Article

A New Porosity Evaluation Method Based on a Statistical Methodology for Granular Material: A Case Study in Construction Sand

Binghui Wang [1], Shuanglong Xin [1], Dandan Jin [2,*], Lei Zhang [1], Jianjun Wu [1] and Huiyi Guo [1]

1 Faculty of Architecture and Civil Engineering, Jiangsu University of Science and Technology, Zhenjiang 212100, China; wbhchina@126.com (B.W.); 18362896535@163.com (S.X.); lei.zhang@just.edu.cn (L.Z.); 13686907142@163.com (J.W.); 19708893106@163.com (H.G.)
2 Faculty of Civil Engineering and Mechanics, Jiangsu University, Zhenjiang 212013, China
* Correspondence: jddnjut@163.com

Abstract: Sand porosity is an important compactness parameter that influences the mechanical properties of sand. In order to evaluate the temporal variation in sand porosity, a new method of sand porosity evaluation based on the statistics of target sand particles (which refers to particles within a specific particle size range) is presented. The relationship between sand porosity and the number of target sand particles at the soil surface considering observation depth is derived theoretically, and it is concluded that there is an inverse relationship between the two. Digital image processing and the k-means clustering method were used to distinguish particles in digital images where particles may mask each other, and a criterion for determining the number of particles was proposed, that is, the criterion of min(Dao). The execution process was implemented by self-written codes using Python (2021.3). An experiment on a simple case of Go pieces and sand samples of different porosities was conducted. The results show that the sum of the squared error (SSE) in the k-means method can converge with a small number of iterations. Furthermore, there is a minimum value between the parameter Dao and the set value of a single-particle pixel, and the pixel corresponding to this value is a reasonable value of a single-particle pixel, that is, the min(Dao) criterion is proposed. The k-means method combined with the min(Dao) criterion can analyze the number of particles in different particle size ranges with occlusion between particles. The test results of sand samples with different densities show that the method is reasonable.

Keywords: porosity; observed porosity; target particles; digital image; k-means clustering algorithm

1. Introduction

As a granular material, sand is widely present in natural sites. The study of its mechanical characteristics is one of the important contents of civil engineering, and it is also the foundation of engineering research such as site seismic analysis and dynamic responses in traffic engineering. The mechanical properties of soils are determined by their microstructural properties. The evaluation of compaction parameters such as porosity is an important consideration in studies of the mesostructural properties of soils. Porosity is also an important parameter in assessing the static/dynamic and also conduction or transport properties of sand [1]. Conventional geotechnical testing methods can only determine the initial or final porosity of a specific location within a soil sample by sampling but cannot simultaneously determine the distribution and change rule of porosity of the entire soil sample. To overcome this limitation, scholars have introduced other testing methods, such as CT scanning technology [2–4], nuclear magnetic resonance [5], digital image technology [2,6–8], and resistivity testing technology [9–13].

With the rapid development of computer technology and digital imaging technology in recent years, digital imaging technology, which was previously applied to the study of

soil's mesostructural properties and porosity evaluation abroad, has been widely applied to geotechnical engineering research. In 1976, Oda [14] proposed a statistical method of a void ratio in which particle centroids are connected into a polygon, which is often referred to as the Oda method. Later, this method was improved by Bhatia et al. [15] and Frost et al. [16] Delaunay triangulation meshes were used instead of polygon meshes to study the mesopore distribution law of sand, and the method was applied for the rapid determination of beach sand particle size [17,18]. Yoshimoto et al. [19] developed a brightness identification method to evaluate sand saturation. In China, digital imaging technology has been used to investigate soil displacement and deformation in the early stages. For example, Zhang Ga et al. [20] and Li Yuanhai et al. [21] analyzed the deformation field of soils based on the correlation analysis of images of sand particle motion. As research progressed, studies on microstructure were gradually carried out. Liu Jinghui et al. [22] analyzed the change in sand porosity during a stress process by using image technology based on triaxial tests. Zhou Jian et al. [23,24] introduced the process of digital image processing in more detail and proposed some criteria to distinguish sand particles and pores, and they applied them to the analysis of mesostructure characteristics and transport characteristics of meso-sand particles in the liquefaction process of sand. Chen Haiyang et al. [25] and Zhu Changqi [26] used image analysis technology to study the geometric shape characteristics of calcareous sand particles and the internal structural characteristics of the particles. Li Wei et al. [27] studied the crack propagation law of compacted expansive soil under wet–dry cycling using image analysis technology. Digital imaging technology was also used to identify the rock pore structure based on optical images of borehole walls in the site hole [28]. Image analysis technology combined with artificial intelligence is also an important research direction to evaluate the porosity of rock and soil, for example, in the work of Fahad et al. [29]. Bondi et al. [30] adapted Artificial Neural Networks/machine learning to predict porosity or bulk density values. Obviously, the porosity (or void ratio) obtained by the above research method based on plane images was different from the actual porosity for the boundary due to the presence of vessel walls or cut surfaces, and it was rarely discussed how one would try to minimize this distinction.

The aim of this study was to perform target particle identification and statistics based on digital images and to propose a new method for evaluating sand porosity and other compactness indices by considering a certain visual depth of the image. Among them, the target particle refers to the particles within a specific particle size range. First, the relationship between porosity and the number of target particles obtained by digital image statistics was theoretically analyzed, and a clustering algorithm for the digital image was introduced and its specific procedure established. The effectiveness of the algorithm was demonstrated by tests.

2. Relationship between the Porosity and Target Particle Number

Suppose a cuboid container is filled with sand; the volume of the container is $V = BS$, where S is the vertical sectional area perpendicular to the width direction, as shown in the left figure in Figure 1. If the volume of sand particles in the container is V_s, the pore volume is $V_v = V - V_s$. The basic definition of sand porosity n_3 is as follows:

$$n_3 = \frac{V_v}{V} = 1 - \frac{V_s}{V} \tag{1}$$

where the subscript '3' of n_3 indicates the actual porosity of the three-dimensional sand to make it easy to distinguish.
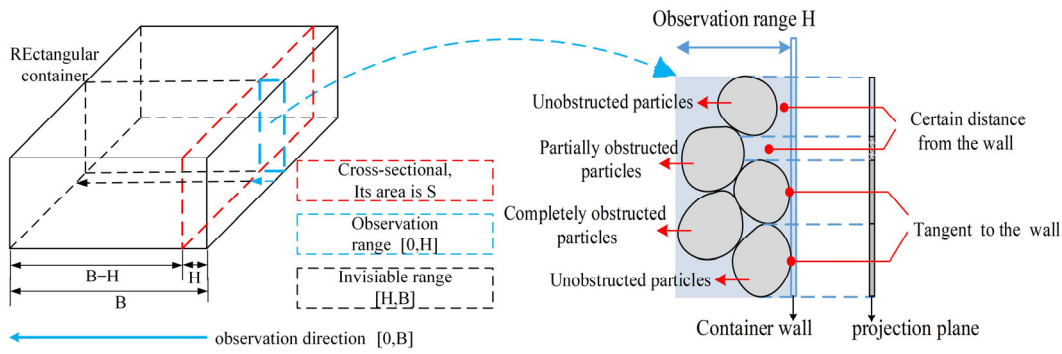
Figure 1. Sketch of the surface layer of sand.

