

Article

An Integrated Approach for Accurate 3D Point Cloud Recognition of Substation Equipment Using PSO and kNN Classification

Yi Yang¹, Yiwei Sun^{1,*}, Fengda Zhang¹ and Meng Liu¹

¹ Transmission and Transformation Technology Center, State Grid Shandong Electric Power Research Institute, Jinan 250003, Shandong, China

* **Correspondence:** sddkysyw@163.com

Abstract: Low efficiency and poor accuracy are caused by missing data in traditional 3D reconstruction methods. This study suggests a new 3D point cloud recognition technique for substation equipment based on 3D laser scanning point clouds, which combines the k-nearest neighbour (KNN) classification algorithm and particle swarm optimisation (PSO) algorithm, to address these issues. The particle swarm optimisation algorithm optimises the coefficient weights of each subspace feature. The k-nearest neighbour classification algorithm is then used to finish the classification. To confirm the superiority and accuracy of the suggested approach, the impact of the point cloud subspace's size and loss rate on the recognition effect is examined experimentally and contrasted with the enhanced iterative nearest point algorithm. With an average recognition time of 0.19 seconds and a recognition accuracy of over 95%, the experimental results demonstrate the method's good performance in terms of efficiency and accuracy, opening up a wide range of potential applications.

Keywords: Substation equipment, PSO, KNN, 3D point cloud recognition, 3D reconstruction

1. Introduction

The development of laser scanning technology has made 3D point clouds a crucial tool for gathering real-world scenes. Nonetheless, a popular topic in recent research is how to effectively use 3D point cloud data for equipment management and identification [1–3]. The high requirements for data integrity in traditional 3D reconstruction methods make it difficult to achieve ideal results in real scenes, which hinders their widespread application in power systems [4].

The safe and steady operation of the substation's equipment is essential to the dependability of the entire power grid because it is a vital link in the power system that facilitates the transmission and distribution of electric energy. The quantity and complexity of substation equipment are growing along with the power system's scale. The manual inspection method used in the traditional equipment management approach has numerous drawbacks, such as low efficiency, high cost, ease of miss-take, etc.

Researchers both domestically and internationally are currently devoting a great deal of time to studying substation equipment's 3D identification technology. suggested an algorithm that employs a 3D laser to gather point cloud data about the equipment, improves the Ada Boost algorithm with a support vector machine for identification, and simulates the results to confirm the algorithm's superiority. The findings demonstrate that there is good recognition accuracy with training and testing

accuracy of over 98%. In order to identify device point cloud data, [5] suggested a Hough voting algorithm. To determine the targets' similarity, the distance histograms in the model library were compared. The similarity threshold is used to determine the initial recognition results. [6] describes how to use a panoramic camera and a 3D scanning system to get multi-angle scene photos and point cloud data of outdoor equipment and substations. Using Autodesk 3d-Max tools, the substation equipment, buildings, and the surrounding terrain environment are modelled in a 3D scene in accordance with predetermined modelling standards. [7] suggested an early warning system that uses UAVs and laser scanning to detect the surroundings of substations intelligently.

Nonetheless, there are still certain obstacles to be solved in the field of 3D point cloud recognition of substation equipment. One such issue is missing data, where the actual 3D point cloud data that have been collected are frequently absent, particularly in some intricate areas of the equipment surface. It is imperative that the missing data be handled and the model's ability to adjust to partial data be improved. Due to the variety of device shapes and materials, one of the difficulties in designing a generic and accurate recognition model is recognition accuracy. Targeted processing might be necessary for various equipment types and conditions. Real-time requirements: It's crucial to monitor equipment status in real-time in power systems. More effective algorithms and models are required because traditional approaches might not be able to meet the demand for quick response to equipment changes [8].

In this paper, we propose a PSO algorithm combined with kNN classification algorithm for 3D point cloud recognition of substation equipment, aiming to solve the issues in traditional methods based on 3D laser scanning point clouds. To increase classification accuracy, we optimise the coefficient weights of every subspace feature in the point cloud data by implementing the particle swarm optimisation algorithm.

