggregate centroids for each grade (mm)
SD_{mean}	Mean value of the shortest distance between the edges of aggregates for the same grade (mm)
AR	Area ratio of different grade aggregates
QSA	Quantity of small aggregate in the shortest distance range between centroids
LAE_{major}	Length of major axes of ellipse with the same standard second-order central moment as the void (mm)
LAE_{minor}	Length of minor axes of ellipse with the same standard second-order central moment as the void (mm)
EE	Eccentricity of ellipse with the same standard second-order central moment as the void
ACP	Area of convex polygon surrounding void (mm^2)
PAV	Proportion of the area where the void coincides with the convex polygon
VA	Void area (mm^2)
CRA	Circumscribed rectangular area (mm^2)
RCR	Ratio of circumscribed rectangular area to enclosed area
PV	Perimeter of void (mm)
ECR	Equivalent circle radius of void (mm)
VDV	The variance of distance between void edge to center of void edge
ADV	Average distance from void edge to center of void edge (mm)
NV	Number of voids
$GD_{0.25}$	Gaussian descriptors ($\lambda = 0.25$)
$GD_{0.5}$	Gaussian descriptors ($\lambda = 0.5$)
$GD_{0.75}$	Gaussian descriptors ($\lambda = 0.75$)
GD_1	Gaussian descriptors ($\lambda = 1$)
$GD_{1.25}$	Gaussian descriptors ($\lambda = 1.25$)
$GD_{1.5}$	Gaussian descriptors ($\lambda = 1.5$)
$GD_{1.75}$	Gaussian descriptors ($\lambda = 1.75$)
GD_2	Gaussian descriptors ($\lambda = 2$)
RMA	Ratio of asphalt mortar area to cross-sectional area of specimen (%)
RAA	Ratio of aggregate area to cross-sectional area of specimen (%)
RVA	Ratio of void area to cross-sectional area of specimen (%)

Since the process of modeling the passing percentage of each sieve are similar, the modeling of the passing percentage on the 1.18 mm sieve is selected for detailed introduction, and the rest

are briefly displayed.

4. Results and discussion

4.1. Gradation prediction model by mesostructure parameters

4.1.1. Modeling of Passing percentage on 1.18 mm sieve

The explained variable is selected as the passing percentage on 1.18 mm sieve ($P_{1.18}$), and the featurization is carried out on the feature pool. Finally, there are 20 remaining features. All features are normalized, and stepwise regression is carried out to obtain the model:

$$P_{1.18mm} = \frac{4.41}{ACP+1} + 3.92\sqrt{NV} + 1.77AD^2_{\min(9.5-4.75\text{mm})} + \frac{1.81}{AD_{\min(1.18-0.6\text{mm})}+1} + 9.43 \tag{26}$$

When the significance level is 0.05, the adjoint probability of F-statistic is less than 0.05, that is, the equation is significant.

For the BPG test, the OLS regression is performed on the logarithm of the square term of the residual and the independent variable, and the pro (F-statistic) is 0.023. Under the significance level of 0.05, the original assumption that the equation is not significant is rejected, that is, there is a heteroscedasticity.

In order to solve the heteroscedasticity problem, the original independent variables are retained, and the weighted least square method is used for regression. Firstly, the weight of each independent variable is calculated. After trial calculation, it is found that the following formula has high goodness of fit, $R^2 = 0.335$.

