Study of Semantic Similarity Evaluation Methods for Combined Assembly Models

Xufan Wang, Hu Qiao^{*}, Sibo Hu, Liang You, Zhijie Wang

School of Mechatronic Engineering, Xi'an Technological University, Xi'an, Shaanxi, 710021, China *Corresponding author

Keywords: Combined assembly model retrieval, Assembly features, Similarity assessment

Abstract: In order to better realize the effective reuse of manufacturing information of 3D CAD models for manufacturing domain, a semantic-based research of combined assembly model retrieval method is proposed. Firstly, the labelled CAD model is transformed into a structured CAD model with the assembly features as the semantic object; then the attribute information for evaluating the similarity between two manufacturing features is extracted to construct a multi-attribute fused manufacturing feature similarity weighted evaluation model, which is used to construct a complete bipartite graph with manufacturing features as nodes; the optimal matching algorithm is used to calculate the optimal Finally, the optimal matching algorithm is used to calculate the similarity of the assembly model, which is used as the basis for evaluating the similarity of the model. The results show that the method can better achieve semantic-based combined assembly model retrieval, and can effectively support the reuse of manufacturing information of 3D CAD models in manufacturing oriented fields.

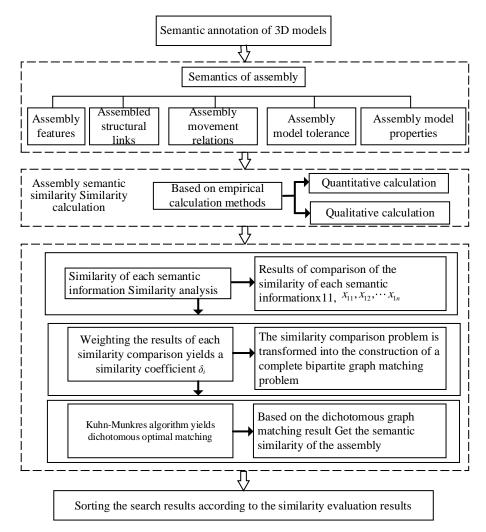
1. Introduction

With the development of computer technology and the use of advanced manufacturing techniques in the production of life, the information available in manufacturing companies' models is growing geometrically. This has led to new questions about how users can efficiently find the part models they need in the mass of information optimized combinations, and just 20% as a result of design innovation [1]. Therefore, most companies build CAD model libraries with the aim of having easy access to the models they already have during product development in order to identify the models that meet their specific functions. However, due to the sheer number of models, it is not advisable to view these models by manual browsing, so a reasonable and effective retrieval method is required [2].

In recent years 3D model retrieval research is moving towards fused semantic retrieval to better support the development of the model reuse domain. Sridhar et.al [3] used a semantic representation based on the machining feature association relationship of a part to represent this association relationship as manufacturing feature semantics and organize it to form a Model Dependency Graph (MDG) to further compare the model similarity and extract the required model. Tsaip [4] used a fuzzy set theory to evaluate the semantic similarity of 3D models. The input of engineering

