Similarity Analysis of 3D CAD Model based on Optimal Matching Algorithm

Zhiqiang Zhang[1*], Hongying Xue[2] and Yifan Zhang[3]

[1]Centre for Continuing Education, Linzhou Architectural Vocational and Technical College, Anyang, Henan 455000, China, zzqeiyk@163.com
[2]Department of basic courses, Linzhou Architectural Vocational and Technical College, Anyang, Henan 455000, China, xhyeiyk@163.com
[3]Anyang No.1 Middle School,Anyang, Henan 455000, China, zyfzbd@126.com

Corresponding author: Zhiqiang Zhang, zzqeiyk@163.com

Abstract. This paper studies the feature extraction method of CAD (Computer Aided Design) model. The attribute adjacency graph is used to represent the three-dimensional model and the calculation method for the similarity of the model faces is given. The difference in the number of sides is used to calculate the shape similarity of the faces, and the face similarity matrix is constructed. The adjacency correspondence of model faces is introduced, and the similarity calculation method of the adjacency correspondence is given. The focus is on the calculation process of the shape similarity of the model faces, which lays the foundation for model retrieval. The algorithm in this paper supplements the classic shape distribution algorithm, optimizes the random point selection method, reduces the number of sampling points and increases the area ratio feature between the model surface and the surface, which can greatly improve the algorithm time while improving the resolution of the complex model. Finally, offline feature extraction technology and online model matching technology are used to evaluate the overall similarity of the 3D CAD model. At the same time, other strategies and algorithms are used to search for the optimal matching faces of the two models, and detailed face matching methods, flowcharts and experimental procedures are given. After the algorithm summarizes and studies the geometric topological properties of the model, it builds the corresponding adjacency matrix based on the attribute adjacency graph, changes the search idea, and converts the problem of finding small graphs in large graphs into the use of simulated annealing algorithm to maximize the associated attribute graph of the model.

Keywords: Optimal matching; three-dimensional model; CAD; similarity analysis
DOI: https://doi.org/10.14733/cadaps.2022.S4.135-145

1 INTRODUCTION

The development of the processing and manufacturing industry has resulted in mass production and use of model components, which has led to the development and application of CAD model retrieval technology. In the process of model processing and application, how to efficiently retrieve and reuse the CAD model that users need has always been a key problem. As the requirements become more complex, the difficulty and complexity of model retrieval are correspondingly increased, which makes the existing model retrieval technology no longer able to meet the needs of users. Therefore, it is more important to seek better model retrieval methods and explore new model retrieval techniques. With the transformation and upgrading of manufacturing to digital and intelligent "intelligent manufacturing", three-dimensional models have become the core of product design, manufacturing and analysis. By reusing existing models, design costs can be effectively reduced, product innovation can be accelerated, and enterprise competitiveness can be improved. At present, the accumulation of a large number of 3D models has put forward higher requirements for product data management and reuse modes. Content-based 3D model similarity assessment has become a key technology for intelligent design and product data management and reuse. Three-dimensional models are more and more widely used because they contain multi-level information collections and rich visual details: product lifecycle management based on three-dimensional digital models is an important development trend in modern product design and manufacturing; at the same time, companies are in the production process Accumulating a large number of rich 3D models, studying how to quickly and effectively find the required model from the massive CAD model of the product, fully digging and discovering the information and knowledge contained in it, and reusing it, is to speed up the product development process. One of the important ways is to shorten the product manufacturing response cycle and improve the quality of product development.