$$\begin{aligned}\ln(e^2) = &-\frac{2}{ACP+1} + 2.52\sqrt{NV} - 0.06AD^2_{\min(9.5-4.75\text{mm})} - \frac{2.67}{AD_{\min(1.18-0.6\text{mm})}+1} - \frac{2.42\sqrt{NV}}{ACP+1} \\ &- \frac{3.83AD^2_{\min(9.5-4.75\text{mm})}}{ACP+1} - \frac{2.47}{(ACP+1)(AD_{\min(1.18-0.6\text{mm})}+1)} - 0.52\end{aligned} \tag{27}$$

$$weight = 1/\sqrt{e^{-\frac{2}{ACP+1}+2.52\sqrt{NV}-0.06AD^2_{\min(9.5-4.75\text{mm})}-\frac{2.67}{AD_{\min(1.18-0.6\text{mm})}+1}-\frac{2.42\sqrt{NV}}{ACP+1}-\frac{3.83AD^2_{\min(9.5-4.75\text{mm})}}{ACP+1}-\frac{2.47}{(ACP+1)(AD_{\min(1.18-0.6\text{mm})}+1)}-0.52}}$$

The weight is calculated by above formula, and the weighted least square method is used for regression. The new model is shown as Eq (28).

$$P_{1.18mm} = \frac{4.82}{ACP+1} + 3.54\sqrt{NV} + 1.57AD^2_{\min(9.5-4.75\text{mm})} + \frac{1.82}{AD_{\min(1.18-0.6\text{mm})}+1} + 9.54 \tag{28}$$

The BPG test is performed again, and it is found that the adjoint probability of F-statistic is 0.635. Under the significance level of 0.05, the original hypothesis is accepted, that is, there is no heteroscedasticity.

4.1.2. Modeling of Passing percentage for other sieves

1) Modeling of Passing percentage on 19 mm sieve

$$P_{19mm} = 3.10\sqrt{AR_{19-16mm}} - 0.03RMA^2 - \frac{0.02}{AD_{mean(2.36-1.18mm)}+1} + 3.09 \tag{29}$$

2) Modeling of Passing percentage on 16 mm sieve

$$P_{16mm} = 5.97\sqrt{PAQ_{16-13.2mm}} + 1.32\sqrt{AA_{19-16mm}} - \frac{0.55}{AD_{\min(4.75-2.36\text{mm})}+1} - 0.67AD^2_{\min(1.18-0.6\text{mm})} - \frac{0.57}{RMA+1} - \frac{0.21}{RAA+1} + 6.87 \tag{30}$$

3) Modeling of Passing percentage on 13.2 mm sieve

$$P_{13.2mm} = -4.68AD^2_{\min(16-13.2\text{mm})} - 4.25AR_{13.2-9.5mm} - \frac{2.79}{LAE_{\text{minor}}+1} + \frac{2.38}{RMA+1} - 2.68AD^2_{mean(1.18-0.6\text{mm})} + 8.63 \tag{31}$$

4) Modeling of Passing percentage on 9.5 mm sieve

$$P_{9.5mm} = -\frac{231.65}{AA_{13.2-9.5mm}+1} - \frac{3.18}{GD_{1.25}+1} - 7.06AA^2_{16-13.2mm} + \frac{225.60}{QSA_{13.2-9.5mm}+1} - \frac{6.24}{RAA+1} - \frac{2.75}{AD_{\text{mean}(9.5-4.75\text{mm})}+1} + 1.34GD^2_{0.25} + 5.02 \tag{32}$$

5) Modeling of Passing percentage on 4.75 mm sieve

$$P_{4.75mm} = 12.64PA^2_{9.5-4.75mm} - 22.95PAQ^2_{1.18-0.6mm} + 4.84SD_{mean(13.2-9.5\text{mm})} + 15.02AR^2_{2.36-1.18mm} + 10.09\sqrt{PA_{2.36-1.18mm}} - 5.39\sqrt{NV} - 2.25\sqrt{RAA} + 45.07 \tag{33}$$

6) Modeling of Passing percentage on 2.36 mm sieve

$$P_{2.36mm} = 4.37\sqrt{NV} + \frac{3.85}{AA_{13.2-9.5mm}+1} - \frac{8.22}{AR_{4.75-2.36mm}+1} + \frac{5.65}{AD_{mean(9.5-4.75\text{mm})}+1} + \frac{3.28}{RAA+1} + 21.26 \tag{34}$$