The volume of all sand particles can be expressed as the sum of the product of the individual sand particle volume v_{si} of each particle size and the corresponding particle number.

$$V_s = \sum_{i=1}^{m} N_i v_{si} \tag{2}$$

where m is the number of types of sand particle size.

If a sand particle is taken as the target particle and we let b be the mass percentage of the target particle mass to the total mass of the sand, then the target particle volume expression is obtained on the basis of the unchanged target particle mass, that is, $\rho_b V_{sb} = \overline{\rho} V_s b$.

$$V_{sb} = C_b V_s b \tag{3}$$

$$C_b = \frac{\overline{\rho}}{\rho_b} \tag{4}$$

In which $\overline{\rho}$ is the average particle density and ρ_b is the target particle density of sand. Suppose the sand particle density of each particle size is the same, which is acceptable; then, $C_b = 1$. Similar to Equation (2), the target particle volume can be expressed as

$$V_{sb} = N_b v_{sb} \tag{5}$$

where v_{sb} is the individual target particle volume and N_b is the number of target particles. The total volume of sand particles expressed with the number of target particles was obtained by plugging Equation (3) into Equation (5), namely,

$$V_s = \frac{v_{sb}}{C_b b} N_b \tag{6}$$

This was obtained as follows by plugging into Equation (1).

$$n_3 = 1 - \frac{v_{sb}}{C_b b V} N_b \tag{7}$$

For a particular type of sand, C_b and b are usually unchanged. The above equation shows that the porosity of sand can be characterized by counting the number of target particles in sand.

Due to the difficulty of counting target particles inside the sand pile, it is only possible to observe target particles on the surface of the sand.

Figure 1 shows the case in which the sand was placed in a rectangular container with a width of B and a cross-sectional area of S. At an observation depth of H, this includes sand particles in contact or not with the container wall. Some of the sand particles that are not in contact with the container wall can be partially observed and are partially obscured. Therefore, the sand in the container could be divided into two parts, visible sand and invisible sand, and the range [0, B] of the sand body was divided into the sum of the range of sand surface [0, H] and the interior of sand [H, B]. If all of the sand particles in

the container were uniformly distributed, the observed porosity n_3^f of sand, which can be understood as the porosity obtained from surface observation, was the same as n_3. Using N_b^f to represent the number of target particles in the range of the sand surface $[0, H]$, the number of target particles at the sand surface is given by

$$N_b = N_b^f \frac{B}{H} \tag{8}$$

This was obtained as follows by plugging into Equation (7).

$$n_3 = n_3^f = 1 - \frac{v_{sb}}{C_b b S H} N_b^f \tag{9}$$

From the above equation, it was concluded that the porosity of the sand n_3 could be evaluated by the number of target particles on the surface of the sand N_b^f, assuming that the sand particles are uniformly distributed along the direction of observation.

It was noteworthy that in the observation range, some partially obscured sand particles could still be observed through the pores between sand particles, even if sand particles blocked each other. Therefore, the surface layer of sand mentioned in this paper not only includes the sand particles close to the container wall but also those that were not close to the container wall but could be observed, which are called the partially obscured particles, as shown in Figure 1. The observable sand particles were in the range of observable depth H. The view of this article is different from the mainstream view of Zhou Jian et al. [24].

Next, in order to count the number of target particles through the captured graphics, we compared the target particles with spheres and introduced sphere volume parameter v_{sphere}, that is,

$$v_{sb} = m v_{sphere} = m \frac{\pi d_b^3}{6} = \frac{2}{3} m \left[\pi (d_b/2)^2 \right] d_b = \frac{2m}{3} \bar{s}_{sb} d_b \tag{10}$$

where m indicates the ratio of the volume of sand particles to the volume of the sphere in the case that $m = v_{sb}/v_{sphere}$, which can also be called sphericity [31], and the value is related to the shape of particles, which can be referenced in [32]. d_b is the diameter of the sphere particles and simultaneously represents the height of the target sand particles in the observation direction, and $\bar{s}_{sb} = \pi (d_b/2)^2$ is the cross-sectional area of the sphere particles.

By plugging Equation (10) into Equation (9), the following equation is obtained:

$$n_3 = n_3^f = 1 - \frac{2m}{3C_b b} \frac{d_b}{H} \frac{\bar{s}_{sb}}{S} N_b^f \tag{11}$$

Considering that $N_E = S/\bar{s}_{sb}$ is the ratio of the window area to the standard area of sand particles, the equation was transformed into

$$n_3 = n_3^f = 1 - \frac{2m}{3C_b b D_b} \frac{N_b^f}{N_E} \tag{12}$$

Here, $D_b = H / d_b$, which is the ratio of the observable maximum depth range H to the target particle diameter d_b in the observation direction, and it is a parameter of relative depth. D_b is related to the sand compactness and grain composition. The denser the sand, the more likely it is that the particles will block each other, and the lower the D_b will be. The smaller the density, the greater the D_b will be. When the visible frequency is 1, it means that only the sand surface within the diameter range of a single particle close to the observed surface is considered. In addition, it is also related to the grain composition of sand target particles, and the higher the proportion of fine particles in sand and the smaller the pores between particles, the less likely it is that the deep particles can be observed. More target particles are obstructed by fine particles, and the image on the projection plane forms target

particles with holes. This is related to the light intensity during observation, which will not be discussed at this time.

The observed porosity can be obtained by digital image analysis of the digital image of the sand surface taken by a digital camera. Depending on the amount of information required, the digital image of the sand surface with a certain observation depth can be obtained with different image-acquisition requirements and processing methods. If the particle information with the "terrain" feature needed to be obtained, as shown in Figure 2a, the "terrain" information was mainly characterized by the grey value of pixel points, that is, points with the same grey value had a similar "topographic height". Obtaining this information required sophisticated image-acquisition and processing technology, such as the watershed image-processing technology. If only the target particle profile information was to be obtained, the image-acquisition and processing technology requirements were not high, and the statistics of the target particle number could be analyzed using a binary image, as shown in Figure 2b.
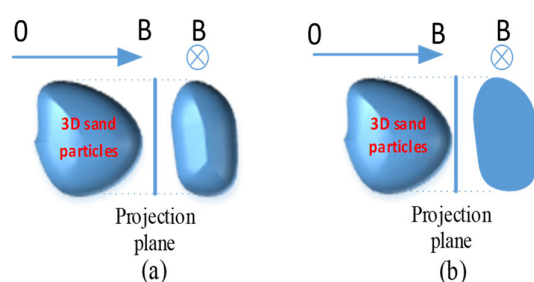

Figure 2. Two types of the dimensionality reduction processes for sand particles: (**a**) terrain image; (**b**) binary image.