2 RELATED STUDIES

3D model similarity evaluation is the basis of many applications, and it is also a hot and difficult point in interdisciplinary research. Similarity evaluation has a wide range of applications, and many applications can be derived only through model retrieval. For example, through similarity comparison, instance-based modeling, scene modeling, model collaborative design management, automatic model classification, model navigation, and auxiliary processes can also be realized. Planning and processing quotation, etc. Deng et al. [1] proposed a shape distribution algorithm, which constructs an easy-to-compare probability distribution curve by defining the shape function related to the geometric attributes of the model sampling points. After comparing several shape function experiments, it is found that using the distance between two random points on the model as the shape function is stable and efficient. This method has low requirements on the quality of the 3D model, is insensitive to noise and non-manifolds, and is suitable for rough clustering of models. Lin et al. [2] applied shape distribution to physical model retrieval in the field of mechanical CAD, and found that models that are generally similar but have very different details and characteristics have similar shape distribution curves and tend to be normal. Therefore, the direct application of the shape distribution algorithm has the disadvantage of insufficient discrimination. Jayanti et al. [3] made improvements on the basis of the shape distribution algorithm by using the characteristics of the solid model. According to whether the line of the sampling point pair passes through the inside of the model, the probability distribution curve is divided into three types: internal, external and mixed to increase the discrimination of the model. Due to the large amount of calculation in the intersection algorithm, Zhao et al. [4] proposed to construct the shape distribution curve of the model bounding box on the basis of calculating the shape distribution curve of the model itself, and the difference was used as the convex hull difference curve of the model. Experiments show that this method has a good degree of discrimination for models with internal characteristics (such as thin-walled parts). Ambrosini et al. [5] proposed a method to improve the shape distribution method by using the symmetry of the three-dimensional model. Based on the analysis of the main plane of the model, this method judges the positional relationship between the pair of sampling points and the main plane, including two points on the same side of the main plane positively, on the same side or different sides of the main plane, and divide the shape distribution curve into the three categories which improve the discrimination of the similarity comparison of symmetric models.

Mian et al. [6] carried out attribute labeling on the basis of the adjacent graph of the surface, in which the node records the type of the surface such as plane, conical surface, etc., and the edge records the attributes such as length and unevenness. On this basis, geometric invariants are extracted to form feature vectors, which are used to filter irrelevant models. We use the eigenvalue spectrum of the graph topology to recursively perform subgraph segmentation, and use adjacency matrix instead of graph matching to evaluate similarity. Some researchers have adopted similar attribute tags and introduced constraint conditions in the matching stage [7, 8]. For example, similar faces require the same type, and similar edges require attribute matching. The introduction of constraint conditions simplifies the process of inexact graph matching, and the similarity calculation can be performed only by using the number of matching nodes and edges. Some scholars divide the plane adjacent graph hierarchy into a hierarchical structure centered on a certain plane, and use numerical calculation instead of graph matching to simplify the calculation, but the selection of the center plane requires interactive definition [9]. Simultaneously, the simulated annealing algorithm is used for precise matching to find the largest common subgraph to determine the local feature similarity. However, the similarity calculation only considers the number of nodes. The input local model features are expressed as subgraphs, and the model in the design warehouse is expressed as the big picture, in order to find matching subgraphs in the big picture, a local search method of Lie group gradient flow is proposed [10].

3 3D CAD MODEL CONSTRUCTION BASED ON OPTIMAL MATCHING ALGORITHM

3.1 3D CAD Framework

The 3D CAD solid model is a typical mesh representation, composed of disordered spatial triangles, which is an approximation of the precise model. Figure 1 shows the 3D CAD model framework based on the optimal matching algorithm.

Gaussian curvature and mean curvature (hereinafter referred to as curvature) describe the local geometric properties of the surface and are important differential geometric characteristics of the surface. As a piecewise continuous linear model, the triangular mesh model has no continuous curvature. For mesh surfaces, it is usually necessary to estimate the curvature of each vertex. The curvature of the points inside the mesh can be obtained by interpolation.

$$A(x1, x2, x3) = \frac{|x2 - x1| * |x3 - x1|}{2} \tag{1}$$

According to the characteristics of free-form surfaces being smooth and continuous, the distribution of curvature should not produce abrupt changes. However, because the model expressed in STL format usually has certain noise, noise will cause the generation of abnormal values of curvature; in order to prevent these abnormal values from affecting the effect of subsequent algorithms, it is necessary to filter the curvature and correct the abnormal values of curvature. The following describes the filtering and correction process with the modified Gaussian curvature as an example, and the same method is used for the average curvature.

$$S=\sum_{i=1}^{N}(A*x(i))=A*(x1+x2+...+x(N)) \tag{2}$$

After calculating and correcting the curvature of each vertex, the Gaussian curvature of each triangle on the surface must be described quantitatively. The free-form surface expressed in STL format essentially uses a set of interconnected triangles to approximate an accurate surface. The triangles are flat and have a curvature of 0, but each triangle approximates a small piece of surface, and the curvature of the surface is not zero. Therefore, a curvature value can be assigned to each triangle to approximate the curvature of the corresponding surface.