7) Modeling of Passing percentage on 0.6 mm sieve

$$P_{0.6mm} = -4.57PAQ_{4.75-2.36mm} + 3.93NV + \frac{1.63}{GD_{1.25}+1} + 1.60AA^2_{4.75-2.36mm} + \frac{0.77}{AA_{16-13.2mm}+1} + 7.45 \tag{35}$$

8) Modeling of CA ratio

$$CA = -\frac{0.14}{AA_{16-13.2mm}+1} + 0.1TAA + \frac{0.08}{TAA+1} - \frac{0.05}{AD_{mean(13.2-9.5\text{mm})}+1} + \frac{0.03}{RAA+1} - 0.09PAQ^2_{4.75-2.36mm} + 0.05GD^2_{1.25} - 0.05AD^2_{\text{mean}(9.5-4.75mm)} + 0.32 \tag{36}$$

9) Modeling of FA ratio

$$FA = -\frac{0.11}{AA_{16-13.2mm}+1} - \frac{0.001}{RMA+1} + 0.11 \tag{37}$$

4.2. Mesostructure parameter prediction model by gradation

Because there are too many mesostructure parameters, modeling all the parameters will make the article lengthy. Therefore, the average distance between aggregates of 9.5–4.75 mm is selected for modeling in this study.

There are 800 mesostructure images for each gradation, therefore the mesostructure indexes predicted by gradation are a range instead of a fixed value. The difference is characterized by mean and standard deviation. The explained variables and explanatory variables are shown in Tables 7 and 8.

Table 7. Explained variable.

Variable abbreviation	Variable name
$AD_{mean(9.5–4.75\ mm)}$	Mean value of average distance between aggregate centroids for 9.5–4.75 mm grade aggregates (mm)
SD_{AD}	Standard deviation of $AD_{mean(9.5–4.75\ mm)}$

Table 8. Explanatory variable.

Variable number	Variable name
P_{19mm}	Passing percentage on 19 mm sieve (%)
P_{16mm}	Passing percentage on 16 mm sieve (%)
$P_{13.2mm}$	Passing percentage on 13.2 mm sieve (%)
$P_{9.5mm}$	Passing percentage on 9.5 mm sieve (%)
$P_{4.75mm}$	Passing percentage on 4.75 mm sieve (%)
$P_{2.36mm}$	Passing percentage on 2.36 mm sieve (%)
$P_{1.18mm}$	Passing percentage on 1.18 mm sieve (%)
$P_{0.6mm}$	Passing percentage on 0.6 mm sieve (%)
CA	CA ratio (Bailey method)
FA	FA ratio (Bailey method)

(i) Modeling of AD_{mean}(9.5–4.75 mm)

Feature engineering is carried out on the feature pool, and 10 features are finally left. Normalization and stepwise regression are carried out for all features, and the following model is obtained:

$$AD_{mean(9.5-4.75\text{mm})} = -\frac{2.80}{P_{9.5mm}+1} + 1.74P_{1.18mm} - 1.50P_{19mm}^2 - \frac{2.41P_{19mm}^2}{P_{9.5mm}+1} + 1.17P_{1.18mm}P_{19mm}^2 + 44.68 \tag{38}$$

At the significance level of 0.05, the concomitant probability of F-statistic is less than 0.05, that is, the null hypothesis that all structural parameters are significantly 0 is rejected, and the equation is significant.

For the BPG test, OLS regression is performed on the square term of residuals and independent

variables. Significance test is performed on the equation to check whether heteroscedasticity existed. The test Pro(F-statistic) was 0.278. At the significance level of 0.05, the null hypothesis that the equation is not significant is accepted, that is, there is no heteroscedasticity.

(ii) Modeling of SD_{AD}

Feature engineering is carried out on the feature pool, and 10 features are finally left. Normalization and stepwise regression are carried out for all features, and the following result is obtained:

$$SD_{AD} = -0.96\ln(P_{4.75mm}+1) + \frac{0.374}{P_{2.36mm}+1} + 0.19\ln(P_{19mm}+1) + 2.62 \tag{39}$$

When the significance level is 0.05, the adjoint probability of F statistic is less than 0.05. The original assumption that all structural parameters are significant is 0 is rejected, that is, the equation is significant.

