19th WORLD CONFERENCE OF THE ASSOCIATED RESEARCH CENTRES FOR THE URBAN UNDERGROUND SPACE
Belgrade, Serbia, November 4-7, 2025.

Paper No: 2.8.105
DOI: 10.46793/ACUUS2025.2.8.105

EXTRACTION OF INDIVIDUAL ROCK DISCONTINUITIES BASED ON THREE-LAYER CLUSTERING ALGORITHM FROM 3D POINT CLOUDS

Bingyi Pan[1], Wei Wu[2]*, Yong Huang[3], Baolin Chen[4], Hehua Zhu[5]

Abstract: Accurate characterization of rock discontinuities is essential for tunnel stability analysis, as discontinuities play a fundamental role in governing the mechanical behavior and failure mechanisms of the surrounding rock mass. The advancement of remote surveying methods has enabled the rapid identification of discontinuities and extraction of pertinent geometric information from 3D point clouds of the rock mass. Nevertheless, traditional methods that rely on point features are insufficient in their ability to eliminate fragmented non-structural regions during the process of individual discontinuity extraction. This paper presents a novel three-layer clustering approach that aims to extract individual discontinuities from 3D point clouds automatically. To begin with, the rock mass point cloud is filtered to extract core points and remove non-structural regions. Subsequently, an improved density-based spatial clustering of applications with noise (IDBSCAN) is employed to identify planar units within core points. To merge adjacent fragmented planar units, these units are re-clustered based on their spatial relationships, resulting in the identification of individual discontinuities. A weighted modified k-means++ method is then used to cluster individual discontinuities according to their respective normal vectors. Consequently, the orientation, set, and trace of each individual discontinuity can be obtained respectively through geometric analysis. This method is applied to one tunnel case and compared with previous studies. The results suggest that the proposed method has a high level of reliability and accuracy when it comes to automatically extracting individual discontinuities from 3D point clouds, providing a more robust foundation for tunnel design and construction.

Keywords: Individual discontinuities, Rock mass, 3D point clouds, Automatic extraction

1. INTRODUCTION

The identification of discontinuities is crucial for evaluating the strength, stability, and mechanical behavior of the rock mass. Conventional measurement methods primarily rely on geological engineers manually measuring using geological compasses or scanlines. However, these methods have significant drawbacks as they are time-consuming, pose risks to geological engineers, and cannot measure certain areas (Chen et al., 2016). Recently, rapid advancements in remote surveying technologies have enabled the use of laser scanners or photogrammetry to acquire 3D point clouds of the rock outcrop. This offers safe and highly precise non-contact measurement methods capable of capturing areas that are inaccessible to manual measurement.

In recent times, numerous studies have focused on extracting discontinuities from 3D point clouds of the rock outcrop obtained through photogrammetry or laser scanning methods (Chen et al., 2021; Singh et al., 2023). Certain studies employ a manual point selection method from 3D point clouds, followed by plane fitting to identify discontinuities (Assali et al., 2016; Drews et al., 2018). This approach necessitates operators with professional geological expertise and is time-consuming. Moreover, automatic methods for discontinuity identification have garnered attention, primarily focused on clustering-based and region growing-based approaches (Hu et al., 2020; Kong et al., 2020). Singh et al. (2021, 2022) introduced sinusoidal waves and local point descriptors as supplements to the point features used in clustering methods. Subsequently, researchers performed density

[1] Bingyi Pan, Department of Geotechnical Engineering, Tongji University, Shanghai, China, 2111040@tongji.edu.cn.
[2] Wei Wu, Department of Geotechnical Engineering, Tongji University, Shanghai, China, weiwu@tongji.edu.cn.
[3] Yong Huang, China Railway First Survey and Design Institute Group Co., Ltd., Xi'an, China, 32186254@qq.com.
[4] Baolin Chen, Zhejiang Institute of Communications Co., Ltd., Hangzhou, China, chenbaolin1992@126.com.
[5] Hehua Zhu, Department of Geotechnical Engineering, Tongji University, Shanghai, China, zhuhehua@tongji.edu.cn.

clustering on the grouped points to identify individual discontinuities. However, due to the relatively uniform nature of 3D point clouds in the rock outcrop, this method fails to effectively eliminate non-structural regions. The region growing method utilizes the nearest neighbor relationship among point clouds for direct growth, which is affected by the global threshold and tends to incorporate fragmented non-structural surface regions. Additionally, Ge et al. (2022) suggested using artificial neural networks for automatic grouping of discontinuities. However, data labeling is still necessary for each rock outcrop point cloud.