The following are some benefits of the 3D point cloud recognition technique presented in this paper, which is based on the kNN classification algorithm and PSO algorithm:

- **Efficiency:** The point cloud data's features are optimised through the introduction of the PSO algorithm, which raises the classification's accuracy and efficiency.
- **Generality:** The model can be customised to fit various kinds and shapes of equipment thanks to the combination of the kNN classification algorithm.
- **Experimental evidence:** Based on numerous experiments, the approach suggested in this paper outperforms the conventional iterative nearest point algorithm in terms of efficiency and recognition accuracy.

2. 3D Identification of Substation Equipment

The subspace feature of the point cloud is the cosine value of the angle between the centre of mass of the entire point cloud and the line connecting the centres of mass of each subspace point cloud with respect to the positive direction of the Z-axis. This allows for the identification of the 3D point cloud of substation equipment. The subspace is divided according to the point features, and the features of the subspace are extracted to form the feature vector. The PSO algorithm is used to optimise the coefficient weights of each subspace feature, and the kNN is used to obtain the classification results.

2.1. kNN Classification Algorithm

The kNN algorithm's basic idea is that a sample is included in a category if it has k nearest neighbours in the space that are also included in that category [9]. The eigenvalue distance is used in the kNN algorithm to classify a subset of neighbouring kNNs.

The kNN algorithm's classification is influenced by the value of k . The test sample will count the categories in a narrow range and is prone to classification errors if k is very small [10]. In general, the cross-test is used to determine the k value. If k is large, the test samples will count the number

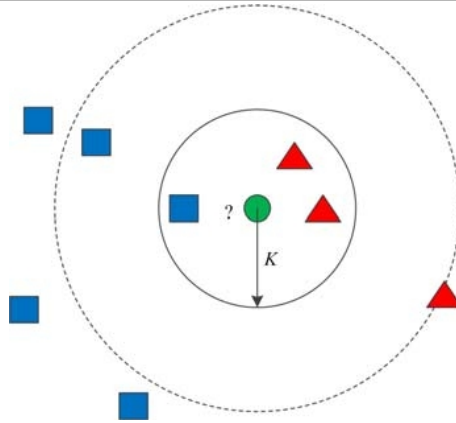


Figure 1. K-nearest Neighbour Decision Diagram

of categories in a larger range, with more statistical information, which can avoid classification errors. However, the classifier efficiency is poor. Figure 1 displays the k Nearest Neighbours Decision Diagram.

The test data is input and the distance between the features of the test data and the features of the template is calculated to find the k most similar data, assuming that the data in the template and labels are known [11]. The class of the test data is the one with the greatest number of k values. The following are the steps involved in implementing the algorithm:

- 1) Calculate the distance d between the test data T and the template data MT , and choose the Euclidean distance as shown in Eq. (1), [12]:

$$d(T, MT) = \sqrt{\sum_{i=1}^n (T_i - MT_i)^2}. \quad (1)$$

- 2) Sorting by distance, with k set to 3, yields the first k values with the smallest distance [13].
- 3) Determine how often each of the k sample categories occurs.
- 4) In the test data's predicted classification, the k samples are assigned to the category with the highest frequency. The category with the smallest distance is the predicted category if the frequencies are equal.

2.2. Weights for PSO Optimisation Subspace Features

PSO is a global evolutionary heuristic algorithm for multi-objective, nonlinear, and multivariate problems that is based on group intelligence [1]. To find the optimal solution, the PSO algorithm iterates over the initial particles, updates the particle positions and velocities, and tracks the optimal particles in the space. Usually, reaching the desired number of iterations is the criterion for ending the iteration. PSO's benefits of simple programming, quick convergence, and great accuracy make it a popular choice. The PSO algorithm is used in this paper to optimise the feature weights within each subspace. In Figure 2, the flowchart is displayed.

Modifying the coefficient weights of each subspace feature improves the recognition and classification accuracy of features. The classification error rate is defined as a fitness function when employing PSO to optimise the feature weights, and the final positions of the particles represent the weights occupied by the corresponding features. The following is the procedure for implementing an algorithm:

- 1) Set the particle's initial position and speed.
- 2) Update the particle's position and velocity in accordance with Eqs (2), (3) and (4), [14].