$$\frac{(1-r1)a}{p}+\frac{(1-r2)b}{p}+\sqrt{r1*r2}=1 \tag{3}$$

Take out the triangles in turn, and judge whether the triangle has been used according to the used flag. If it is not used, you can define the triangle as a meta area (the smallest area), and define the type of the meta area according to the triangle type, and record the meta area at the same time in the boundary.

$$\begin{cases} u(x,y)=1-\dfrac{|x-y|}{\max(x,y)} \\ v(x,y)=1-\dfrac{|x-y|}{\max(x+1,y+1)} \end{cases} \tag{4}$$

Find the triangles connected to the 3 sides of the meta-area, and determine the type of these triangles. If they are the same as the meta-area, add the triangle to the meta-area and update the boundary of the area; if the type of the connected triangle is different from the meta-area, then the corresponding edge of the element region is defined as the final boundary of the region, and the final boundary does not grow any more. Repeat the above growth process until all the boundaries become the final boundaries, at which time the growth of the region ends.

$$d=\sqrt{(x1-x2)^2+(y1-y2)^2+(z1-z2)^2} \tag{5}$$

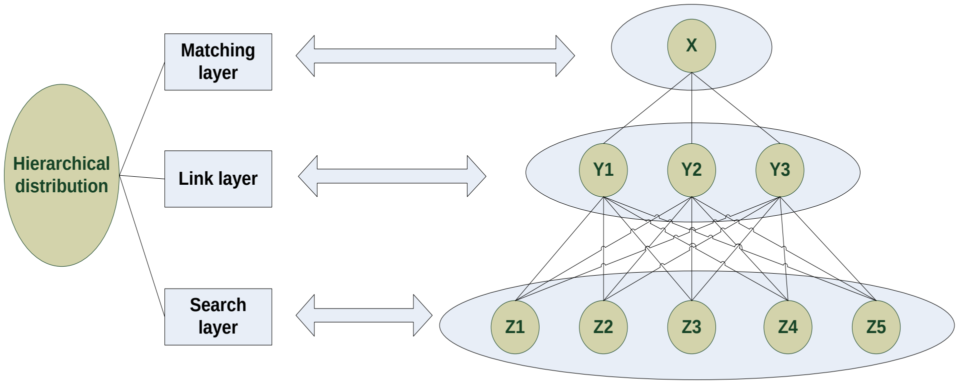

Figure 1: Three-dimensional CAD model framework based on optimal matching algorithm.

Through the segmentation of free-form surfaces and the extraction of shape descriptors, a free-form surface is described as a group of regions with relatively simple shapes. The similarity between two free-form surfaces can be measured by comparing the similarity between two groups of local regions. The comparison of the two local areas of the surface is the comparison of two 7-

dimensional generalized vectors. Since each dimension of a vector represents a different meaning, when implementing generalized vector comparison, each dimension of the vector is processed separately by rules.

3.2 Optimal Matching Algorithm

The similarity between the curved surfaces should be determined by the optimal scheme of the overall matching between the areas of the two curved surfaces, rather than the optimal press-fitting between the individual areas, in order to obtain the 2 free-form surfaces for comparison. The overall optimal matching between regions here is to find the most human match of the similarity matrix.

$$dist(x,y)=\begin{cases}\sum_{i=1}^{n}|x(i)-y(i)|^2, x>y \\ \max|\mathrm{x(i)\text{-}y(i)}|,\mathrm{x<y}\end{cases} \tag{6}$$

The result of the calculation is to select the corresponding region pair from the similar caries matrix as the optimal matching. The similarity value of the two surfaces is the sum of the column-similarity values of these optimal matching K domains. The overall similarity depends on the matching degree between the source model face and the target model face.

$$z(i,j)=\begin{bmatrix}\frac{x(1)-x(1)}{x(1)+x(1)} & 0 & 0 \\ 0 & \ldots & 0 \\ 0 & 0 & \frac{x(i)-x(j)}{x(i)+x(j)}\end{bmatrix} \tag{7}$$

If the matching degree between all corresponding faces is higher, then the greater the overall similarity is. The optimal face matching sequence M is obtained by accumulating the shape similarity between the source model surface and the target model surface.