Given the limitations of traditional discontinuity identification methods in effectively eliminating non-structural regions, this paper employs a three-layer clustering approach to extract individual discontinuities from 3D point clouds of the rock mass and perform grouping. Initially, the point normal vector of each point is acquired as a point feature by conducting principal component analysis (PCA) on the point clouds. Subsequently, the point clouds of the rock mass are filtered to extract the core points and remove non-structural regions. Next, a three-layer clustering method is applied to identify and group the individual discontinuities. Lastly, geometric analysis is employed to determine the orientation, set, and trace information for each individual discontinuity.

2. METHODOLOGY

2.1. Point normal vector calculation

To efficiently obtain the point normal vector, this paper utilizes a combination of the k-nearest neighbor (KNN) algorithm with a fixed number of nearest neighbor points and the principal component analysis (PCA) method. By specifying the number of nearest neighbor points *k*, the KNN algorithm rapidly retrieves the *k* nearest neighbor points, including the point itself, from the point clouds of the rock mass. The fundamental concept involves minimizing the sum of squared distances between the nearest neighbor points and the fitting plane to determine the fitting plane, and subsequently utilizing its normal vector as the point normal vector for that specific point. This approach is equivalent to performing PCA on a matrix formed by the 3D coordinates of the nearest neighbor points. The covariance matrix of the 3D coordinate matrix needs to be calculated first using the following equation:

$$C = \frac{1}{k-1} P^{\cdot} P \tag{1}$$

where P is the 3D coordinate matrix of the nearest neighbor points in k rows and three columns (after centralization); k is the number of the nearest neighbor points (including the query point itself); C is the covariance matrix of P, with three rows and three columns. The covariance matrix C is decomposed into its three eigenvalues λ_1, λ_2, λ_3 ($\lambda_1 \geq \lambda_2 \geq \lambda_3$). The unit eigenvector corresponding to λ_3 is considered as the point normal vector of that query point.

2.2. Three-layer clustering approach

2.2.1. Core point identification

To eliminate non-structural regions and filter the point clouds of the rock mass, the core points are initially identified as the points to be clustered from the point clouds. Ester et al. (1996) employed the concept of nearest neighbor point density to search for core points, but without fully accounting for the curvature information. In this study, the core points must satisfy two conditions:

Condition 1: the flatness coefficient $\eta = \dfrac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$ must be below the set threshold to roughly identify points with lower curvature (figure 1). Only points that meet the flatness threshold condition $\eta < \eta_{\max}$ ($\eta_{\max}$ is a predetermined constant) are retained for the density calculation;

Condition 2: the density of all neighbor points meeting condition 1 for the queried point should surpass the density of all points in the point clouds that meet condition 1, which is calculated using the following equation:

$$\frac{k_f}{k} > \frac{N_f}{N} \tag{2}$$

where k_f is the number of the nearest neighbor points that meet the flatness condition, k is the number of the nearest neighbor points for each point, N_f is the number of points among all points that satisfy the flatness condition, and N is the number of all points in this dataset.

Evaluating these two conditions suggests that local curvature information can effectively eliminate non-structural regions and identify core points.

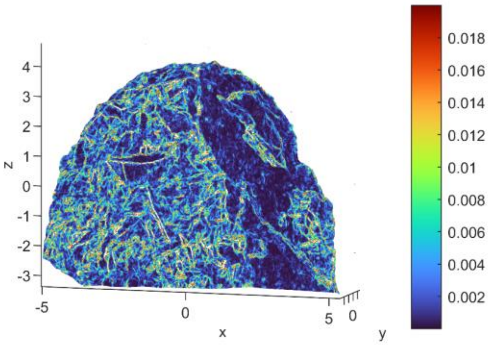


Figure 1. *Flatness index of the tunnel face.*

2.2.2. Core point clustering

After obtaining the core points, it is necessary to cluster them initially to obtain planar units. The traditional DBSCAN algorithm clusters based on the density of the point cloud, which does not consider the coplanar features of the discontinuity point cloud. To take full advantage of the coplanarity of the discontinuity point cloud, the calculation of the angle between the k-nearest neighboring points is added to perform the density connectivity judgment. The IDBSCAN algorithm is adopted here, with the following steps:

1. Select the core point with the highest flatness coefficient as the initial clustering point and search for core points with density direct connection (using the k-nearest neighbor relationship) to join the cluster. At the same time, the added points must meet the requirement that the angle between their normal vector and the fitting plane normal vector of the existing points in the cluster is below the threshold;
2. Repeat step 1 until there are no more core points with density direct connection;
3. Choose the core point with the highest flatness coefficient from the remaining core points as another initial clustering point for further clustering;
4. Iteratively perform step 1-3 until all core points have been traversed.