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (pbest_{id} - x_{id}^k) + c_2 r_2 (gbest - x_{id}^k). \quad (2)$$

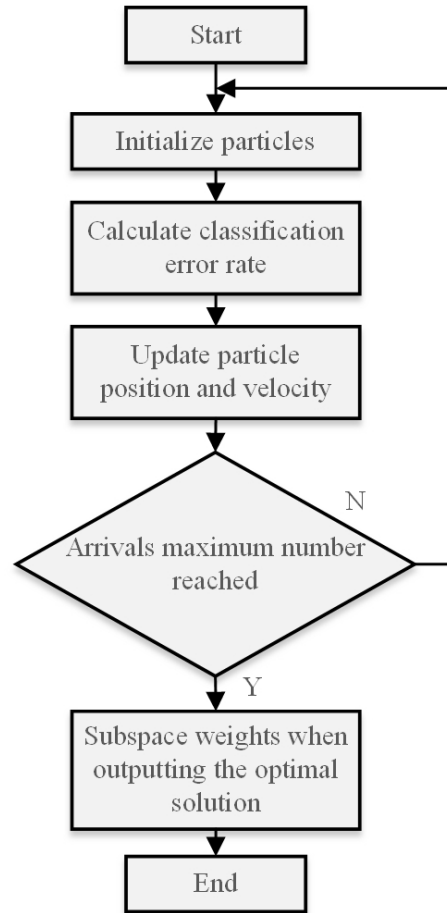


Figure 2. PSO Algorithm Flow

$$x_{id}^{k+1} = x_{id}^k + V_{id}^{k+1}. \quad (3)$$

Here, the coefficients for self-learning and social learning are c_1 and c_2 , random numbers between 0 and 1 are r_1 and r_2 , the position and speed of the k -th iteration are x_{id}^k and V_{id}^k , respectively, and the individual and global optimal points in the population's evolutionary process are $pbest_{id}$ and $gbest$, respectively.

- 3) Calculate the fitness function according to Eq. (4).

$$\text{fitness} = 1 - \frac{m}{N}, \quad (4)$$

where N is the total number of test data; m is the number of correctly recognised data; and function is the misrecognition error rate.

- 4) Both global and individual optimality are being updated.
- 5) When the number of iterations is reached, you should stop and output the parameters [15]. Repeat steps 2) through 4) until the maximum number of iterations is reached if the number of iterations has not been reached.

3. Analysis of Experimental Results

3.1. Experimental Parameters

Following the creation of the template feature library, the kNN algorithm is used to classify the devices and the PSO algorithm is used to optimise the weights of the coefficients of each subspace feature in order to increase the accuracy of classification and recognition. The computer has an 8GB RAM and an i5-2450m CPU. Matlab 2014 is the software for image processing.

Device Type	Length/m	Wide/m	High/m
500kV _{CB}	1.457	5.124	6.245
500kV _{CT}	1.326	1.368	6.128
500kV _{DS}	1.478	13.456	8.256
500kV _{ES}	1.568	3.856	5.678
500kV _{LA}	1.824	1.534	6.538
500kV _{PT}	0.985	1.002	6.035
500kV _{LT}	2.215	2.231	6.854

Table 1. Equipment Dimensions

Side length	Number	Before optimization		After PSO optimization	
1.01	55	70.03	0.065	83.25	33.02
0.9	126	79.32	0.035	88.26	45.01
0.6	252	79.85	0.025	88.25	62.08
0.5	432	78.06	0.028	88.06	91.85
0.4	960	88.12	0.032	95.34	120.18

Table 2. Classification Effect (Different Subspace Sizes)

100 test data points are the maximum number of iterations that the PSO algorithm can run. Circuit breakers (CB), current transformers (CT), disconnect switches (DS), earth switches (ES), lightning arresters (LA), voltage transformers (PT), traps (LT), etc. are among the substation equipment dimensions that are statistically displayed in Table 1, [2].