The local structure retrieval problem of the 3D CAD model is transformed into a graph matching problem of the attribute adjacency graph of the local structure to be retrieved and the property adjacency graph of the retrieved CAD model to solve, and then the graph matching problem is transformed into a constraint optimization problem to solve based on two cases. There are many ways to solve iterations, such as mean field theory, relaxation labels, genetic algorithms, and so on. Considering the speed and performance of the iteration, this paper adopts the mean field theory to carry out iterative calculations. The mean field theory is a method of replacing the summation of individual effects with the average action effect. This method can simplify the study of complex problems and put a high-level. The multi-dimensional difficult-to-solve problem is transformed into a low-dimensional problem. Given the initial mapping matrix M, we use the mean field theory to iterate to obtain the optimal solution that finally converges to the objective function. Figure 2 shows the specific optimal match in the model attribute distinction of graph structure. In order to improve the matching efficiency, when designing the algorithm, the vertex mapping matrix has been improved. In the calculation of the vertex attributes of the attribute adjacency graph, more surface attributes are introduced. The main attributes are the geometry type of the surface, the direction of the surface, and the convexity of the surface. As long as one main attribute of the surface in the two CAD models is different, it is impossible to match the two vertices of the adjacent graph.

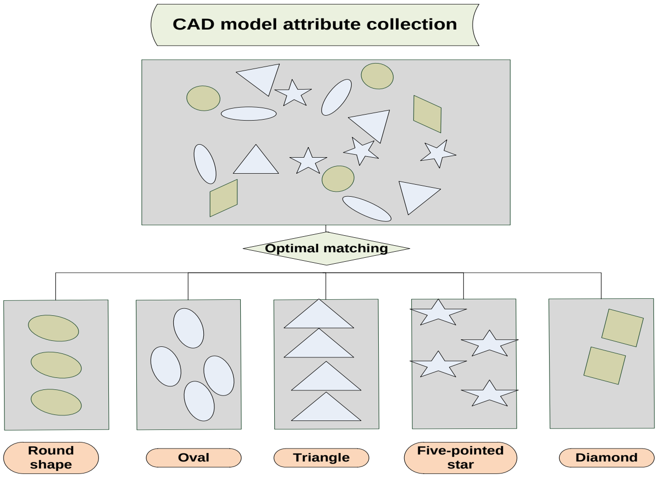

Figure 2: Model attribute differentiation for optimal matching of a specific graph structure.

4 APPLICATION AND ANALYSIS OF 3D CAD MODEL BASED ON OPTIMAL MATCHING ALGORITHM

4.1 3D CAD Model Simulation

In order to verify and evaluate the effectiveness and performance of the algorithm, Microsoft Visual Studio is used as the integrated development environment, and Open CASCADE91 is used as the geometry platform. A 3D CAD model library has been constructed. Among them, the 3D CAD model library contains 200 typical CAD models such as discs, shafts, and cabinets. The complexity of the models varies from a few to a few hundred. Figure 3 shows the histogram of matching values of different topological classes of the 3D CAD model. The topological feature between regions refers to the connection relationship between the regions.

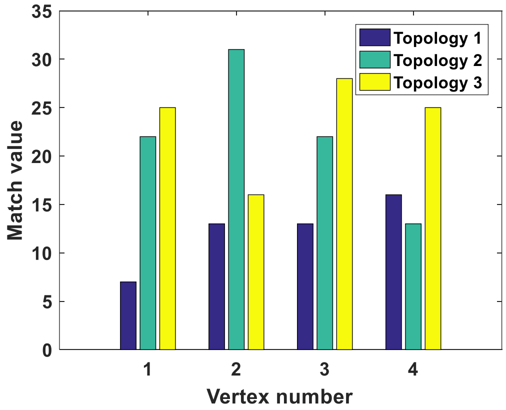

Figure 3: Histograms of matching values of different topology classes of 3D CAD models.