To evaluate the porosity n_3 of sandy soil using the observed porosity formula n_3^f derived in this article, the following points should be noted.

1. In the evaluation of porosity using particle count statistics, the average particle size was used as the particle size. Therefore, this method takes the average particle size as the equivalent particle size of the target particles.

2. In the analysis of the porosity evaluation, the influence of the target particle shape was not considered, and the target particles were approximated using equiaxed particles. Therefore, this method is more suitable when the roundness or sphericity of the particle shape is relatively high.

3. For sand particles with a relatively large range of particle size, it is more appropriate to choose particles with larger sizes as the target particles. As an example of actual sand in Sections 4.2 and 4.3 of the text, the range of the sand particle size is 0.075–3 mm, and sand particles in the 2–3 mm particle size group were selected as the target particles.

Therefore, Formula (12) provides a relationship between the observed porosity and the number of target particles, which can be used to determine the porosity change of the sand pile, in the case that its properties change with static/dynamic loading even under a sand flow condition, using digital imaging technology based on images obtained without the strict control of lighting conditions.

3. Implementation of Digital Image Processing Based on Particles Statistical

The observed porosity was obtained by digital image processing and the analysis of the sand at the surface under a certain observation depth. When counting the number of target particles in a digital image, the target particle appeared in the digital image as an independent connected domain if it was a single particle, which could be easy to be identified, as shown in Figure 2b. When sand particles blocked each other in the observation direction, there was a connected domain containing overlapping projection areas of multiple target particles in digital images, as shown in Figure 3a. According to the definition of observed porosity, this connected domain needed to be distinguished. In this

study, the k-means clustering algorithm was used to determine the number of centroids based on the particle size of target particles.
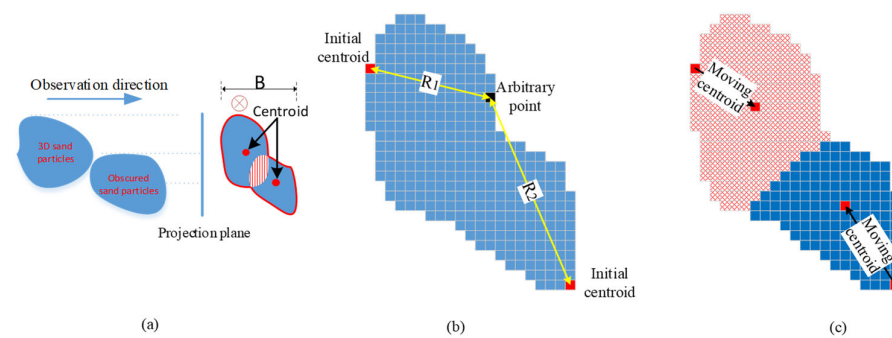

Figure 3. K-means clustering algorithm diagram: (**a**) binary image of two particles; (**b**) the distance between all of the pixels of target particles and the position of the initial centroid; (**c**) the center point moves to a new center point position.

Hartigan et al. [33] proposed the k-means clustering algorithm in 1979, and it has been continually improved since then [34]. So far, it has been widely used as a classical algorithm. The main idea is to use distance as the evaluation index of similarity; that is, the closer the two objects are, the more similar they are. The method was used here with each pixel of a connected domain as the object and the centroid as the clustering center, and the average pixel of the target particle was taken as the criterion to distinguish the type and number of connected domains, which were analyzed by comparing the distances between each pixel in the connected domain and the centroid.

The specific steps for using k-means clustering algorithm are as follows:

1. The digital image was captured by a camera, and the binary image of target particles was obtained through the image pre-processing, such as the binarization transformation under specific thresholds, as well as the expansion and corrosion operation in morphology, which are used to remove the pores formed by small particles blocking the target particles, which will not be elaborated here. Figure 3b shows a binary image of two particles in Figure 3a.

2. The image's connected domain and related information were obtained, and then, the initial centroid number k was determined by the number of pixels in each connected domain. And based on the number of pixels, the number of initial centroids was calculated. The initial centroid location was selected arbitrarily. As shown in Figure 3b, the initial centroid number of k = 2 was determined by dividing the number of pixels in the connected domain by the number of one-particle pixels, which is in a certain range and can implemented by a loop in the program. And the result was between 1 and 2 times the average number of target pixels.

3. The distance between all of the pixels of target particles and the position of the initial centroid was calculated, as shown by R_1 and R_2 in Figure 3b. According to the principle of the shortest distance from the centroid, all of the pixels were classified into two categories, as shown in the oblique shaded area and the blue area in Figure 3c. Here, the Euclidean distance is used for calculating distance, rather than other distances such as the Minkowski distance, for the shape of sand particles approaching a sphere.

4. The center point moves to a new center point position, which is the position of the average distance calculated in the case of the previous cluster center, as shown in Figure 3c.

5. Steps 2 and 3 were repeated until the centroid no longer needed to be moved. The final centroid was taken as the center, and the pixel from which the smallest distance was taken as the target particle. During the above iterative steps, the sum of the squared error (SSE) was also calculated to determine the quality of the cluster assignments.

The calibration of standard pixel points of target particles was based on the given shooting conditions, including the shooting distance, the camera's focal length, the camera's

sensor pixels, and other conditions, to obtain the pixels number of the target particle. Two points should be considered: 1. The shooting conditions of the target particle calibration were in accordance with the shooting conditions in the actual test process. 2. It was necessary to calibrate a large number of target particles and finally determine the number of standard target particle pixels.