Three factors are examined in order to confirm the algorithm's superiority: 1) the impact of the point cloud subspace's size on recognition outcomes; 2) the impact of the point cloud's absence on recognition outcomes; and 3) a comparison with the enhanced ICP algorithm.

3.2. Effect of Subspace Size on Recognition Results

In theory, the number of subspaces within the point cloud indicates the higher the dimension of the feature vector; conversely, the smaller the subspace of the point cloud, the smaller the side lengths of the subspace. In comparison to standard devices, the results of classification and recognition are more accurate, and vice versa [6].

The following describes how edge length affects various subspaces. The point cloud data was normalised and all point clouds were scaled using a 3*3*6 cube before subspace segmentation. Next, if distinct subspace edge lengths are selected, there will be variations in the number of subspaces, point cloud feature vector lengths, and classification effects. The classification results for subspaces with side lengths of 1.0, 0.8, 0.6, 0.5, and 0.4 are displayed in Table 2.

Before and after PSO optimisation, Table 2 computes the classification accuracy and time for various subspace sizes. Figure 3 displays the accuracy of classification both before and after PSO optimisation for various subspace edge lengths, illustrating the impact of subspace size on the classification outcomes.

Figure 3 demonstrate the notable improvement in classification accuracy following PSO optimisation. Although the PSO algorithm's iterative process requires time, it can be utilised to optimise subspace feature weights and increase classification accuracy within a reasonable time frame. It is easier to have a "dimensional catastrophe" the shorter the edge length, the more subspaces there are overall, and the larger the feature vector dimensions.

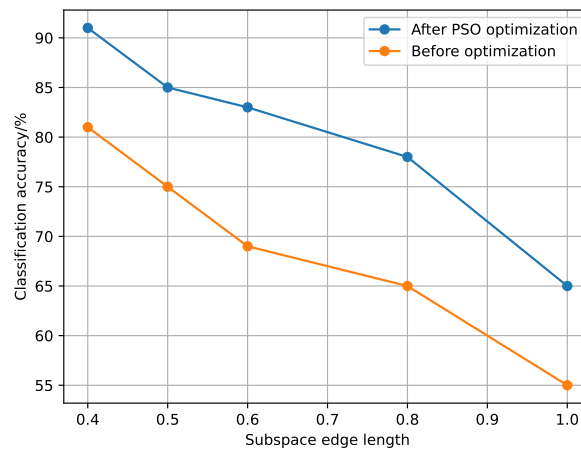


Figure 3. Classification Accuracy of Different Edge Lengths Before and After Optimization

Device name	Improve ICP		Method in the text	
	Yes	Average	Yes	Average
500KV _C B	Yes	82.05	Yes	0.15
500KV _C T	Yes	56.67	Yes	0.18
500KV _D S	Yes	150.86	Yes	0.22
500KV _E S	Yes	130.54	Yes	0.25
500KV _L A	Yes	92.56	Yes	0.19
500KV _P T	Yes	71.25	Yes	0.22
500KV _L T	Yes	126.85	Yes	0.20
...
	Accuracy	Average	Accuracy	Average
	98.25%	85.32%	97.21%	0.20

Table 3. Comparison of Different Algorithm Recognition

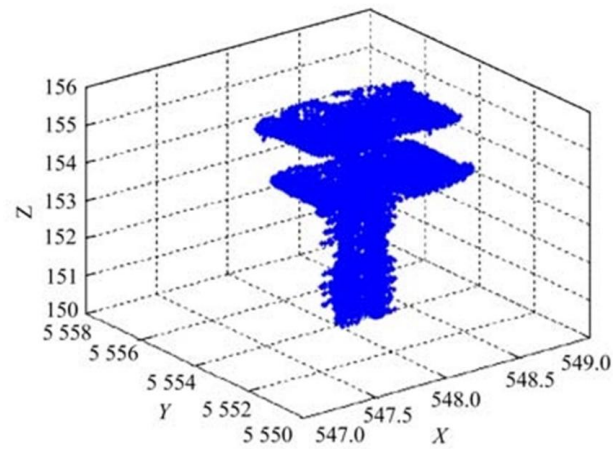
3.3. Impact of Missing Point Clouds on Recognition Results

The paper analyses the point cloud missing performance of the identification methods. Ideal point cloud data is acquired by the substation apparatus. Using a circuit breaker as an example, the text states that 10%, 20%, and 30% of cloud data are lost at random points in time. As seen in Figure 4, the aforementioned three cases are recognised and categorised using the identification technique described in the paper. It can be correctly identified when 10% or 20% of the data are missing, and incorrectly identified when 30% of the data are missing.