Since the object of this paper is a triangular mesh model, different resolutions will have a certain degree of influence on the surface segmentation. In order to make the quantity describing the topological characteristics between the surface regions have a certain robustness to the surface resolution, this paper uses the relative lengths of the common edges of various types of regions and the regions to describe the topological connection characteristics of the regions.

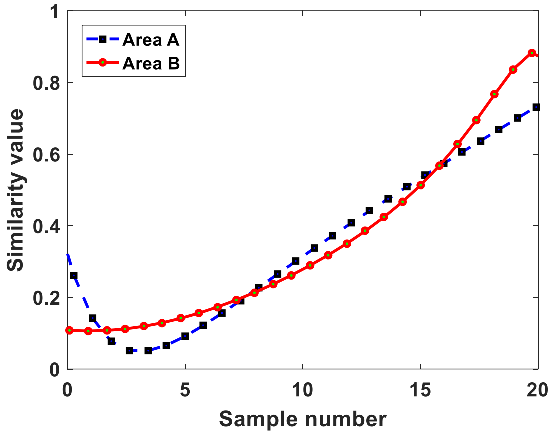

Figure 4: The similarity curve of different adjacent regions of the CAD mesh model.

As shown in the text, areas are adjacent, the types of these 5 areas are marked by capital numbers in the figure. Figure 4 shows the similarity curves of different adjacent regions of the CAD mesh model. The topological connection characteristics of the area are expressed by 4 numbers less than 1, and the essence is to express the area R with these 4 numbers. The distribution of various surrounding areas and the closeness is connected with the R area. When measuring the shape similarity between two models, other strategies are generally used to find the optimal surface matching sequence of the two models. Figure 5 shows the comparison of search error rates before and after recognition by the optimal matching algorithm. In order to measure the effectiveness of the method proposed in this paper, two sets of comparative experiments were carried out.

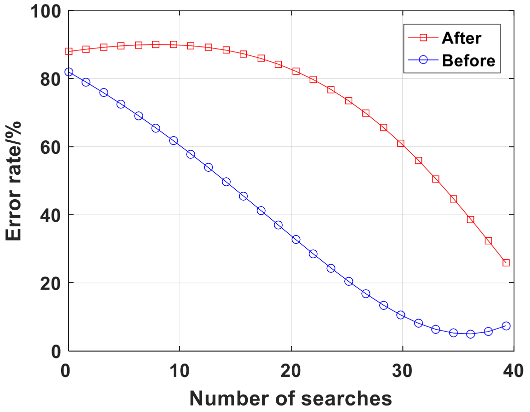

Figure 5: Comparison of search error rates before and after recognition by the optimal matching algorithm.

In experiment, we use the formula to calculate the similarity between the source model surface and the target model surface, construct the surface similarity matrix between the source model and the target model; we use the common algorithm to search the surface similarity matrix to find the source model and the target optimal face matching sequence between models, and at the same time, formulas are used to calculate the similarity between the source model and the target model. In experiment, the method proposed in this paper is used to calculate the similarity between the source model and the target model.

4.2 Case Application and Analysis

In order to fully verify the model similarity evaluation performance of the algorithm in the article, 20 classic reuse model structures are selected, with 70% similarity as the evaluation threshold, the human general model library is searched for models with classic structures, and each classic structure is performed 50 times. We repeat the experiment, take the average, and then average the experimental results of the 20 classic structures to get the average recall rate and average precision rate of each algorithm. The accuracy value distribution of the obtained CAD model joint surface is shown in Figure 6. The ideal result of the curve is a constant straight line with precision equal to. For the experimental results, the closer the curve position is to the ideal curve, the better the similarity evaluation performance of the algorithm is and the higher the retrieval accuracy of the model is.

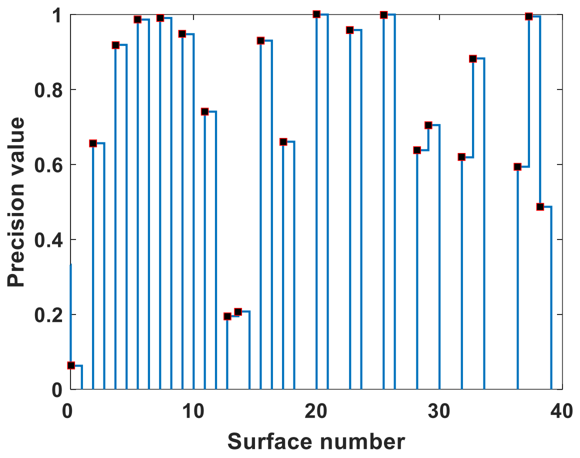

Figure 6: The accuracy value distribution of the CAD model joint surface.