Clustering the core points allows for comprehensive consideration of the nearest neighbor angle relationship, resulting in the acquisition of planar units.

2.2.3. Planar unit clustering

Planar units are clusters of core points, which can still exhibit fragmentation. Additionally, there might be multiple planar units associated with a single individual discontinuity. Hence, it is necessary to cluster planar units. The clustering of planar units involves the following steps:

1. Retrieve the k nearest neighbor points for all the points within each planar unit and determine the connectivity of the planar units based on the intersection relationship among these nearest neighbor points;
2. Utilize the planar unit with the highest flatness coefficient as the initial clustering unit to search for connected planar units and incorporate them into the cluster. The added planar units must also fulfill the condition that the angle between their normal vector and the fitting plane normal vector of the existing points in the cluster is below the threshold;
3. Repeat step 2 until no more connected planar units are found, and conclude the clustering process;
4. Choose the planar unit with the highest flatness coefficient from the remaining planar units as the initial clustering unit for further clustering;
5. Iterate through steps 2-4 until all planar units have been processed.

Following the clustering of planar units, each resulting cluster can be considered as an individual discontinuity. The identification process of individual discontinuities comprehensively considers local point features and local plane features, effectively eliminating fragmented and undulating non-structural regions.

2.2.4. Individual discontinuity clustering

The number of sets is an important parameter for discontinuities. The clustering of individual discontinuities considers the structural characteristics of the rock mass in a more comprehensive manner compared to direct point clustering. This is achieved by removing non-structural regions during clustering, leading to more accurate clustering results. The fast k-means++ algorithm proposed by Wu et al. (2020) is highly efficient and capable of achieving automatic optimal grouping. However, it is based on point clustering. Here, a modified k-means++ algorithm is employed for weighted clustering based on individual discontinuities. The process involves the following steps:

1. Calculate the total number of the point clouds belonging to each individual discontinuity as the clustering weight of each individual discontinuity;
2. Select the individual discontinuity normal vector with the highest weight as the initial clustering center;
3. Choose the normal vector of the individual discontinuity that has the farthest distance from its closest known initial clustering center as an additional initial clustering center, and repeating this step until all k' initial clustering centers have been determined;
4. Divide the remaining individual discontinuities into clusters according to their nearest initial clustering centers, and k-means algorithm is applied to identify the final clustering center.

The distance d between different individual discontinuities in the clustering process is defined as follows:

$$d\left(n_i, n_j\right) = \arccos\left(\left|n_i \cdot n_j\right|\right) \tag{3}$$

Where, n_i and n_j represent the normal vectors of two individual discontinuities, respectively. To obtain the optimal number of sets k'_{opt}, a fast silhouette algorithm is adopted (Wu et al., 2020), which is efficient for large scale of data. When the silhouette value reaches its maximum, the corresponding number of sets k' can be considered as the optimal number of sets k'_{opt}.

2.3. Geometric analysis

Geometric analysis of the extracted individual discontinuities allows for obtaining their orientation, set, and trace information. Orientations are determined through a polar projection transformation of the normal vectors of individual discontinuities, resulting in the dip direction and dip angle. The sets are determined through the three-layer clustering process. To accurately represent the intersection between the discontinuities and the rock outcrop, the intersection lines formed by the intersection of the disk models representing the individual discontinuities and the rock outcrop are acquired. These intersection lines are subsequently projected onto a 2D plane to obtain the traces. The proposed trace mapping method extensively utilizes the extracted individual discontinuity information and ensures a one-to-one correspondence between the discontinuities and the exposed traces, preventing the occurrence of one discontinuity corresponding to multiple traces.

3. APPLICATION

This paper adopts point clouds of one tunnel face collected in railway tunnel in Tibet, China. Due to the complicate site conditions, it is very convenient to utilize digital photogrammetry technology combined with structure from motion (SfM) to acquire point clouds. Nine photos were taken in field from different angles, which were then used for 3D reconstruction to obtain point clouds. The CloudCompare software was used to pre-process the raw point clouds. The point clouds, which consist of 235681 points, are shown in figure 2.