semantic information, processing requirements, quality standards and material properties, etc., is performed through interaction. Cardone [5] proposes a feature semantic based 3D model retrieval method, where the feature semantics of the 3D model: machining process, material type, dimensions and amount being machined, are represented as spatial vectors. Elinson [6] uses the machining feature semantics, such as shape, volume and number of machined surfaces, as attributes of graph vertices and the interaction between feature semantics as edges connecting the vertices, starting from the machining design intent of the 3D model. Lupinetti [7] clusters the 3D assembly models in the model library according to the semantics of mechanical motion characteristics, motion subsets and connection types of 3D assembly models, and then gradually layers the model parts by combining the hierarchical semantics of 3D assembly models. It makes it possible for the user to gradually narrow down the search scope. Qiao et al. [8] In response to the problems of semantic mismatch, poor accuracy and low efficiency in existing 3D assembly model retrieval methods, a 3D assembly model retrieval method based on 3D assembly information is proposed. Firstly, for the retrieval of assembly information, the assembly information is represented by a symbolic code, and the 3D model that matches the assembly design intention is found by retrieving the code. Zhao et al. [9] In order to remedy the shortcomings of traditional retrieval models based on keyword matching, a semantic information retrieval model of domain ontology is explored. The ontology of quadruples is described, a resource mapping scheme is given, an ontology concept expansion strategy is developed, an ontology concept similarity calculation algorithm is described, and an experimental comparison analysis is carried out. Chen et al. [10] The results show that the ontology semantic retrieval model has a higher accuracy and completeness rate than the traditional retrieval model, and has certain theoretical and practical value. Wang et al. [11] The semantic tree is constructed to establish the hierarchical relationship between concepts, and the expansion of keyword semantics and the corresponding retrieval method are proposed. The similarity between keyword semantics and model semantics is calculated based on Word Net, and the models under the nodes with strong semantic similarity are returned to reduce the possibility of empty retrieval results. Ma et al. [12] The semantic tree of design intent of each model is established according to the modeling information, and the ontology semantic model tree and search index based on the 3D model semantic tree database are established, and then the semantic similarity of the corresponding model is calculated by comparing the similarity between the target search term set and the semantic tree nodes, which solves the semantic gap problem of the content-based retrieval method; By matching the target retrieval terms with semantic annotated terms for similarity calculation, the retrieval search completion rate is improved. Chen et al. [13] An information retrieval model for the field of mechanical design is developed using semantic web technology and intelligent subject technology, and the architecture of the model and its key technologies are analyzed. The model enables good reusability of information, and can obtain accurate and comprehensive information from a large number of web resources, and provide information with different levels of detail and personalized services for different users, in preparation for the future implementation of an information integration system based on the semantic web. Guo et al. [14] address the key issues of establishing a knowledge database based on semantic retrieval methods, knowledge extraction and semantic metrics, and focus on semantic-based 3D model retrieval on the basis of establishing a 3D model ontology knowledge network. Metzler et al. [15] combine the language modeling and inference network approaches into a single framework. The resulting model allows structured queries to be evaluated using language modeling estimates. Miriam et al. [16] investi-gates the definition of an ontology-based Information Retrieval model, oriented to the exploitation of domain Knowledge Bases to support semantic search capabilities in large document. Liang et al. [17] proposed a representation of assembly structural data including topological structure, assembly semantics, and geometrical information. Then enrich assembly design ontology for knowledge

captured and shared in Web Ontology Language 2 Description Logicand Semantic Web Rule Language. And next, define the matching strategies and similarity assessment for two matched models. Tang et al. [18] response to the complex structure of the assembly model representation, which is not directly used for similarity evaluation, the assembly model is simplified by representing the assembly composition tree and assembly constraint diagram, and then the similarity evaluation of the assembly model is achieved based on subtree matching and diagram matching methods. In order to meet the needs of collaborative assembly planning, Wang et al. [19] used assembly nodes to describe product assembly information and established an assembly linkage diagram model of the product. Kim et al. [20] have developed an ASD (assembly design) prototype system for capturing assembly intent and connection intent using the Web Ontology Language (OWL) and the semantic Web rule language (SWRL) for detailed classification of assembly semantics.



2. Semantic Information Based 3D Model Similarity Evaluation Method

Figure 1: Similarity evaluation method for 3D assembly models based on assembly semantics

This paper first classifies the semantics of the assembly models, and then, based on the empirical calculation method and the characteristics of the assembly semantics, the assembly semantics are calculated quantitatively and qualitatively respectively to derive the calculation value of the similarity of the assembly semantics. Finally, based on the bipartite graph matching method and

Kuhn-Munkres algorithm, the similarity evaluation of the model assembly semantics in the 3D model library is carried out to extract the 3D models that match the input model assembly semantics. The technical route of this paper is shown in Figure 1.