The image process is shown in Figure 4. It consists of six main steps: basic parameter input, image reading and windowing, image pre-processing, connected domain search and statistics, particle statistics by k-means clustering and the criterion of min(*Dao*), and result output. The basic parameters that needed to be input included only the standard pixels of the target particles, which were determined by the previous calibration process of the average pixel points of target particles. The image pre-processing included the steps of greyscale, image enhancement, and image binarization, where the image enhancement used histogram equalization and the bimodal valley in the histogram automatically selected by the program was used as the threshold in image binarization. The related information of the connected domain included the number of connected domains, the number of pixels of each connected domain, and the coordinates of each pixel point. The output results included the number of windows, the number of target particles, the value of N_{2E}, the value of N_2/N_{2E}, and the observed porosity.
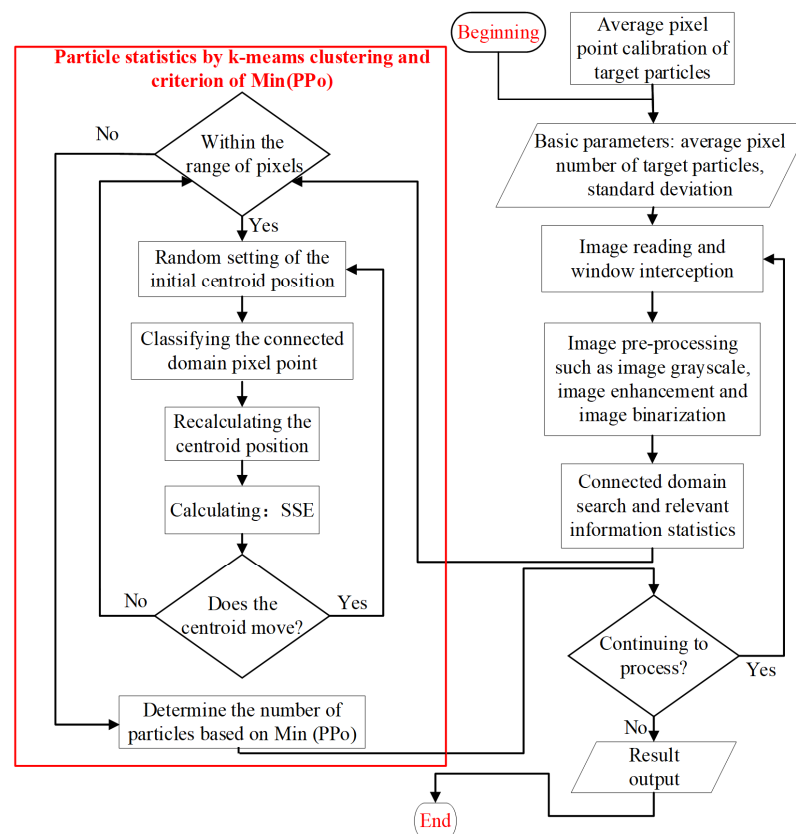

Figure 4. Main steps of the image processing program.

Through the above analysis, by giving specific individual particle pixels P_i, the area of occlusion can be analyzed. But due to differences in particle size and shape, only the average value P_m and the standard deviation δ can be obtained, not a specific value. Determining the suitable P_i value is still a problem.

Let us assume the simplest case, in which there are M particles of equal particle size with a total number of pixels as Ap_i, where a is the multiple coefficients of the standard deviation, which depends on the complexity of the target particle shape, and the subscript sign i is the number of single-particle pixel values. When there is no occlusion between

particles in a test image, it is obvious that there is $Ap_0 = P_i \times M$. But when there is partial occlusion between particles, $Ap_i < P_i \times M$. So the difference between the two areas is the sum of the obstructed area, i.e., $Dao = Ap_0 - Ap_i = P_i \times M - Ap_i$. For example, the shaded area between the two centers in Figure 3a represents the area of occlusion between the two particles.

Let us take a step further: the sizes of M particles are not the same, and there is a range of individual particle pixels where $P_i = [P_m - a\delta, P_m + a\delta]$. In this situation, the Dao should be carefully considered as it contains important information, that is, the minimum value of Dao within a specific range of P_i is as a judgement condition to obtain the actual number of particles, which is the criterion of min(Dao). A simple example will be used for further discussion in Section 4.1 below.

The above algorithm is implemented in Python, and the corresponding code is written by the authors.

4. Verification and Evaluation of the Method on Target Particle Statistics Based on Digital Images

4.1. Simple Case for Recognizing Circle Particles Partially Obstructed

In order to verify the correctness of the program written and to analyze the differences in identifying obstructed particles, as well as the pattern of determining the number of particles, including obstructed particles, a simple test was conducted by laying 38 black Go pieces with an average diameter of about 23 mm on a white paper. Among them, there were six pairs of Go pieces on the white paper with varying degrees of mutual occlusion between them, as shown in Figure 5a. The results of binarization and morphological processing of the image are shown in Figure 5b,c, respectively. Based on Figure 5c, the average area of the Go pieces can be obtained as $P_m = 116{,}618$ pixels, and the variance $\delta = 1667$. The results obtained by fitting an equal diameter circle are shown in Figure 5d.

(**a**) image taken   (**b**) after binarization

(**c**) after morphological processing (**d**) after simulating an equal diameter circle
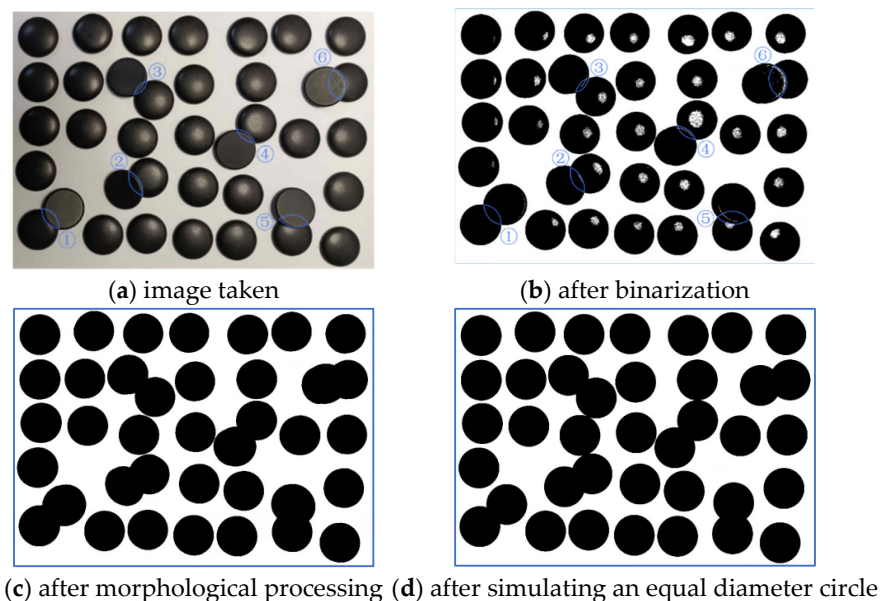
Figure 5. DIP processes of the mode 1–3 test.

In order to evaluate the application effect of the k-means method, the results of analyzing the sixth pair of overlapping particle graphs using this method are given, as shown in Figure 6; that is, the graph of SSE convergence result with the increase in iteration times. It can be seen that all results converge rapidly after about 10 iterations. The figure also shows the position movement diagram of the two cluster center points when the set value of the number of go sub-pixel points is 23,000, as shown by the "×" point under the white arrow in the blue-background binary diagram in the figure.
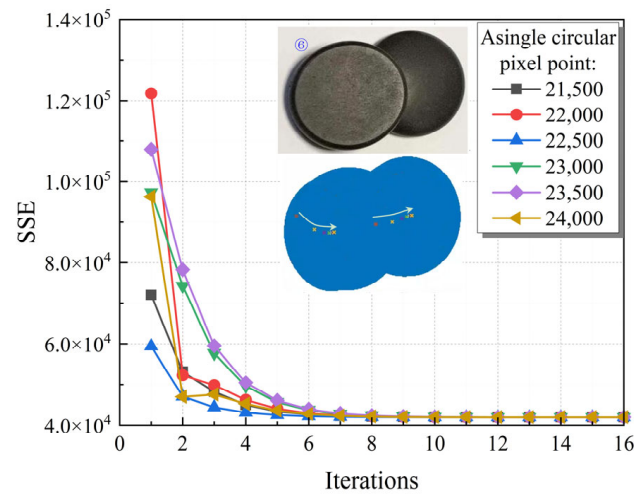
Figure 6. Convergence of SSE with iterations.