For the BPG test, OLS regression is performed on the square term of residual and independent variables. The equation significance test is performed on the equation to check whether there is heteroscedasticity problem. The F-statistic is 0.013. Under the significance level of 0.05, the original hypothesis that the equation is not significant is rejected, that is, there is heteroscedasticity.

In order to solve the heteroscedasticity problem, the original independent variables are retained, and the weighted least square method is used for regression. Firstly, the weight of each independent variable is calculated. After trial calculation, it is found that the following formula has high goodness of fit, $R^2 = 0.683$.

$$\begin{aligned} &\ln(e^2) = 4.20\ln(P_{4.75mm}+1) + \frac{2.55}{P_{2.36mm}+1} + 2.13\ln(P_{19mm}+1) + 5.08\ln(P_{4.75mm}+1)\ln(P_{19mm}+1) - 0.26 \\ &weight = 1/\sqrt{e^{4.20\ln(P_{4.75mm}+1)+\frac{2.55}{P_{2.36mm}+1}+2.13\ln(P_{19mm}+1)+5.08\ln(P_{4.75mm}+1)\ln(P_{19mm}+1)-0.26}} \end{aligned} \tag{40}$$

The weight is calculated by the above formula, and the weighted least square method is used for regression to obtain:

$$SD_{AD} = -0.95\ln(P_{4.75mm}+1) + \frac{0.17}{P_{2.36mm}+1} + 0.18\ln(P_{19mm}+1) + 2.46 \tag{41}$$

The BPG test is performed again. It is found that the adjoint probability of F statistic is 0.415. Under the significance level of 0.05, the original hypothesis is accepted, that is, there is no heteroscedasticity.

4.3. Model verification

An example is added to demonstrate the use scenario of the above model. The model examples established in Sections 4.2 and 4.3 are demonstrated respectively. The calculation example of the mesostructure parameters prediction based on gradation adopts the model of predicting the passing percentage of the 4.75 mm sieve. For the gradation prediction, the mean value of average distance between aggregate centroids for 9.5–4.75 mm aggregates is used as an example.

4.3.1. Prediction of the passing percentage of the 4.75 mm sieve

Background: in the road detection process, it is necessary to obtain the gradation information of a section of the road.

1) Obtaining sectional image

The core sample of the pavement is cut to obtain 10 section images at different positions. The purpose is to make the average value of the calculated mesostructure parameters more stable, that is, the average value can better represent the state of the core sample.

2) Image processing and calculation of mesostructure parameters

The image processing method developed in this research is used to process 10 cross section images, and 10 void images and aggregate images are obtained respectively. These images are imported to the calculation program the obtain the mesostructure parameters, and 7 parameters are extracted from the mesostructure parameters system as Eq (33). The parameters are shown in Table 9.

Table 9. Mesostructure parameters for the 10 images.

Image number	$PA_{9.5\text{-}4.75\ mm}$	$PAQ_{1.18\text{-}0.6\ mm}$	$SD_{mean(13.2\text{-}9.5\ mm)}$	$AR_{2.36\text{-}1.18\ mm}$	$PA_{2.36\text{-}1.18\ mm}$	NV	RAA
1	0.0083	0.2009	12.9817	0.1608	0.0012	162.0000	74.4998
2	0.0070	0.3203	9.7246	0.3150	0.0014	171.0000	72.8321
3	0.0069	0.2118	10.0213	0.2204	0.0013	149.0000	75.4631
4	0.0077	0.2516	7.3240	0.2588	0.0015	189.0000	74.4665
5	0.0076	0.2253	10.2477	0.2218	0.0012	192.0000	74.0706
6	0.0061	0.2419	4.9142	0.1628	0.0018	200.0000	73.2175
7	0.0086	0.2281	7.7761	0.2485	0.0010	182.0000	74.5839
8	0.0059	0.2260	6.6271	0.2882	0.0014	131.0000	75.0626
9	0.0067	0.2313	11.3000	0.2069	0.0014	187.0000	73.0016
10	0.0049	0.2026	15.0923	0.1748	0.0013	221.0000	75.4926
Mean	0.0070	0.2340	9.6009	0.2258	0.0014	178.4000	74.2690