2.1. Semantic Classification of Combinatorial Assembly Models

The assembly semantics of the model mainly includes the assembly semantics and the attribute semantics of the parts that make up the assembly model. The assembly semantics is mainly expressed in some assembly characteristics of the part model during the assembly process, while the attribute semantics of the parts lies in the material properties of the parts themselves, the part type, the part name, etc. There are several types of assembly semantics between the combined assembly models, each semantic type includes different semantic words, based on the semantic combined assembly model retrieval, the corresponding semantic words are compared for similarity, in order to achieve efficient retrieval efficiency, the semantic classification of the assembly semantics and the attributes of the parts themselves is needed.

2.1.1. Classification of Assembly Model Assembly Semantics

(1) Classification of assembly features

Assembly model manufacturing features mainly include model holes, cavities, slots, steps, etc. In the process design process, tolerances and surface roughness have a close relationship with the assembly between the part model. The specific assembly features are classified as shown in Figure 2. The correlation between assembly features plays an important role in the evaluation of the similarity of the assembly model [21], and models with the same combination of design features are more similar in terms of process reuse, provided that the geometry is similar [22].

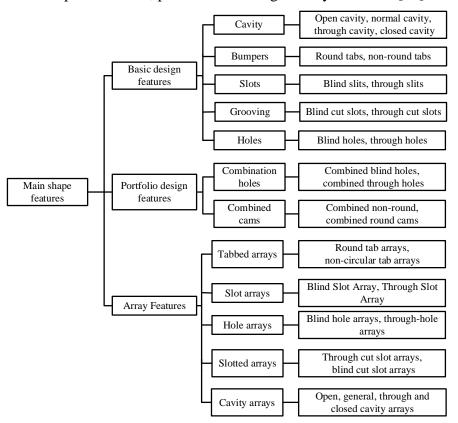


Figure 2: Classification of model assembly features

(2) Assembly structure association semantics

For a combined assembly model, which is generally assembled from a number of part models, the parts belonging to an assembly have a close connection with each other and are able to work together to achieve certain functions of the designed product or assembly. Therefore, before the CAD assembly model can be divided into modules, it is necessary to analyse and evaluate the degree of association between the assembled parts. The connection and fit of the assembled parts determines the degree of association between the assembly structure of the parts. According to the connection type and attribute information in the mechanical assembly model, the connection between the parts is used to define the semantics of the association of the part structure, as shown in Figure 3.

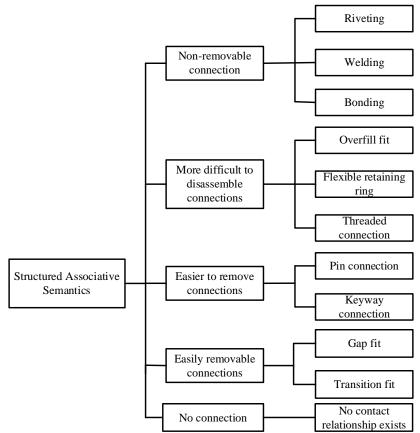


Figure 3: Semantic classification of assembly associations

(4) Assembly model tolerance/surface roughness

In the process of process design, the tolerance or surface roughness of the model is closely related to the choice of the assembly feature processing method. For two local structures with exactly similar geometry, if the dimensional accuracy differs significantly, the corresponding dimensional processing process will produce a large difference, so the accuracy class assembly semantics plays a key role in the evaluation of model similarity.

2.1.2. Semantic Classification of Assembly Model Part Attributes

The material properties of the 3D assembly models vary and have different strain resistance. Depending on the product being applied to different working conditions, the material types are, steel: cast steel, alloy steel and carbon steel, cast iron: ductile iron, malleable iron and grey cast iron, bearing alloys: lead-based bearing alloys and tin-based bearing alloys, copper alloys: deformed copper alloys and cast copper alloys. As shown in Figure 4.