The number of particles obtained from the fitting result of the medium-diameter circle in Figure 5d is 38, and the minimum value of the relationship between the number of obstructed pixels *Dao* and P_i mentioned earlier was used as the judgment condition to determine the number of particles. In order to discuss the reasonableness of the judgment condition, the relationship between *Dao* and Pi within the range of $P_i = [P_m - a\delta, P_m + a\delta]$ is shown in Figure 6. The execution results show that there is a minimum point in the relationship between the number of obstructed particle pixels and the input single-particle pixel number P_i. The number of particles corresponding to this minimum point is used to match the actual number of Go pieces, as shown in Figure 7.
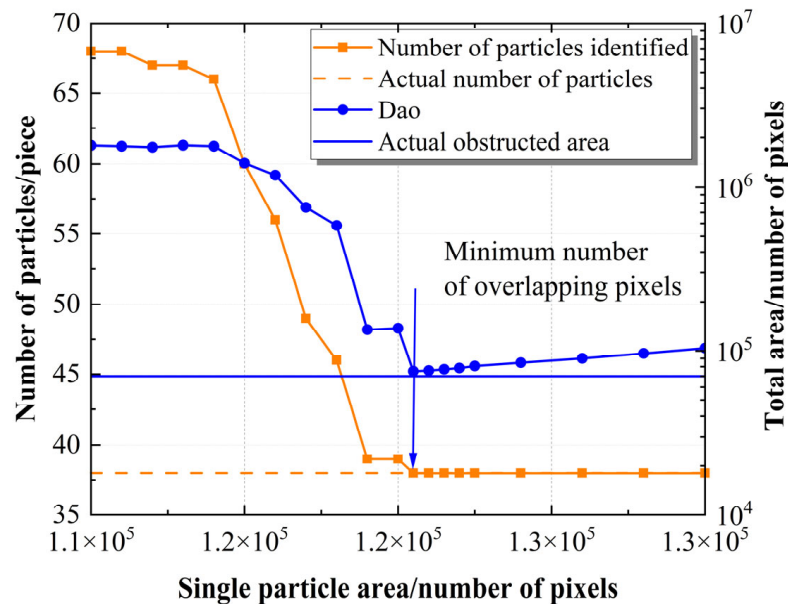

Figure 7. The change in particle number counted and *Dao* with the set value of the single-particle pixel number.

In addition, careful comparison shows that the overlapping area in the simulated image of the same diameter circle is very close to the actual obstructed area, as shown in Figure 6. The obstructed Go piece and other unobstructed Go pieces are curved upwards on white paper, and the pixel area obtained from the shooting is close. Therefore, the actual obstructed area can be calculated as the distance between the centers of the obstructed Go pieces, as shown in Figure 5a,b.

In summary, using the minimum value of the $Dao \sim P_i$ relationship as the criterion to determine the number of particles is reasonable.

4.2. Sand, Test Step, and Calibration

Quartz sand with a specific gravity of G_s = 2.69 and maximum and minimum dry densities of 1.84 g/cm^3 and 1.53 g/cm^3 was used; the particle size distribution of sand used in the validation test was as follows: particle size 2–3 mm accounted for 14.09%, particle size 1–2 mm accounted for 29.99%, particle size 0.5–1 mm accounted for 48.15%, particle size 0.25–0.5 mm accounted for 6.92%, and particle size 0.075–0.25 mm accounted for 0.85%. The corresponding maximum and minimum porosities were 0.419 and 0.315, respectively. Sand particles with particle sizes of 2~3 mm were selected as target particles, with a mass percentage of b = 14.09%, and they were dyed black and then remixed with uniform stirring to make test objects containing black target particles. They were scattered in an acrylic soil box with transparent walls and internal dimensions of $L \times B \times H'$ = 29 cm $\times$ 19 cm $\times$ 9 cm. A Nikon D810 camera with about 36.35 million effective pixels and a manual focus was used.

The test steps were as follows:

1. The average pixel point of target particles (particles dyed black) was calibrated while determining the shooting distance and the focal length of the camera.

2. We mixed the target particles with unstained particles of other sizes and mixed them evenly as much as possible to prepare the sand sample.

3. We weighed a sample with a mass of m_s (unit: g) and then layered it into a soil box. After leveling the surface of the sand, we measured the stacking height h in cm units to obtain the volume V_s of the sand. We calculated the corresponding porosity according to the following equation and recorded it.

$$n = 1 - \frac{m_s}{LBhG_s\rho_w}$$

In the formula, ρ_w is the density of water, taken as 1.0 g/cm^3.

4. We took clear photos in the direction perpendicular to the height and length plane of the box. Digital images were taken at the same focal length and shooting distance as the calibration in step 1, and we recorded the name of the photo corresponding to the porosity.

5. We gently tapped the wall of the soil box to slowly compact the sand and then tested the stacking height h of the sand, substituted it into the formula to recalculate the porosity, and recorded it. The soil box was vibrated to compact the sand, the height of the sand was measured, and the porosity was calculated.

6. We repeated steps 2 and 3 above until the sand no longer solidified when the wall of the soil box was lightly tapped.

7. Using the processing method established in Section 3, we processed the corresponding porosity photos to obtain the observed porosity. This article conducted a total of 4 sets of 21 working conditions tests, and the porosity of each set of sand was measured as shown in Table 1.

When calibrating the average pixel of target particles, the photographed subjects were 699 target particle, which rarely touched each other on white paper, and the obtained average pixel of target particles was 604 with a standard deviation of 158; the pixels' range used in the program was (440,770).

Table 1. Results of tests.

Test Group Number	Test Number	Porosity n	N_2/N_{2E}	n_3^f	Test Group Number	Test Number	Porosity n	N_2/N_{2E}	n_3^f
Test 1	1	0.421	0.314	0.427	Test 3	1	0.427	0.325	0.407
	2	0.419	0.313	0.428		2	0.404	0.330	0.397
	3	0.410	0.323	0.410		3	0.396	0.339	0.380
	4	0.404	0.324	0.408		4	0.388	0.336	0.387
	5	0.400	0.332	0.393		5	0.383	0.340	0.378
Test 2	1	0.410	0.321	0.412		6	0.375	0.323	0.410
	2	0.402	0.324	0.408	Test 4	1	0.410	0.328	0.400
	3	0.393	0.329	0.398		2	0.402	0.335	0.387
	4	0.385	0.337	0.385		3	0.393	0.334	0.389
	5	0.380	0.336	0.385		4	0.385	0.336	0.385
						5	0.380	0.325	0.407

*4.3. Results Analysis and Method Evaluation*

Three rectangular windows of each digital image were captured at fixed positions. The number of target particles in each window was taken as the value N_2, and the observed porosity n_3^f was calculated by substituting the calculated value N_{2E} of the sum of all the window areas into Equation (12). This calculation process was very important, and it was also the basic requirement to meet the consistency of the sample estimation.