3) Data preprocessing

Based on the prediction model, the data transformation is performed. According to the maximum and minimum value of the mesostructure parameter database, the data is normalized. The preprocessed date is shown in Table 10.

Table 10. Preprocessed data.

Data transformation	Data before transformation	Transformed data	Normalized data
$PA_{9.5\text{-}4.75\ mm}{}^{2}$	0.0070	0.0000	-0.1538
$PAQ_{1.18\text{-}0.6\ mm}{}^{2}$	0.2340	0.0547	-0.2566
$SD_{mean(13.2\text{-}9.5\ mm)}$	9.6009	9.6009	-0.9688
$AR_{2.36\text{-}1.18\ mm}{}^{2}$	0.2258	0.0510	-0.5852
$PA_{2.36\text{-}1.18\ mm}{}^{1/2}$	0.0014	0.0369	-0.0534
$NV^{1/2}$	178.4000	13.3566	-0.1103
$RAA^{1/2}$	74.2690	5515.8910	0.7594

4) Calculation of prediction results

The above values are substituted into Eq (33) to calculate the passing percentage of the 4.75 mm sieve as shown in Eq (42).

$$\begin{aligned} P_{4.75mm} &= 12.64PA_{9.5-4.75mm}^2 - 22.95PAQ_{1.18-0.6\text{mm}}^2 + 4.84SD_{mean(13.2-9.5\text{mm})} + 15.02AR_{2.36-1.18mm}^2 \\ &+10.09\sqrt{PA_{2.36-1.18mm}} - 5.39\sqrt{NV} - 2.25\sqrt{RAA} + 45.07 \\ &12.64\times(-0.1538) - 22.95\times(-0.2566) + 4.84\times(-0.9688) + 15.02\times(-0.5852) + 10.09\times(-0.0534) \\ &-5.39\times(-0.1103) - 2.25\times 0.7594 + 45.07 = 33.8841 \end{aligned} \tag{42}$$

Finally, the prediction passing percentage of 4.75 mm sieve is 33.8841%, while the actual value used in this example is 30%. The prediction error is acceptable.

4.3.2. Predicted average distance between 9.5–4.75 mm aggregate

Background: the gradation of asphalt mixture is known, and the corresponding mesostructure parameters should be analyzed.

1) Gradation data processing

Firstly, gradation is converted into the mass percentage on the sieve, and then the aggregate smaller than 0.6 mm is removed to calculate the mass percentage on the other sieves.

Table 11. Gradation data.

Variable abbreviation	Variable name	Value
$P_{19\ mm}$	Mass percentage on 19 mm sieve(%)	0
$P_{16\ mm}$	Mass percentage on 16 mm sieve(%)	0
$P_{13.2\ mm}$	Mass percentage on 13.2 mm sieve(%)	3.4
$P_{9.5\ mm}$	Mass percentage on 9.5 mm sieve(%)	21.6
$P_{4.75\ mm}$	Mass percentage on 4.75 mm sieve(%)	30.5
$P_{2.36\ mm}$	Mass percentage on 2.36 mm sieve(%)	24.2
$P_{1.18\ mm}$	Mass percentage on 1.18 mm sieve(%)	14.9
CA	CA ratio (Beret index)	5.4
FA	FA1 ratio (Beret index)	0.3

As in the previous example, the data transformation and normalization are performed. The data for mean value and standard deviation are shown in Tables 12 and 13.

Table 12. Data for mean value prediction.