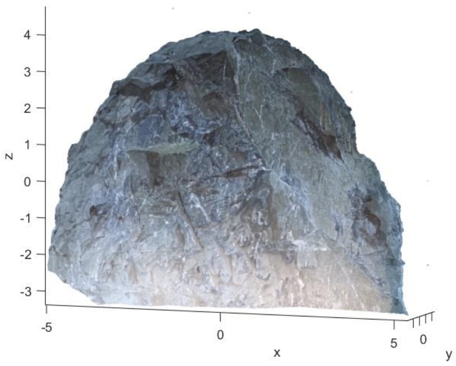


Figure 2. *Point clouds of the tunnel face.*

146 individual discontinuities are identified using the proposed method, while 513 individual discontinuities are identified using the fast k-means++. The discontinuities identified using the proposed method contain an average of 1081 points and a maximum of 48116 points, while the discontinuities identified using the fast k-means method contain an average of 459 points and a maximum of 76102 points. The results are shown in figure 3, where different colors represent distinct individual discontinuities identified by each method. The results demonstrate that the proposed method effectively eliminates a significant number of broken regions, whereas the fast k-means++ preserves these regions and generates numerous small non-structural planes.

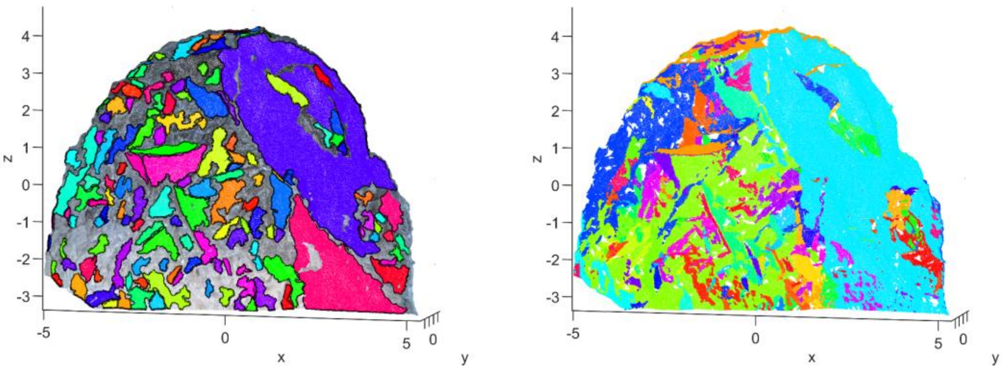


Figure 3. *Individual discontinuities extracted by (a) the proposed method, and (b) fast k-means++. (Each color represents an individual discontinuity.)*

The fast silhouette algorithm is used to calculate the fast silhouette values for different numbers of sets in both this method and the fast k-means++, as presented in figure 4. The results indicate that the optimal number of sets k'_{opt} for this method is 5, whereas for the fast k-means++, it is 3. Manual discrimination confirms that the tunnel face should possess 5 sets of discontinuities, providing evidence that clustering based on individual discontinuities outperforms clustering based on point normal vectors.