In experiment, the common algorithm is used to search the face similarity matrix F. To find the optimal surface matching sequence between the source model and the target model, the time complexity is 0. The structural similarity between models is affected by the neighborhood structure between the model components. If the neighborhood structure between the model faces is similar, then this group of corresponding faces is similar. It can be considered that the two faces are structurally similar. By accumulating the similarity of the neighborhood structure between the model faces, it can be judged out of model with the structural similarity. Although the time complexity of experiment 2 is higher than that of experiment 1, the effect of experiment 2 is better than that of experiment 1 in terms of measuring the similarity between models. Figure 7 shows the dependence curve of the algorithm stability with the model vertex. The experimental results show that when the common algorithm is used to calculate the model similarity, the 10 models in the experiment cannot be distinguished, and the method proposed in this article can effectively measure the difference of these models; when calculating the similarity of the key models, this article shows the calculation result of the proposed method is higher than that of the ordinary algorithm.

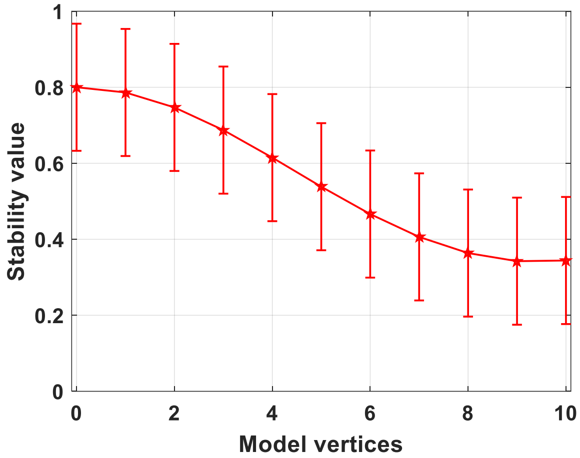

Figure 7: Algorithm stability vs. model vertex dependence curve.

Some three-dimensional CAD models were randomly selected. The number of faces of the models was 17, 25, 31, 42, 46, 33, 41, 25, 17, 13 and 45. They were placed in the model library to search and input the model itself. Figure 8 shows the optimal matching algorithm test deviation ladder diagram of the CAD model. Due to the vertex subdivision method, the model search with a large number of faces has certain advantages in speed, but the stability is lower, and vertex subdivision search speed of poorly performing models is significantly reduced.

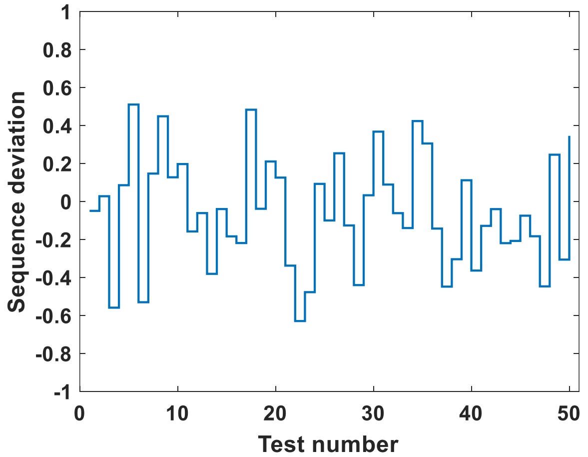

Figure 8: The optimal matching algorithm test deviation ladder diagram of the CAD model.

Compared with the algorithm-based similarity evaluation method, the algorithm in this paper searches for the optimal matching sequence of the two models, which was searched models with multi-level similarities in the model library. The search results are more similar in perception, while the algorithm-based model was used. In the experimental results of the similarity evaluation algorithm, the model is not the expected reuse model. Compared with the model similarity evaluation algorithm based on the network ontology language, the number of models searched by the algorithm in the article is more, and some of them have high complexity. Models are also searched, which is convenient for designers to better reuse and in-depth mining of models.