Data transformation form	Data before transformation	Transformed data	Normalized data
$1/(P_{9.5mm}+1)$	21.6	0.0442	-0.9239
$P_{1.18mm}$	14.9	14.9000	0.5444
P_{19mm}^2	0	0.0000	-1.0000

Table 13. Data for standard deviation prediction.

Data transformation form	Data before transformation	Transformed data	Normalized data
$Ln(P_{4.75mm}+1)$	30.5	3.4500	0.0132
$1/(P_{2.36mm}+1)$	24.2	0.0397	-0.8714
$Ln(P_{19mm}+1)$	0	0.0000	-1.0000

2) Calculate the mean and standard deviation

The data in Table 12 is substituted into Eq (38) to calculate the mean value.

$$\begin{aligned} AD_{mean(9.5-4.75\text{mm})} &= -\frac{2.80}{P_{9.5mm}+1}+1.74P_{1.18mm}-1.50P_{19mm}^2-\frac{2.41P_{19mm}^2}{P_{9.5mm}+1} \\ &+1.17P_{1.18mm}P_{19mm}^2+44.68 \\ AD_{mean(9.5-4.75\text{mm})} &= -2.8\times(-0.9239)+1.74\times 0.5444-1.5\times(-1) \\ &-2.41\times(-1)/(-0.9293)+1.17\times 0.5444\times(-1)+44.68=46.8507 \end{aligned} \tag{43}$$

The data in Table 13 is substituted into Eq (39) to calculate the standard deviation:

$$\begin{aligned} SD_{AD} &= -0.95\ln(P_{4.75mm}+1)+\frac{0.17}{P_{2.36mm}+1}+0.18\ln(P_{19mm}+1)+2.46 \\ SD_{AD} &= -0.95\times 0.0132+0.17\times(-0.8714)+0.18\times(-1)+2.46=2.1194 \end{aligned} \tag{44}$$

The mean value is 46.8507 and the standard deviation is 2.1194.

3) Calculate the confidence interval of mesostructure parameters

Based on the t-distribution table, when the significance level is 0.1 and 800 samples, the value of t-statistic is 1.645. It is substituted into the following formula to calculate the confidence interval:

$$\begin{aligned} &\left[\overline{X}-t_{\frac{\alpha}{2}}(n-1)\frac{\sigma}{\sqrt{n}},\overline{X}+t_{\frac{\alpha}{2}}(n-1)\frac{\sigma}{\sqrt{n}}\right] \\ &[46.8507-1.645\times\frac{2.1194}{\sqrt{800}},46.8507+1.645\times\frac{2.1194}{\sqrt{800}}] \\ &=[46.7274,46.9739] \end{aligned} \tag{45}$$

The confidence interval for predicting the average distance between 9.5–4.75 mm aggregates is [46.7274, 46.9739].

In this paper, the prediction of mesostructure parameter and gradation are of great significance for future construction gradation identification and digital modeling. However, this paper also has some limitations, which are mainly reflected in the following two aspects:

1) At present, the gradation identification ability of asphalt mixture is limited, which makes it difficult to accurately identify aggregates with the diameter smaller than 1.18 mm;

2) Although we have 132 specimens by CT scan at present, it is still difficult to meet the requirements of accurate prediction. In the future, we will continue to expand the amount of data and improve the prediction accuracy of the model.

5. Conclusions

In this paper, the mesostructure images of asphalt mixtures with 64 gradations are obtained by industrial CT, and the meso structure parameters are extracted. The correlation models between gradation and mesostructure parameters are established and verified based on econometric method. The conclusions can be drawn as follows:

1) Based on digital image processing method, void and aggregate parameters are extracted from 64 types of asphalt mixtures. At the same time, gradation parameters are proposed, including percentages retaining on sieve, CA ratio and FA ratio. 102 mesostructure parameters and 10 gradation parameters are selected for modeling.