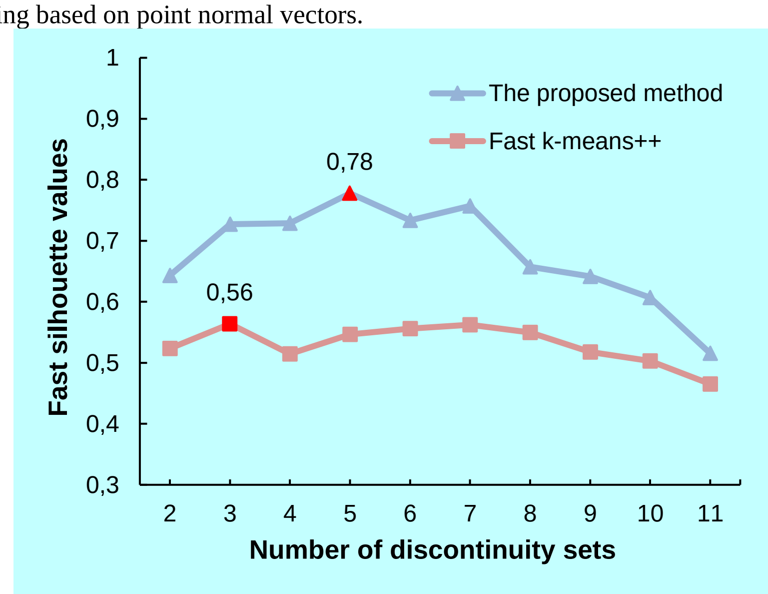
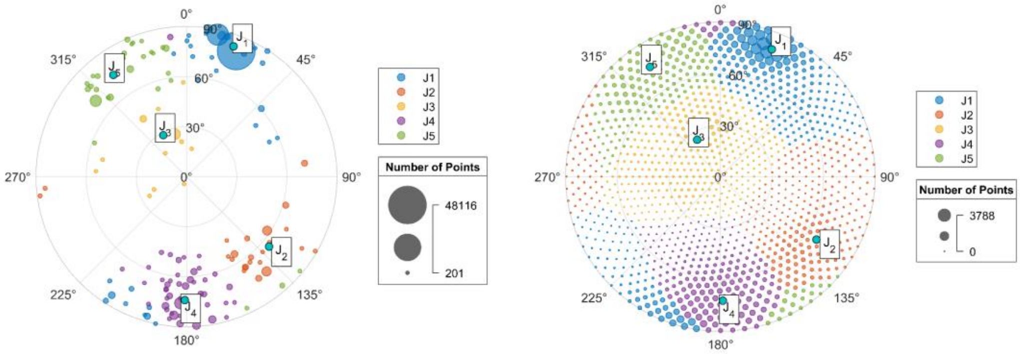


Figure 4. *Comparison between the proposed method and the fast k-means++.*

To facilitate comparison, the optimal number of sets k'_{opt} is selected as 5 for both the proposed method and fast k-means++ (figure 5). The orientation results of each set are presented in table 1. It is evident that the dip directions and dip angles of each set exhibit no significant differences, with a maximum dip direction deviation of 7 degrees and a maximum dip angle deviation of 3.1 degrees.



(a) (b)

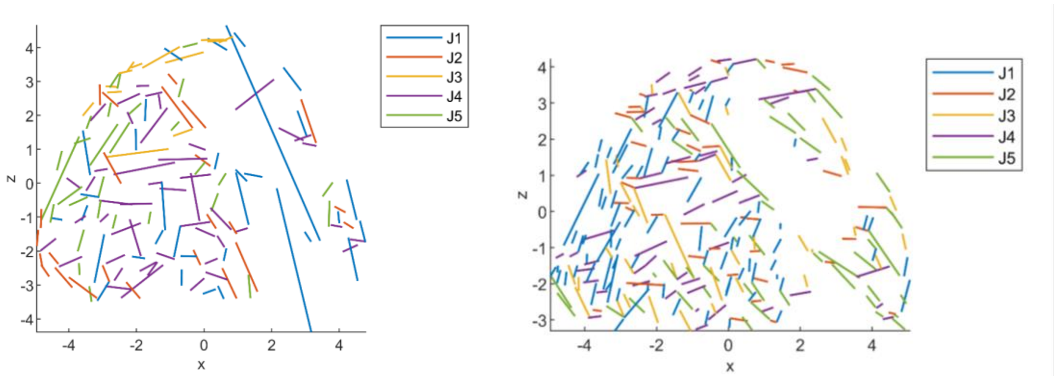
Figure 5. *Cluster results of the tunnel face using (a) the proposed method, and (b) fast k-means++.*

Table 1. *Comparison of discontinuity orientations using the proposed method with the fast k-means++.*

Set	The proposed method		Fast k-means++		$\Delta\|\delta\|$	$\Delta\|\theta\|$
	δ/θ (°)	Number of discontinuities	δ/θ (°)	Number of discontinuities		
J_1	19.8/82.9	33	22.0/79.8	108	2.2	3.1
J_2	130.3/64.7	24	123.3/66.7	128	7.0	2.0
J_3	330.7/28.4	13	327.7/25.3	84	3.0	3.1
J_4	180.8/74.0	47	178.9/72.2	86	1.9	1.8
J_5	324.3/75.0	29	327.3/75.8	107	3.0	0.8

Note: $\Delta|\delta|$ and $\Delta|\theta|$ represent the dip direction and dip angle differences respectively.