Figure 8 shows the image processing of working conditions 1–3 of the test experiments, including the original photograph and the image of window ②, the image after greyscale, binarization, and the image after particle statistics and simulation. From Figure 8, it can be seen that there are many black particles obstructed by other particles, such as the black particles shown in the red oval box in the figure, which are clearly obstructed by other particles. These particles were all counted as equal-diameter circular particles. In the statistical analysis process, the method described in Section 4.1 was used to determine the number of sand particles. Table 1 lists the results of four sets of test experiments.
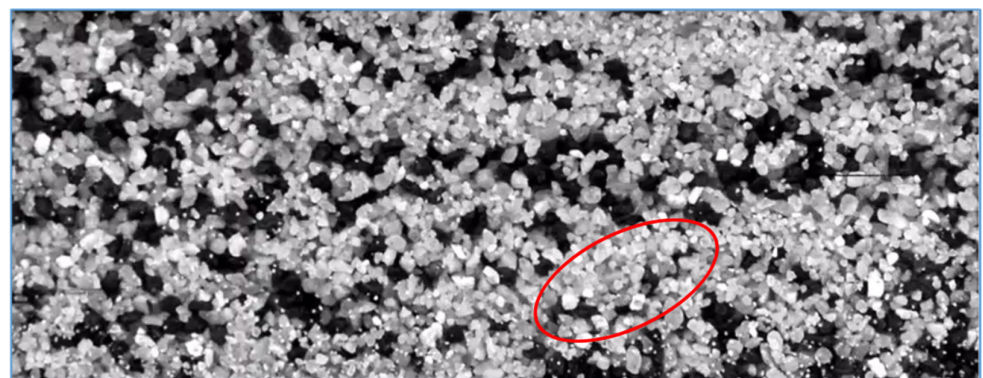
Figure 9a shows the relationship between the actual porosity n and N_2/N_{2E}. We found that there was an inverse relationship between the two, which is consistent with the inverse relationship shown by Formula (12) derived from the theory. Then, the values of each parameter in Formula (12) are given based on the corresponding physical meanings. $C_b = \overline{\rho}/\rho_b = 1$ is acceptable, as mentioned above. A mass percentage b = 14.09% was obtained by testing. Both referring to reference [30] and considering the shape of sand in this test, sphericity m = 0.85 is acceptable. For the parameter D_b, there is currently no reference, and no experiments have been conducted. If D_b = 2.2, it can be obtained that $2m/3C_b b D_b$ = 1.828. After analyzing the above parameters, we found that the slope value obtained by Formula (12) is very close to that obtained using the test data in Figure 9. The significance of this value exceeding 2.0 is to indicate that the observable depth is greater than twice the particle size; this aspect is feasible and requires further research.

Figure 9b shows the relationship between the actual porosity and the observed porosity; their average could be fitted using a linear relationship with a high degree of correlation, and the value of R^2 was 0.92.

Although there was a good linear relationship between the actual porosity and the mean value of the observed porosity, there was relatively large discrepancy in the statistical intermediate process, which was manifested by the following two aspects: one was the discreteness between the different windows of the same image; the other was the discreteness between different images. The main reason was that target particles are unevenly distributed in the plane perpendicular to the observation direction. Therefore, in the actual image-processing process, capturing a sufficient number of windows was conducive to obtaining a good linear relationship between the observed porosity and the actual porosity.
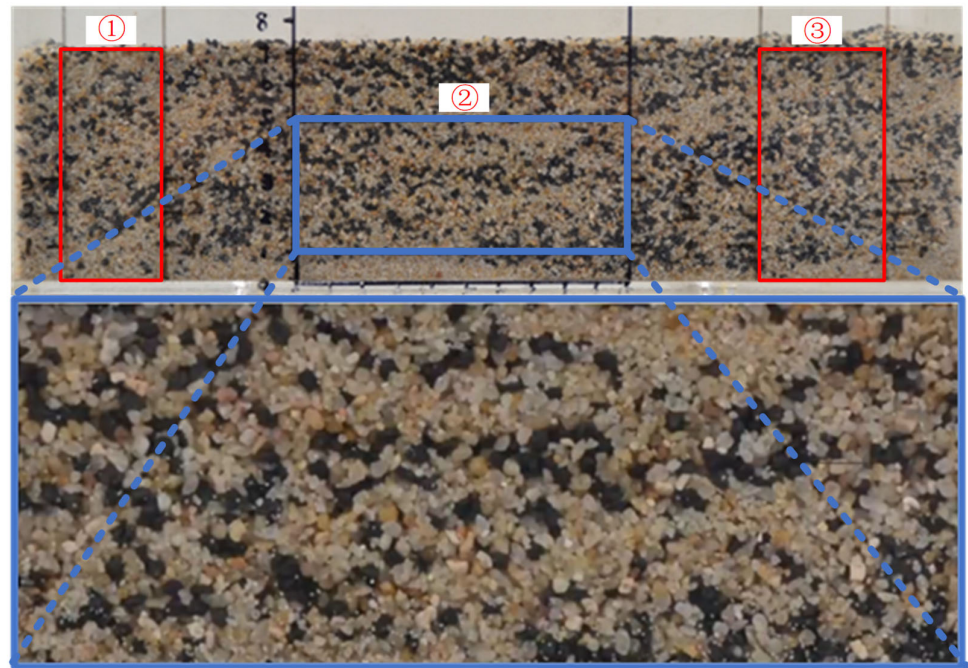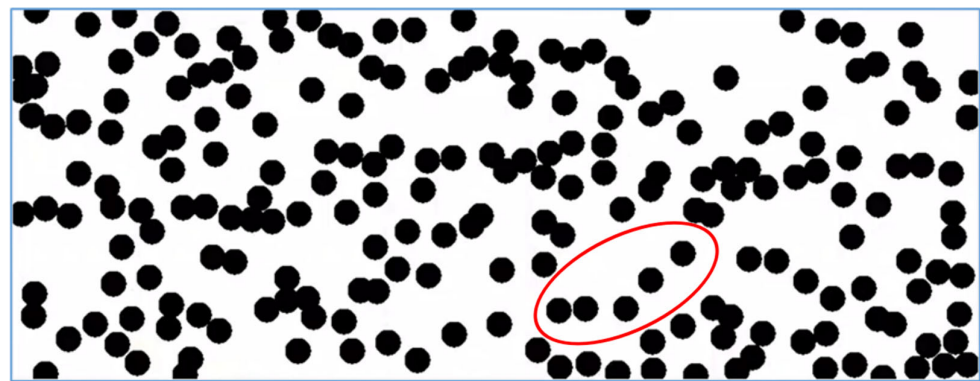