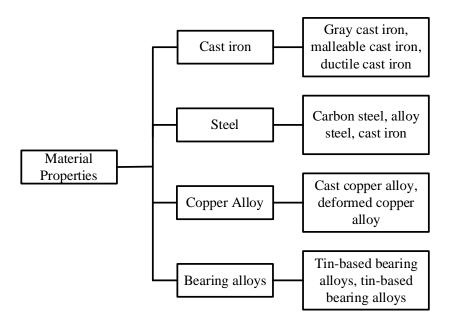


Figure 4: Semantic classification of part attributes

2.2. Similarity Calculation for Combined Assembly Models

According to the previous description, this paper designs an assembly feature descriptor that can describe different levels of information. The feature is defined as a 5-dimensional vector, and the similarity comparison between two models is a comparison of two 5-dimensional vectors. In order to achieve the similarity evaluation between models, it is first necessary to quantify each component of the feature descriptor, and then to fuse multiple feature attributes to construct a similarity evaluation of assembly features. As each component element represents a different meaning, each component element is evaluated for feature similarity according to the following rules.

(1) Comparison of similarity of assembly characteristics

If the assembly features of the two models are of different types, the similarity of the two features is considered to be 0, the features are not similar. If the features are of the same type, the similarity of the types is 1. Let the two compared features be T_1 and T_2 , and the similarity of the types is denoted by S_T .

$$S_{F} = \begin{cases} 0, & F_{1} \neq F_{2} \\ 1, & F_{1} \neq F_{2} \end{cases}$$
(1)

(2) Similarity calculation for assembly connection types

Connection types provide a clear representation of the fit relationships between assembly models and have an important influence in assembly semantic retrieval. Structural association semantic types can be divided into: non-disassembled connections, harder-to-disassemble connections, easier-to-disassemble connections, easily disassembled connections and no connections called parent categories. The more difficult to disassemble connections can be divided into: interference fits, elastic retainers, threaded connections, easier to disassemble connections can be divided into: pin connections, keyway connections, easy to disassemble connections can be divided into: gap fits, transition fits etc. They are referred to as sub-categories. The semantic search is compared according to the harder to disassemble connections, easier to disassemble connections and easily disassembled connections. For the subclasses in these three parent categories, the semantic similarity comparison method with reference to literature [23,24] is used to develop similarity evaluation values. This is shown in Figure 5.

	1.1	1.2	1.3	2.1	2.2	3.1	3.2
1.1	1	0.9	0.8				
1.2		1	0.7				
1.3			1				
2.1				1	0.5		
2.2					1		
3.1						1	0.5
3.2							1

Figure 5: Similarity calculation for assembly connection types

The values in the diagrams 1.1 to 1.3 indicate interference fit, elastic retaining ring and threaded connection, 2.1 to 2.2 indicate pin connection and keyway connection, 3.1 to 3.2 indicate gap fit and transition fit respectively The values in the diagrams indicate the value of the similarity between the two two subclasses of the semantics of the connection relationship. if the parent classes are different, the similarity is 0.

(3) Calculation of similarity of motion relations for combined assembly models

When performing model retrieval, the user input model, and the retrieved output model need to satisfy the above motion relationship if there is relative motion, otherwise they cannot be connected. Assuming that the motion types of input model *i* and model *j* in the model library are T_i and T_j respectively, the similarity of the motion types of model *i* and model *j* is:

$$S_T = \begin{cases} 1 & t(T_i) = t(T_j) \\ 0 & t(T_i) \neq t(T_j) \end{cases}$$
(2)

Where,

Function t - denotes the name of the motion type.

If input model *i* has the same type of motion as model *j* in the model library, then $S_T = 1$. If input model *i* has a different type of motion than model *j* in the model library, then $S_T = 0$.

(4) Similarity calculation for model tolerances or surface roughness

The similarity of tolerance or surface roughness depends on the grade of tolerance or surface roughness. Let the highest level of tolerance in the basic attributes of the two compared manufacturing features be IT_1 , IT_2 , and the minimum value of surface roughness be R_1 and R_2 , respectively.

$$S_{R} = \frac{1}{2} \left(1 - \frac{|IT_{1} - IT_{2}|}{\max(IT_{1}, IT_{2})} \right) + \frac{1}{2} \left(1 - \frac{|R_{1} - R_{2}|}{\max(R_{1}, R_{2})} \right)$$
(3)

(5) Calculation of similarity of material types

Different parts play different roles in the assembly model and require different material properties. The main material types considered here are, steel: cast steel, alloy steel and carbon steel, cast iron: ductile iron, malleable iron and grey cast iron, bearing alloys: lead-based bearing alloys and tin-based bearing alloys, copper alloys: deformed copper alloys and cast copper alloys. The material types of the main functional parts in the model library are mainly considered for model retrieval, other sub-assemblies are not taken into account.