This method obtains 146 traces by finding the intersection line formed by the intersection of the disk models representing discontinuities and the exposed tunnel face. According to the method of Zhang et al. (2020), 265 3D traces are obtained based on 3D curvature and projected onto the 2D plane of the tunnel face. The average length of the traces identified by this method is 0.77m and the maximum length is 6.93m, while the results of Zhang et al. (2020) have an average length of 0.54m and a maximum length of 2.22m. A comparison (figure 6 and figure 7) reveals that the intersection method is superior in capturing longer traces that represent discontinuities. Conversely, the method based on 3D curvature yields numerous short traces, many of which are contour segments of the same discontinuity.



(a) (b)

Figure 6. *Trace mapping using (a) the proposed method, and (b) the method of Zhang et al. (2020).*

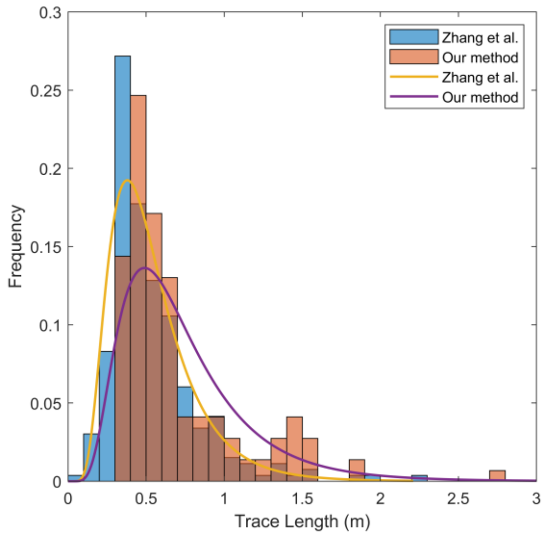


Figure 7. *The trace length histograms and fitted log-normal distribution using the proposed method and the method of Zhang et al. (2020).*

4. DISCUSSION

According to the analyses and comparisons of the results, the proposed method shows excellent accuracy in automatically identifying discontinuities in the rock mass. A major advantage of this method over the fast k-means++ algorithm, which directly clusters discontinuities based on point features, is that it directly identifies individual discontinuities without the need for prior point grouping, thus avoiding errors in identifying individual discontinuities due to incorrect grouping. At the same time, areas of fragmentation due to blasting or weathering are removed and do not affect the identification of individual discontinuities.

For individual discontinuity clustering, the proposed method inherits well the advantages of the fast k-means++ algorithm for the density-based pre-selection of initial clustering centroids. In addition, fast silhouette algorithm employed helps to determine the optimal clustering automatically. Furthermore, the trace generated from the intersection between the disk model and the exposed surface of the rock mass is more consistent with the facts.

It should be noted that this method additionally considers the nearest neighbor curvature information to accurately identify individual discontinuities, but has not yet fully considered structural features, which may lead to the presence of random discontinuities in the results. Regarding computational efficiency, the proposed three-layer clustering method involves more aggregation operations than standard clustering algorithms. For large-scale point clouds (approaching one million points), the computation time increases from tens of seconds to several minutes. While this is a notable increase, the performance still meets the real-time needs of most engineering applications, representing a reasonable trade-off for the higher accuracy achieved.

The proposed algorithm is designed to be robust in complex geological conditions, such as those with significant fragmentation from blasting or weathering. Its core strength lies in the initial filtering of non-structural regions, which allows it to isolate and identify major discontinuities even in a noisy environment. The quality of the input data also plays a role. While the core point identification step provides some resilience to noise and variations in point density, extremely sparse or low-quality data could still affect the reliability of normal vector calculations and, consequently, the accuracy of the final discontinuity extraction.

5. CONCLUSION

A three-layer clustering method for extracting individual discontinuities is presented in this paper. The effectiveness of the proposed method is demonstrated through its application to a tunnel face case and comparison with existing methods. The proposed method efficiently eliminates non-structural regions and identifies individual discontinuities. Furthermore, the results of the optimal grouping calculation based on individual discontinuity align more closely with manual judgment than those based on point normal vectors, underscoring the significance of prioritizing the identification of individual discontinuities. It is worth noting that the extraction process currently does not utilize point cloud color information. Future research endeavors will enhance the accuracy of individual discontinuity identification by integrating color and texture information.

6. ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China [42272338, 41827807, 41902275]; Shanghai Sailing Program [18YF1424400]; Joint Fund for Basic Research of High-speed Railway of National Natural Science Foundation of China, China Railway Corporation [U1934212]; China State Railway Group Co., Ltd. [P2019G038]; Department of Transportation of Zhejiang Province [202213]; China Railway First Survey and Design Institute Group Co., Ltd. [19-21-1, 2022KY53ZD(CYH)-10]; China Railway Tunnel Group Co., Ltd. [CZ02-02-08]; PowChina Hebei Transportation Highway Investment Development Co., Ltd. [TH-201908]; Sichuan Railway Investment Group Co., Ltd. [SRIG2019GG0004]; The Science and Technology major program of Guizhou Province [2018]3011.

7. REFERENCES

[1] Assali, P., Grussenmeyer, P., Villemin, T., Pollet, N., & Viguier, F. (2016). Solid images for geostructural mapping and key block modeling of rock discontinuities. *Computers & Geosciences*, *89*, 21–31. https://doi.org/10.1016/j.cageo.2016.01.002.

[2] Chen, J., Huang, H., Zhou, M., & Chaiyasarn, K. (2021). Towards semi-automatic discontinuity characterization in rock tunnel faces using 3D point clouds. *Engineering Geology*, *291*, 106232. https://doi.org/10.1016/j.enggeo.2021.106232.

[3] Chen, J., Zhu, H., & Li, X. (2016). Automatic extraction of discontinuity orientation from rock mass surface 3D point cloud. *Computers & Geosciences*, *95*, 18–31. https://doi.org/10.1016/j.cageo.2016.06.015.

[4] Drews, T., Miernik, G., Anders, K., Höfle, B., Profe, J., Emmerich, A., & Bechstädt, T. (2018). Validation of fracture data recognition in rock masses by automated plane detection in 3D point clouds. *International Journal of Rock Mechanics and Mining Sciences*, *109*, 19–31. https://doi.org/10.1016/j.ijrmms.2018.06.023.

[5] Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Kdd*, *96*(34), 226–231.

[6] Ge, Y., Cao, B., & Tang, H. (2022). Rock Discontinuities Identification from 3D Point Clouds Using Artificial Neural Network. *Rock Mechanics and Rock Engineering*, *55*(3), 1705–1720. https://doi.org/10.1007/s00603-021-02748-w.

[7] Hu, L., Xiao, J., & Wang, Y. (2020). Efficient and automatic plane detection approach for 3-D rock mass point clouds. *Multimedia Tools and Applications*, *79*(1), Article 1. https://doi.org/10.1007/s11042-019-08189-6.

[8] Kong, D., Wu, F., & Saroglou, C. (2020). Automatic identification and characterization of discontinuities in rock masses from 3D point clouds. *Engineering Geology*, *265*, 105442. https://doi.org/10.1016/j.enggeo.2019.105442.

[9] Singh, S. K., Banerjee, B. P., Lato, M. J., Sammut, C., & Raval, S. (2022). Automated rock mass discontinuity set characterisation using amplitude and phase decomposition of point cloud data. *International Journal of Rock Mechanics and Mining Sciences*, *152*, 105072. https://doi.org/10.1016/j.ijrmms.2022.105072.

[10] Singh, S. K., Banerjee, B. P., & Raval, S. (2023). A review of laser scanning for geological and geotechnical applications in underground mining. *International Journal of Mining Science and Technology*, *33*(2), 133–154. https://doi.org/10.1016/j.ijmst.2022.09.022.

[11] Singh, S. K., Raval, S., & Banerjee, B. P. (2021). Automated structural discontinuity mapping in a rock face occluded by vegetation using mobile laser scanning. *Engineering Geology*, *285*, 106040. https://doi.org/10.1016/j.enggeo.2021.106040.

[12] Wu, W., Zhang, K., & Zhu, H. (2020). A fast automatic extraction method for rock mass discontinuity orientation using fast k-means++ and fast silhouette based on 3D point cloud. *IOP Conference Series: Earth and Environmental Science*, *570*(5), 052075. https://doi.org/10.1088/1755-1315/570/5/052075.

[13] Zhang, K., Wu, W., Zhu, H., Zhang, L., Li, X., & Zhang, H. (2020). A modified method of discontinuity trace mapping using three-dimensional point clouds of rock mass surfaces. *Journal of Rock Mechanics and Geotechnical Engineering*, *12*(03), Article 03.