(**a**) Original image

(**b**) Grayscale image

Figure 8. *Cont.*

(**c**) Binarized and morphological

(**d**) Image after particle statistics and simulation

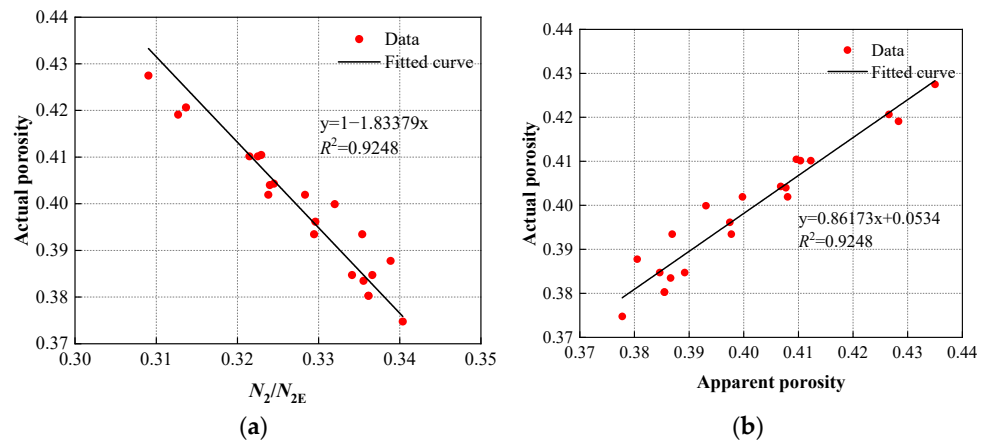**Figure 8.** Processes of the DIP of typical mode test.

(**a**)

(**b**)

Figure 9. Relationship between the sand porosity and observed porosity: (**a**) the relationship between the actual porosity n and N_2/N_{2E}; (**b**) the relationship between the actual porosity and the observed porosity.

5. Discussion

In this study, the idea and theoretical feasibility of evaluating sand porosity using the target particle number statistics method was proposed. However, the following problems remain to be investigated further.

1. The particle size selected in this paper was 2~3 mm, which accounted for about 14% of the total mass. However, when target particles with particle sizes of 0.5~1 mm and 0.25~0.5 mm were used, there were too many target particles in the digital images taken, and they were also too dense. When the number of target particles was large, the entire

screen would appear in the captured digital image and they could not be gratifying. When selecting the target particles, it was necessary to consider the volume ratio of the target particles in addition to the representativeness.

2. The relationship between the observed porosity and the actual porosity was a linear relationship for simplicity, but more complex relationships needed to be considered.

3. The discreteness of the target particle statistics and the optimization of the digital image processing have not yet been analyzed.

4. Black target particles were used in the verification test, which corresponded to the color of the particle shadow. In the actual test, the shadows of a small number of particles were mistaken for target particles; this could be improved by using red target particles [35].

6. Conclusions

In this paper, through the statistics of target particles within a specific range of particle size, the relationship between the porosity of sand and the target particles was theoretically established, and through the simple example of Go, Zi, and the experimental analysis of sand samples, the main conclusions established were as follows:

1. The relationship between sand porosity and the number of target sand particles on the surface was derived, and an inverse relationship between the observed porosity of sand and the number of sand particles was established.

2. A program was implemented based on the k-mean clustering algorithm, and a criterion was established to determine the reasonable number of pixels in the projection area of a single particle, that is, min(*Dao*). The applicability of the k-mean method is effectively illustrated by a simple example of a Go player. The SSE parameters that evaluate the quality of the method converge rapidly with the number of iterations. The relationship between the *Dao* parameter and the pixel point set value Pi clearly shows that the pixel number corresponding to the minimum value of parameter *Dao* is the reasonable pixel number in a single-particle projection area.

3. By introducing the above theoretical formula and analysis program, a method to determine the statistics related to target particles based on a digital image was realized, and the results of tests carried out on sand with different porosities show that the observed porosity was inversely related to the number of sand particles. The feasibility of evaluating the actual porosity of sand based on the statistics of target particle number at the sand surface was verified.

Author Contributions: B.W.: conceptualization, methodology and writing—original draft; S.X.: software, visualization and writing—review and editing; D.J.: conceptualization, data curation and writing—review and editing; L.Z.: visualization, writing—review and editing; J.W.: supervision, visualization; H.G.: supervision, visualization. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflicts of interest.

Notation

The following symbols are used in this paper:

Symbol	Description	Symbol	Description
V	Volume of the container	$\bar{\rho}$	The average particle density
V_s	Volume of sand particles	ρ_b	The target particle density
V_v	Volume of pore	N_b	The number of target particles
V_{sb}	Volume of the target particle	N_b^f	The number of target particles on the surface of the sand
v_{si}	Volume of the individual sand particle	m	$m = v_{sb}/v_{sphere}$
v_{sb}	Volume of the individual target particle	d_b	The diameter of the sphere particles
v_{sphere}	Sphere volume parameter	$\bar{s}_{sb}$	The cross-sectional area of the sphere particles
B	Width of the container	N_E	$N_E = S/\bar{s}_{sb}$
S	Cross-sectional its area	D_b	$D_b = H / d_b$
H	The observation depth	SSE	The sum of the squared error
n_3	Sand porosity	P_i	Single particle pixel number
n_3^f	The observed porosity	Dao	The sum of the obstructed area