2) Featurization used in this paper including data augmentation, standard deviation screening and correlation coefficient screening. The feature quality after featurization is greatly improved and the number of explanatory variables of the model can be reduced.

3) According to featurization and stepwise regression, the gradation prediction models based on mesostructure parameters are established. The econometric hypothesis test is carried out on the models. For the models with heteroscedasticity, the weighted least square regression is used to solve the heteroscedasticity problem.

4) The passing percentage of the 4.75 mm sieve and the mean value of average distance between aggregate centroids for 9.5–4.75 mm aggregates are used as examples. The prediction error of passing percentage is acceptable, and the confidence interval of mesostructure parameters are also given.

Acknowledgments

This paper was supported by National Key Research and Development Program of China (No. 2022YFB2602600), National Natural Science Foundation of China (No. 51908168), Postdoctoral Science Foundation of China (No. 2019M651192 and No. 2020T130150), Heilongjiang Postdoctoral Fund (No. LBH-Z19163). All data, models, and code generated or used during the study appear in the published article.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. T. Ma, D. Zhang, Y. Zhang, S. Wang, X. Huang, Simulation of wheel tracking test for asphalt mixture using discrete element modelling, *Road Mater. Pavement Des.*, **19** (2018), 367–384. https://doi.org/10.1080/14680629.2016.1261725
2. C. Zhou, M. Zhang, Y. Li, J. Lu, J. Chen, Influence of particle shape on aggregate mixture's performance: DEM results, *Road Mater. Pavement Des.*, **20** (2019), 399–413. https://doi.org/10.1080/14680629.2017.1396236
3. R. Cao, Y. Zhao, Y. Gao, X. Huang, L. Zhang. Effects of flow rates and layer thicknesses for aggregate conveying process on the prediction accuracy of aggregate gradation by image segmentation based on machine vision, *Constr. Build. Mater.*, **222** (2019), 566–578. https://doi.org/10.1016/j.conbuildmat.2019.06.147