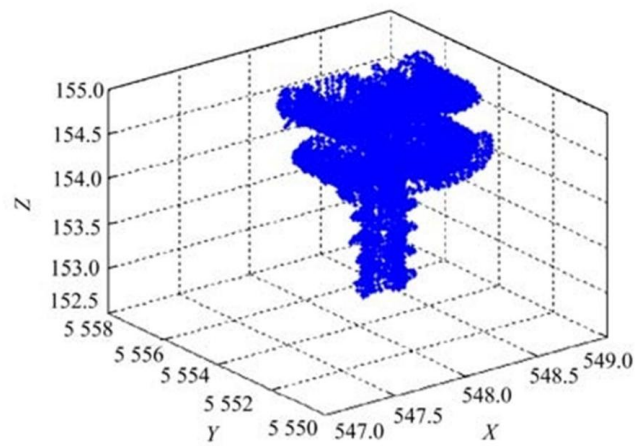
The point cloud was missing too much to be satisfactorily identified, and the method used to conduct the missing test for all test data resulted in an average identifiable missing rate of 13%. All test data has a minimum identifiable missing of 10%, with the majority falling within the 10%–13% range. Consequently, low point cloud missing occasions can be handled by the recognition method described in this paper.

3.4. Algorithm Comparison

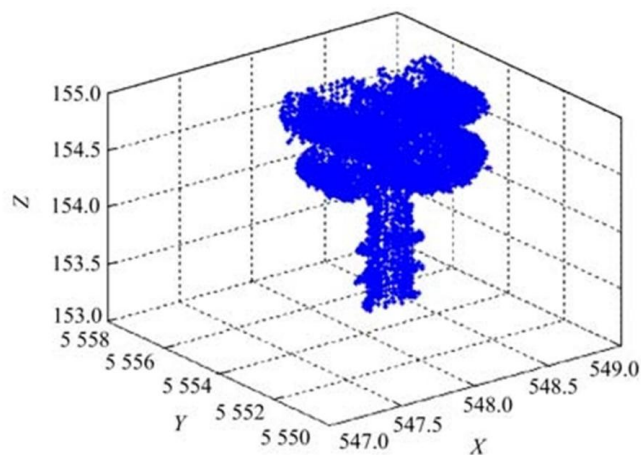
The algorithm in the paper is compared with the enhanced Iterative Closet Points (ICP) algorithm to further demonstrate the algorithm’s efficacy. Based on the traditional ICP algorithm’s higher alignment accuracy, the improved ICP algorithm makes use of the projected contour pre-selection model, which drastically cuts down on recognition time and is currently one of the more effective recognition techniques available. Table 3 displays the recognition results. The subspace has a side length of 0.4 and requires the recognition of 100 point cloud devices.



(a) Missing 10% of 3D views



(b) Missing 20% of 3D views



(c) Missing 30% of 3D views

Figure 4. Point Cloud Missing 3D Map

As shown in Table 3, out of the 100 point cloud devices that needed to be identified, 98 were correctly identified by the improved ICP identification method, with only 2 incorrectly identified. With an accuracy of 97.00%, the paper correctly identified 97 devices and incorrectly identified 3. Regarding the algorithm's recognition time, the suggested method takes one second, while the enhanced ICP takes more than a minute. The suggested method has a faster and more accurate recognition rate, making it perfect for real-world applications.