References

1. Mitchell, J.K.; Soga, K. *Fundamentals of Soil Behavior*; John Wiley & Sons: New York, NY, USA, 2005.
2. Hussaini, S.R.; Dvorkin, J. Specific surface area versus porosity from digital images: High-porosity granular samples. *J. Pet. Sci. Eng.* **2021**, *206*, 108961. [CrossRef]
3. Li, C.S.; Zhang, D.; Wang, H.X.; Shasha, D. 3D mesh generation for soil-rock mixture based on CT scanning. *Rock Soil Mech.* **2014**, *35*, 2731–2736. [CrossRef]
4. Yamada, F.; Tateyama, R.; Tsujimoto, G.; Suenaga, S.; Long, B.; Pilote, C. Dynamic monitoring of physical models beach morphodynamics and sediment transport using x-ray ct scanning technique. *J. Coast. Res.* **2013**, *65*, 1617–1622. [CrossRef]
5. Zhang, K.; Li, N. A new method to replicate high-porosity weak rocks subjected to cyclic freezing-thawing: Sand 3D printing and digital image correlation explorations. *Int. J. Rock Mech. Min. Sci.* **2022**, *157*, 105174. [CrossRef]
6. Shao, L.; Yang, S.; Sang, Y.; Huang, C.; Guo, X. Application of digital image processing technology in dynamic triaxial test of soil mechanics. *J. Theor. Appl. Inf. Technol.* **2013**, *48*, 1358–1365.
7. Shao, L.T.; Liu, X.; Guo, X.X.; Huang, C.; Xue, J. Whole surface deformation measurement of triaxial soil specimen based on digital image processing. *Chin. J. Geotech. Eng.* **2012**, *34*, 409–415.
8. Xu, A.Q.; Xu, W.Y.; Shi, C.; Li, D.L. Micromechanical properties and mechanical parameters of talus deposit based on digital image technology. *Chin. J. Geotech. Eng.* **2012**, *34*, 58–64.
9. Cai, G.J.; Zhang, T.; Liu, S.Y.; Deng, Y.F. Relationship between electrical resistivity and geotechnical characteristic parameters for Jiangsu marine clay. *Chin. J. Geotech. Eng.* **2013**, *35*, 1470–1477.
10. Zhou, M.; Wang, J.G.; Huang, S.B.; Peng, D.; Zhang, L.Q.; Wei, Y. Experimental investigation on influencing factors in soil resistivity measurement. *Rock Soil Mech.* **2011**, *32*, 3269–3275. [CrossRef]
11. Samouëlian, A.; Cousin, I.; Tabbagh, A.; Bruand, A.; Richard, G. Electrical resistivity survey in soil science: A review. *Soil Tillage Res.* **2005**, *83*, 173–193. [CrossRef]
12. Thevanayagam, S. Electrical response of two-phase soil. Theory and applications. *J. Geotech. Eng.* **1993**, *119*, 1250–1275. [CrossRef]
13. Archie, G.E. The electric resistivity logs as an aid in determining some reservoir characteristics. *Trans. AIME* **1942**, *146*, 54–61. [CrossRef]
14. Oda, M.S. Fabrics and their effect on the deformation behaviour of sand. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1976**, *15*, 75.
15. Bhatia, S.K.; Soliman, A.F. Frequency distribution of void ratio of granular materials determined by an image analyzer. *Soils Found.* **1990**, *30*, 1–16. [CrossRef]
16. Frost, J.D.; Kuo, C.Y. Automated determination of the distribution of local void ratio from digital images. *Geotech. Test. J.* **1996**, *19*, 107–117. [CrossRef]
17. Pentney, R.M.; Dickson, M.E. Digital grain size analysis of a mixed sand and gravel beach. *J. Coast. Res.* **2012**, *28*, 196–201. [CrossRef]
18. Baptista, P.; Cunha, T.R.; Gama, C.; Bernardes, C. A new and practical method to obtain grain size measurements in sandy shores based on digital image acquisition and processing. *Sediment. Geol.* **2012**, *282*, 294–306. [CrossRef]
19. Yoshimoto, N.; Orense, R.P.; Tanabe, F.; Kikkawa, N.; Hyodo, M.; Nakata, Y. Measurement of degree of saturation on model ground by digital image processing. *Soils Found.* **2011**, *51*, 167–177. [CrossRef]
20. Zhang, G.; Zhang, J.M.; Liang, D.F. Measurement of soil particle movement in soil-structure interface test. *Chin. J. Geotech. Eng.* **2005**, *27*, 903–907.

21. Li, Y.H.; Zhu, H.H.; Ueno, K.; Mochizuki, A. Deformation field measurement for granular soil model using image analysis. *Chin. J. Geotech. Eng.* **2004**, *26*, 36–41.

22. Liu, J.H.; Hong, B.N.; Zhang, H.B. A new experimental method of soil microstructure changing process. *Rock Soil Mech.* **2003**, *24*, 744–747.

23. Zhou, J.; Chen, X.L.; Jia, M.C.; Feng, Y.W. Dynamic centrifuge tests on macro-micro mechanism of liquefaction of saturated sandy foundation with buried structures. *Chin. J. Geotech. Eng.* **2012**, *34*, 392–399.

24. Zhou, J.; Yu, R.C.; Jia, M.C. Measurement of microstructure parameters for granular soil model using digital image technology. *Chin. J. Geotech. Eng.* **2006**, *28*, 2047–2052.

25. Chen, H.Y.; Wang, R.; Li, J.G.; Zhang, J.M. Grain shape analysis of calcareous soil. *Rock Soil Mech.* **2005**, *26*, 1389–1392. [CrossRef]

26. Zhu, C.Q.; Chen, H.Y.; Meng, Q.S.; Wang, R. Microscopic characterization of intra-pore structures of calcareous sands. *Rock Soil Mech.* **2014**, *7*, 1831–1836.

27. Li, W.; Liu, G.S.; Wang, W.W.; Yao, T.; Sheng, J.H. Crack propagation law of compacted expansive soils under wetting-drying cycles. *Chin. J. Geotech. Eng.* **2014**, *36*, 1302–1308. [CrossRef]

28. Wang, C.; Wang, C.Y.; Wang, Y.T.; Wang, J.H.; Chen, W.; Han, Z.Q. Research on identification and analysis method of rock pore structure based on optical images of borehole walls. *Chin. J. Rock Mech. Eng.* **2021**, *40*, 1894–1901. [CrossRef]

29. Al-Qahtani, F.A. Porosity Distribution Prediction Using Artificial Neural Networks. Master's Thesis, West Virginia University, Morgantown, WV, USA, 2000. Available online: https://researchrepository.wvu.edu/etd/1010/ (accessed on 29 July 2024).

30. Bondi, G.; Creamer, R.; Ferrari, A.; Fenton, O.; Wall, D. Using machine learning to predict soil bulk density on the basis of visual parameters: Tools for in-field and post-field evaluation. *Geoderma* **2018**, *318*, 137–147. [CrossRef]

31. Anusree, K.V.; Latha, G.M. Characterization of sand particle morphology: State-of-the-art. *Bull. Eng. Geol. Environ.* **2023**, *82*, 269. [CrossRef]

32. Alshibli, K.A.; Alsaleh, M.I. Characterizing Surface Roughness and Shape of Sands Using Digital Microscopy. *J. Comput. Civ. Eng.* **2004**, *18*, 36–45. [CrossRef]

33. Hartigan, J.A.; Wong, M.A. Algorithm as 136: A k-means clustering algorithm. *Appl. Stat.* **1979**, *28*, 100–108. [CrossRef]

34. Celebi, M.E.; Kingravi, H.A.; Vela, P.A. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Syst. Appl.* **2013**, *40*, 200–210. [CrossRef]

35. Bloom, M.; Russell, M.J.; Kustau, A.; Mandayam, S.; Sukumaran, B. Measurement of porosity in granular particle distributions using adaptive thresholding. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 1192–1199. [CrossRef]