4. C. Xing, H. Xu, Y. Tan, X. Liu, C. Zhou, T. Scarpas, Gradation measurement of asphalt mixture by X-Ray CT images and digital image processing methods, *Measurement.*, **132** (2019), 377–386. https://doi.org/10.1016/j.measurement.2018.09.066
5. X. Yao, H. Xu, T. Xu, Void distribution, interfacial adhesion and anti-cracking mechanisms of cold recycled asphalt mixture based on AFM and X-ray CT., *Appl. Surf. Sci.*, **606** (2022), 155012. https://doi.org/10.1016/j.apsusc.2022.155012
6. X. Yao, H. Xu, T. Xu, Mechanical properties and enhancement mechanisms of cold recycled mixture using waterborne epoxy resin/styrene butadiene rubber latex modified emulsified asphalt, *Constr. Build. Mater.*, **352** (2022), 129021. https://doi.org/10.1016/j.conbuildmat.2022.129021
7. M. Guo, X. Yin, X. Dun, Y. Tan, Effect of aging, temperature and relative humidity on adhesion between asphalt binder and mineral aggregate, *Constr. Build. Mater.*, **363** (2023), 129775. https://doi.org/10.1016/j.conbuildmat.2022.129775
8. M. Guo, M. Liang, A. Screeram, A. Bhasin, D. Luo, Characterisation of rejuvenation of various modified asphalt binders based on simplified chromatographic techniques, *Int. J. Pavement Eng.*, **23** (2022), 4333–4343. https://doi.org/10.1080/10298436.2021.1943743
9. L. Wang, J. Frost, N. Shashidhar, Microstructure study of WesTrack mixes from X-ray tomography images, *Transp. Res. Record*, **1767** (2001), 85–94.
10. E. Masad, V. Jandhyala, N. Dasgupta, N. Somadevan, N. Shashidhar, Characterization of air void distribution in asphalt mixes using X-ray computed tomography, *J. Mater. Civ. Eng.*, **14** (2002), 122–129. https://doi.org/10.1061/(ASCE)0899-1561(2002)14:2(122)
11. H. Xu, W. Guo, Y. Tan, Internal structure evolution of asphalt mixtures during freeze-thaw cycles. *Mater. Des.*, **86** (2015), 436–446. https://doi.org/10.1016/j.matdes.2015.07.073
12. H. Xu, C. Xing, H. Zhang, H. Li, Y. Tan, Moisture seepage in asphalt mixture using X-ray imaging technology, *Int. J. Heat Mass Transf.*, **131** (2019), 375–384. https://doi.org/10.1016/j.ijheatmasstransfer.2018.11.081
13. E. Arambula, E. Masad, A. E. Martin, Influence of air void distribution on the moisture susceptibility of asphalt mixes, *J. Mater. Civ. Eng.*, **19** (2007), 655–664. https://doi.org/10.1061/(ASCE)0899-1561(2007)19:8(655)
14. W. Jiang, A. Sha, J. Xiao, Experimental study on relationships among composition, microscopic void features, and performance of porous asphalt concrete. *J. Mater. Civ. Eng.*, **27** (2015), 11. https://doi.org/10.1061/(ASCE)MT.1943-5533.0001281
15. N. Wang, F. Chen, T. Ma, Y. Luan, J. Zhu, Compaction performance of cold recycled asphalt mixture using smartRock sensor. *Autom. Constr.*, **140** (2022), 104377. https://doi.org/10.1016/j.autcon.2022.104377
16. Y. Tan, Z. Liang, H. Xu, C. Xing, Research on rutting deformation monitoring method based on intelligent aggregate, *IEEE Trans. Intell. Transp. Syst.*, **23** (2022), 22116–22126. https://doi.org/10.1109/TITS.2022.3175060
17. Y. Tan, Z. Liang, H. Xu, C. Xing, Internal deformation monitoring of granular material using intelligent aggregate, *Autom. Constr.*, **139** (2022), 104265. https://doi.org/10.1016/j.autcon.2022.104265
18. Z. Yue, W. Bekking, I. Morin, Application of digital image processing to quantitative study of asphalt concrete microstructure, *Transp. Res. Record*, **1492** (1995), 53–60.
19. K. Gopalakrishnan, N. Shashidhar, X. Zhong, Attempt at quantifying the degree of compaction in HMA using image analysis, *Adv. Pavement Eng.*, **1** (2005), 1–15. https://doi.org/10.1061/40776(155)18

20. A. R. Coenen, M. E. Kutay, N. R. Sefidmazgi, H. U. Bahia, Aggregate structure characterisation of asphalt mixtures using two-dimensional image analysis, *Road Mater. Pavement Des.*, **13** (2012), 433–454. https://doi.org/10.1080/14680629.2012.711923
21. N. R. Sefidmazgi, L. Tashman, H. Bahia, Internal structure characterization of asphalt mixtures for rutting performance using imaging analysis, *Road Mater. Pavement Des.*, **13** (2012), 21–37. https://doi.org/10.1080/14680629.2012.657045
22. J. Zhu, T. Ma, Z. Lin, J. Xu, X. Qiu, Evaluation of internal pore structure of porous asphalt concrete based on laboratory testing and discrete-element modeling, *Constr. Build. Mater*, **273** (2021), 121754. https://doi.org/10.1016/j.conbuildmat.2020.121754
23. X. Ding, T. Ma, X. Huang, Discrete-element contour-filling modeling method for micromechanical and macromechanical analysis of aggregate skeleton of asphalt mixture, *J. Transp. Eng. Pt. B-Pavements*, **145** (2019), 04018056. https://doi.org/10.1061/JPEODX.0000083
24. C. Xing, B. Liu, Z. Sun, Y. Tan, X. Liu, C. Zhou, DEM-based stress transmission in asphalt mixture skeleton filling system, *Constr. Build. Mater.*, **351** (2022), 128956. https://doi.org/10.1016/j.conbuildmat.2022.128956

AIMS Press