The material type similarity between the combined assembly models is compared by assuming

that the material information of the input model and the model in the model library is M_i and M_j , and the material similarity between the input model and the model in the model library is S_M , then we have

$$S_{M} = \begin{cases} 0, & m(M_{i}) = m(M_{j}) \\ 1, & m(M_{i}) \neq m(M_{j}) \end{cases}$$
(4)

2.3. Evaluation of Semantic Similarity of Assembly of Combinatorial Assembly Models

The semantic similarity evaluation of the combined assembly semantics is calculated using the bipartite graph method, where the input model assembly semantics and the assembly semantics of the models in the model library and the assembly semantics of the models in the model library are matched using the bipartite graph, and then, the optimal match between them is calculated using the Kuhn-Munkres algorithm to derive the semantic similarity value of the combined assembly model assembly.

The weighting coefficients for assembly features, connection types, kinematic relationships, roughness and material types are w_F , w_L , w_T , w_R and w_M and $w_F + w_L + w_T + w_R + w_M = 1$ respectively, with values set at 0.4, 0.3, 0.1, 0.1 and 0.1 respectively, depending on the reference weighting of these assembly semantics when composing the assembly.

2.3.1. Section Titles

Using the results of each type of similarity comparison for each of the five assembly semantics described above as an attribute node, the semantic similarity comparison between two 3D models is transformed into a full bipartite graph optimal matching process for both sets of nodes. A complete bipartite graph is defined as a bipartite graph in which each node in the set of two nodes can be connected to all nodes in the other set.

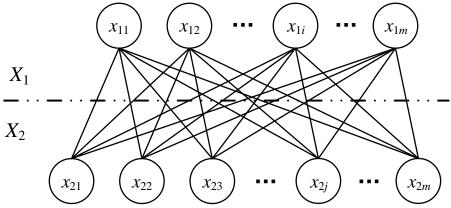


Figure 6: Illustration of a complete dichotomous diagram construction

The assembly semantics of a 3D model can then be represented by a set of attribute nodes, the set of assembly semantics nodes for the 3D model X_1 is $X_1 = \{x_{11}, x_{12}, \dots, x_{1m}\}$ and the set of assembly semantics nodes for the 3D model X_2 is $X_2 = \{x_{21}, x_{22}, \dots, x_{2m}\}$. In this way the comparison of the assembly semantic similarity between models is transformed into a matching process of a full bipartite graph of two sets of nodes. A schematic diagram of complete bipartite graph construction is shown in Figure 6, such that the complete bipartite graph is G = (V, E), where V is the

concatenation of two sets of assembly semantic nodes, i.e. $V = X_1 \cup X_2$. *E* is the edge joining the two sets of nodes. Moreover, the nodes in the same node set are not interconnected, and the nodes between the node set X_1 and X_2 are connected by an edge.

Considering the weight coefficients of each semantics, the similarity of each semantics is calculated to obtain the similarity value δ_{ij} . The similarity value is used as the weight of the bipartite graph, so the weight matrix of the bipartite graph is the similarity matrix of the assembly semantics between the three-dimensional models, as shown in Equation 4. When the number of nodes in the two groups is different, the similarity matrix becomes a square matrix with 0 complements. Where x_{1m} denotes the assembly semantics of the input model and x_{2m} denotes the assembly semantics of the output model.