5 CONCLUSION

This paper systematically studies the theoretical basis of the hierarchical and multi-resolution representation of CAD 3D models, and proposes a reuse-oriented hierarchical structure, similarity comparison algorithms and application schemes. For the solid model, geometric inference is performed by extracting the topological relationship of adjacent surfaces, and a hierarchical structure expressing the feature inclusion relationship is established; for the free-form surface with uniform curvature, the surface is divided into levels, and the qualitative and qualitative and the description of the adjacent relationship are defined. Quantitative indicators are used to describe the degree of curvature from coarse to fine morphology; for composite surface models with sudden curvatures, a hierarchical structure is established by describing the morphological relationship of features, and an assembly reuse strategy based on functional equivalence for geometric replacement is proposed. Using the optimal matching algorithm of the bipartite graph in graph theory, the optimal matching of the two CAD model surfaces is achieved, and then through a series of processing, the similarity value of the two models is obtained to evaluate the similarity of the two models. When evaluating the similarity between the two sides, a variety of influencing factors are comprehensively considered. Because the algorithm is designed based on the geometric attributes of the model's constituent faces and the topological connection relationship between the faces, and the relative area is introduced for calculation, the evaluation of the similarity of the two models has nothing to do with the model size and placement position, and the algorithm has rotation invariance. After experimental testing, the effect is satisfactory.

Zhiqiang Zhang, https://orcid.org/0000-0002-6670-2297
Hongying Xue, https://orcid.org/0000-0001-8456-1419
Yifan Zhang, https://orcid.org/0000-0002-3174-9212

REFERENCES

[1] Deng, Y.; Han, S.-Y.; Li, J.: The design of tourism product CAD three-dimensional modeling system using VR technology, Plos one, 15(12), 2020, 44-45. https://doi.org/10.1371/journal.pone.0244205
[2] Lin, X.; Zhu, K.; Wang, Q.-G.: Three-dimensional CAD model matching with anisotropic diffusion maps, IEEE Transactions on Industrial Informatics, 14(1), 2019, 265-274. https://doi.org/10.1109/TII.2017.2696042
[3] Jayanti, S.; Kalyanaraman, Y.; Iyer, N.: Developing an engineering shape benchmark for CAD models, Computer-Aided Design, 38(9), 2020, 939-953. https://doi.org/10.1016/j.cad.2006.06.007
[4] Zhao, X.; Liu, X.; Zhang, K.: A structure-based 3d CAD model similarity assessment approach, Software and Computer Applications, 2, 2017, 281-284. https://doi.org/10.1145/3056662.3056721
[5] Ambrosini, R.-D.; Wang, P.; O'Dell, W.-G.: Computer-aided detection of metastatic brain tumors using automated three-dimensional template matching, Journal of Magnetic Resonance Imaging, 31(1), 2020, 85-93. https://doi.org/10.1002/jmri.22009
[6] Mian, A.-S.; Bennamoun, M.; Owens, R.: Three-dimensional model-based object recognition and segmentation in cluttered scenes, IEEE transactions on pattern analysis and machine intelligence, 28(10), 2020, 1584-1601. https://doi.org/10.1109/TPAMI.2006.213
[7] Bosché, F.: Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction, Advanced engineering informatics, 24(1), 2020, 107-118. https://doi.org/10.1016/j.aei.2009.08.006

[8] Son, H.; Kim, C.; Kim, C.: 3D reconstruction of as-built industrial instrumentation models from laser-scan data and a 3D CAD database based on prior knowledge, Automation in Construction, 49, 2019, 193-200. https://doi.org/10.1016/j.autcon.2014.08.007

[9] Cai, J.; Martorella, M.; Liu, Q.: Automatic Target Recognition Based on Alignments of Three-Dimensional Interferometric ISAR Images and CAD Models, IEEE Transactions on Aerospace and Electronic Systems, 56(6), 2020, 4872-4888. https://doi.org/10.1109/TAES.2020.3005654

[10] Bosche, F.; Haas, C.-T.: Automated retrieval of 3D CAD model objects in construction range images, Automation in Construction, 17(4), 2018, 499-512. https://doi.org/10.1016/j.autcon.2007.09.001