4. Conclusions

This study proposes a 3D point cloud recognition method for substation equipment based on the particle swarm optimisation algorithm and the k-nearest neighbour classification algorithm. The method performs better in experiments than the conventional iterative nearest point algorithm, with higher recognition accuracy and faster processing speed. This approach has potential applications in the real world and can help advance power system equipment management. In order to better meet the needs of power system equipment management, future research can further optimise the algorithm, expand the scenarios that are applicable, and improve real-time performance. However, the challenges of handling complex scenarios and improving data robustness still need to be solved in practical applications.

Funding

This paper is sponsored by Science and Technology Project of State Grid Shandong Electric Power Company "Research and application of key technologies for panoramic operation and maintenance of digital twin substation" (520626220009).

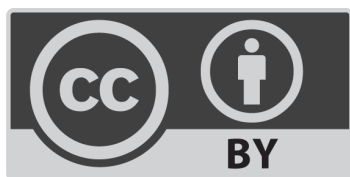
Conflict of Interest

The authors declare no conflict of interests.

References

1. Zhou, J., Liu, Q., Jiang, Q., Ren, W., Lam, K.M. and Zhang, W., 2023. Underwater camera: Improving visual perception via adaptive dark pixel prior and color correction. *International Journal of Computer Vision*, pp.1-19.
2. Nassif, A.B., Talib, M.A., Nasir, Q. and Dakalbab, F.M., 2021. Machine learning for anomaly detection: A systematic review. *Ieee Access*, 9, pp.78658-78700.
3. Ren, X., Ahmed, I. and Liu, R., 2023. Study of Topological Behavior of Some Computer Related Graphs. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 117, pp.3-14.
4. Nazeer, S., Sultana, N. and Bonyah, E., Cycles and Paths Related Vertex-Equitable Graphs. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 117, pp.15-24.
5. Zeng, Y. and Chu, B., The Appropriate Scale of Competition Between Online Taxis and Taxis Based on the Lotka-Volterra Evolutionary Model. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 117, pp.25-36.
6. Zhang, C., Roh, B.H. and Shan, G., 2023, December. Poster: Dynamic Clustered Federated Framework for Multi-domain Network Anomaly Detection. In Companion of the 19th *International Conference on emerging Networking EXperiments and Technologies* (pp. 71-72).
7. Tajanpure, R. and Muddana, A., 2021. Circular convolution-based feature extraction algorithm for classification of high-dimensional datasets. *Journal of Intelligent Systems*, 30(1), pp.1026-1039.

8. Bécue, A., Praça, I. and Gama, J., 2021. Artificial intelligence, cyber-threats and Industry 4.0: Challenges and opportunities. *Artificial Intelligence Review*, 54(5), pp.3849-3886.
9. Zhou, J., Pang, L., Zhang, D. and Zhang, W., 2023. Underwater image enhancement method by multi-interval histogram equalization. *IEEE Journal of Oceanic Engineering*, 48(2), pp.474-488.
10. Mallikarjuna, P.B., Sreenatha, M., Manjunath, S. and Kundur, N.C., 2020. Aircraft gearbox fault diagnosis system: an approach based on deep learning techniques. *Journal of Intelligent Systems*, 30(1), pp.258-272.
11. Zhang, H., Li, Q., Wang, H., Li, Q., Qin, G. and Wu, Q., 2022. A review of energy management optimization based on the equivalent consumption minimization strategy for fuel cell hybrid power systems. *Fuel Cells*, 22(4), pp.116-130.
12. Wang, Y., 2021. Intellectualization of the urban and rural bus: The arrival time prediction method. *Journal of Intelligent Systems*, 30(1), pp.689-697.
13. Mohammed, M.A., Ibrahim, D.A. and Salman, A.O., 2021. Adaptive intelligent learning approach based on visual anti-spam email model for multi-natural language. *Journal of Intelligent Systems*, 30(1), pp.774-792.
14. Ma, Y. and Sharaf, S., 2021. Analysis and synthesis of function data of human movement. *Applied Mathematics and Nonlinear Sciences*, 7(1), pp.917-926.
15. Tajanpure, R. and Muddana, A., 2021. Circular convolution-based feature extraction algorithm for classification of high-dimensional datasets. *Journal of Intelligent Systems*, 30(1), pp.1026-1039.



©2024 the Author(s), licensee Combinatorial Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)