2.3.2. Kuhn-Munkres Algorithms

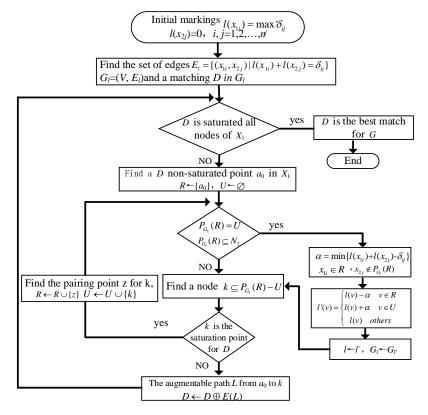


Figure 7: Kuhn-Munkres algorithm

In order to obtain the optimal solution for the semantic similarity evaluation of the two compared 3D models, i.e. to seek the optimal matching of the assignment bipartite graph, the optimal

matching problem of the bipartite graph is transformed into a search for incremental paths based on the Kuhn-Munkres algorithm [25], based on the relevant theory in graph theory, and the bipartite graph reaches the optimal matching when all the nodes are matched. The similarity matrix of the assembly semantics between the 3D models is calculated to find the optimal matching solution.

The flow chart of the Kuhn-Munkres algorithm is shown in Figure 7.

Based on the optimal matching D calculated by KM, the semantic similarity S of the assembly of the two 3D models is calculated using the optimal matching D as follows.

$$S = \frac{\sum_{j=1}^{n} \left(\frac{\tau_{\nu(j)} + \tau_j}{2} \cdot \delta_{\nu(j)j} \right)}{\max(\sum \tau_{\nu(j)}, \sum \tau_j)}$$
(6)

Where V_j is the number of rows that optimally match column *j* in equation 4, $\delta_{V(j)j}$ is the semantic similarity coefficient between the V(j)-th semantics of the 3D model X_1 and the *j*-th semantics of the 3D model X_2 , $\tau_{V(j)}$ is the number of semantics contained in the 3D model X_1 , and τ_j is the number of semantics contained in the 3D model X_2 .

3. Conclusions

This paper completes the similarity evaluation of 3D assembly models based on assembly semantics, carries out the classification and representation of assembly semantics, calculation of assembly semantic similarity and similarity evaluation methods. By constructing a similarity evaluation system for 3D assembly models based on assembly semantics, the assembly design intent of 3D models is better summarised, a calculation method is provided for the comparison of assembly semantic similarity, and the comprehensive similarity of assembly semantics between 3D assembly models is evaluated using dichotomous diagrams and the KM algorithm. Through the semantic retrieval method in this chapter, 3D models that are semantically compatible with the input model and can be assembled can be initially extracted from the model library. The algorithm in this paper can achieve the evaluation of the similarity of 3D models with fused assembly semantics, which is of great scientific significance for the effective reuse of digital manufacturing results, rapid industrial design and the development of new generation CAPP systems.

Future research work includes: (i) optimizing the semantic retrieval mechanism for combined assembly models to improve the accuracy of model retrieval; and (ii) investigating the effect of geometry on model retrieval on the basis of semantic retrieval.

References

[1] Natraj Iyer, Subramaniam Jayanti, Kuiyang Lou, Yagnanarayanan Kalyanaraman, Karthik Ramani. (2004) Threedimensional shape searching: state-of-the-art review and future trends. Computer-Aided Design, 37 (5), 509-530.

[2] Zhansong Wang, Ling Tian, Wenrui Duan. (2014) Annotation and retrieval system of CAD models based on functional semantics. Chinese Journal of Mechanical Engineering, 27 (6), 1112-1124.

[3] Sridhar A, D Srinivasulu, D P Mohapatra. (2013) Model-based test-case generation for Simulink/Stateflow using dependency graph approach //Proceedings of Advance Computing Conference. Ghaziabad: IEEE Computer Society Press, 1414-1419.

[4] Tsai C-Y, Chang C A. (2013) A two-stage fuzzy approach to feature-based design retrieval. Computers in Industry, 56 (5): 493-505.

[5] Cardone A, Gupta S K, Deshmukh A, et al. (2006) Machining feature-based similarity assessment algorithms for prismatic machined parts. Computer-Aided Design, 38 (9): 954-972.

[6] linson A, Nau D, Regli WC. (2003) Feature-based similarity assessment method using convex decomposition //Proeeding of the ASME DETC 2003 Computers and Information in Engineering. Chicago: ACM Press, 1-13.

[7] Lupinetti K, Chiang L, Giannini F, et al. (2017) Regular patterns of repeated elements in CAD assembly model retrieval. Computer-Aided Design and Applications, 14 (4): 516-525.

[8] Qiao H, Qingyun W U, Jiang D U, et al. (2018)3D Assembly Model Retrieval Method Using Assembly Information. Journal of Huaqiao University (Natural Science).

[9] Yanfeng Zhao, Xiaohong Zhou. (2015) Research on Semantic Information Retrieval Model Based on Ontology. Software Engineer, 18 (08): 6-9.

[10] Xiang Chen, Shuming Gao, Song Guo, Jing Bai. (2012) A flexible assembly retrieval approach for model reuse. Computer-Aided Design, 44 (6).

[11] Honglin Wang. (2016) Research on 3D Model Retrieval Based on Semantic Tree. Shandong Normal University.

[12] Lihua Song. (2017) The Research and Development of CAD Model Annotation System Based on Design Intent. Dalian Polytechnic University.

[13] Xiaoyan Chen, Jinan Gu. (2007) Research on Semantic Web-based Information Retrieval Technology for the Mechanical domain. Manufacturing Automation (06): 6-8+59.

[14] Jing Guo. (2013) Some Researches on Semantic-based 3D Model Retrieval Techniques. Northwest University.

[15] Donald Metzler, W.Bruce Croft. (2004) Combining the language model and inference network approaches to retrieval. Information Processing and Management, 40 (5), 735-750.

[16] Miriam Fernandez, Ivan Cantador, Vanesa Lopez, David Vallet, Pablo Castells, Enrico Motta. (2011) Semantically enhanced Information Retrieval: An ontology-based approach. Journal of web semantics: Science, services and agents on the world wide web, 9 (4), 434-452.

[17] Yi Fang Liang, Yan Ru Zhong, Mei Fa Huang, Cong Wen Zeng. (2015) Ontology-Based Assembly Design Representation for Model Reuse. Applied Mechanics and Materials, 4058 (775-775), 484-490.

[18] Weihua Tang. (2010) A Dissertation Submitted to Zhejiang University for the Degree of Master of Engineering. Zhejiang University.

[19] Junfeng Wang, Shiqi Li, Jihong Liu. (2009) Information Model for Collaborative Assembly Planning. Computer Integrated Manufacturing Systems, 15 (04): 670-675.

[20] Kyoung-Yun Kim, David G. Manley, Hyungjeong Yang. (2006) Ontology-based assembly design and information sharing for collaborative product development. Computer-Aided Design, 38 (12), 1233-1250.

[21] Rui Haung, Shusheng Zhang, Xiaoliang Bai, Kaixing Bai, Ximou Zhang. (2013) Manufacturing Semantics Based 3D CAD Model Retrieval Method. Computer Integrated Manufacturing Systems, 19 (06): 1177-1185.

[22] Feiwei Qin, Shuming Gao, Xiaoling Yang, Jing Bai, Quhong Zhao. (2017) A sketch-based semantic retrieval approach for 3D CAD models. Applied Mathematics-A Journal of Chinese Universities, 32 (1), 27-52.

[23] Zarchi M S, A Monadjemi, K Jamshidi. (2014) A semantic model for general purpose content-based image retrieval systems. Computers & Electrical Engineering, 40 (7): 2062-2071.

[24] Lili Wan, Meiqin Liu. (2019) Feature Combination for 3D Model Retrieval Based on Evaluation Indicators. Journal of Computer-Aided Design & Computer Graphics, 21 (11): 1563-1568.

[25] Hongshen Wang, Shusheng Zhang, Xiaolianga Bai et al. (2011) 3D freeform surface similarity assessment algorithm based on region segmentation. Journal of Computer-Aided Des- gin & Computer Graphics, 23 (2): 305-313.