

A 3D Model Search Method Based on Combination of Design Knowledge and Model Features

Jihong Liu^{1,*}, Keyi Wang¹, Chao Fu¹, Yongzhu Hou¹

¹School of Mechanical Engineering and Automation, Beihang University, 100191 Beijing, China

Abstract. 3D models play an important role in modern product design, and there has been a large number of models stored in enterprise database. Reuse of existing models takes a large proportion of new product design. Model search technology helps find existing models which is to be reused. At present there are three model search methods including text-based, content-based and semantic-based. Embedding knowledge into models filled models with abundant semantic information, which supports model search indexing. Embedding knowledge into models filled models with abundant semantic information, which supports model search indexing. A 3D model search method based on combination of design knowledge and model features is proposed. The technical connotation of the proposed method is introduced and an example is given to illustrate the effectiveness of the proposed method.

1 Introduction

Today, we are currently living in the era of knowledge-driven economy. The fierce competition in the modern global market requires that manufacturing companies be able to enhance their competitiveness in terms of time (T), quality (Q), cost (C), innovation (I) and service (S) [1]. The enhancement of competitiveness needs to be achieved through optimization of multi-domain knowledge applications. People have realized the importance of knowledge in the design and manufacture of products, so the enterprises need to effectively manage the existing knowledge, while the knowledge is quickly and accurately provided to the designers to achieve a comprehensive product optimization design [2]. With the development of computer technology in industrial, designers can build a 3D model to carry out more intuitive product design [3]. At the same time, the product of the 3D model in product simulation analysis and digital manufacturing is also in a very important position [4]. It is precisely rich expression of product information that the 3D model have, "full 3D" concept was raised [5]. As it is a representative technology in the full three-dimensional process, Boeing has introduced Model-Based Definition (MBD) technology to integrate the information of the product into a unified 3D model and has been applied in the B787 to achieve disappeared drawings of the design and manufacture.

The purpose of MBD is to annotate 3D models with engineering information on traditional 2D drawings. Those kinds of annotated information describe what the product is and how it should be manufactured. However, currently the design basis behind supportive information

(design knowledge) has not been added into models. So integrating design process information into 3D models can assist designers to understand the product design and reuse them in new products. That is also an effective approach to preserve knowledge, which is precious for an enterprise, and make them more helpful.

The main purpose of the integration of the 3D model and the product design knowledge is to assist the design process through the reuse of the product model and the design process. The search is the most direct way to find the auxiliary resources. Whether the auxiliary design resources can be provided directly to the designer by analysing the search requirements is the main criterion for evaluating a method. At present, there are many researches on the search of 3D models. The existing model search technology is divided into three categories: text, content and semantics [6-9]. The model embedded with design knowledge [10, 11] makes the existing product model with richer semantic information, which also provides index support for model search. When the design knowledge is embedded in the 3D model, it provides the basis for the new search method.

2 The method associated with model and design knowledge

2.1 Classification and organization of knowledge

In order to give full play to the role of product model and engineering knowledge in product development, the literature in [12] describes classification and

* Corresponding author: ryukeiko@buaa.edu.cn

organization technology of the engineering knowledge, according to the structural characteristics of the design knowledge and its form of expression, The knowledge need to be marked to the product model and it should be classified and collated. At the same time, based on the knowledge classification method and knowledge level organization method of product structure and development process, the knowledge organization structure is achieved through XML, so as to realize the computer storage and expression of design knowledge. But it is also an important issue that how to use the marked knowledge to achieve engineering applications. For the three-dimensional model in which design knowledge has embedded, the most straightforward application is to retrieve the 3D model and the corresponding design knowledge related to the design through the 3D model of the embedded knowledge. This application pattern can be applied to multiple scenarios, such as collaborative design, design reuse, design learning and communication between design and manufacturing. It not only can improve the utilization of product design specifications, but also to promote the Information exchange between design process and manufacturing process. On this basis, this paper studies the 3D model search technology based on knowledge.

This paper associates explanatory object and the explanatory knowledge through the structure of a knowledge explanatory tree, as shown in Figure 1. Each explanatory object is explained by a knowledge explanatory tree. The knowledge explanatory tree has the following characteristics: (1) Each node is a knowledge packet, the knowledge package contains the corresponding knowledge item, each knowledge item includes multiple knowledge elements; (2) If a node knowledge package contains intermediate elements, its sub-knowledge package is used to explain its contents; (3) The leaf nodes in each tree shall not contain elements other than the design of the original information.

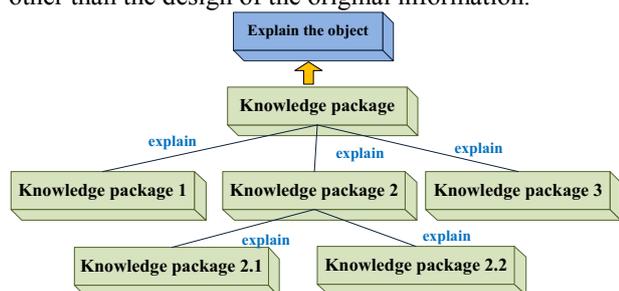


Figure 1. Knowledge explanatory tree

2.2 Knowledge embedded process

Based on the type of the design knowledge, the form of the knowledge representation, the embedding level of 3D model knowledge and organizational structure of 3D model design knowledge mentioned before, this section will elaborate the design knowledge embedding process of 3D model. The process will guide and regulate how the embedded knowledge is completed for an model which don't have embedded knowledge. The whole process is divided into five steps:

Step 1: Determine the model hierarchy. According to the specific model level that is required embedded knowledge, select the appropriate level. As mentioned above, here the model levels are divided into 5 levels, which are products, components (standard, non-standard), parts (standard, conventional, unconventional), features and surface.

Step 2: Select the explanatory object. Based on the level of the model determined in step 1, the explanatory object which is explained by the design knowledge are selected within the range of this hierarchy, such as the specific parts mentioned in the level of the parts, the specific size and tolerance mentioned in the level of the features and the specific surface in the level of the surface.

Step 3: Select the type of knowledge. After the specific explanatory object is determined, the corresponding knowledge package will be constructed by adding the knowledge item. The type of specific knowledge which is embedded is selected based on the knowledge types defined in the above. Each knowledge item belongs to a certain kind of knowledge.

Step 4: Select the form of knowledge representation. Each knowledge item is composed of one or more forms of knowledge, so it is necessary to select the form of knowledge before entering the specific design knowledge content. The forms of knowledge include text, formulas, charts and links.

Step 5: Input the content of knowledge. According to model level, explanatory object, knowledge type and knowledge form which are determined previously, input the corresponding and specific knowledge content, complete embedding of design knowledge.

A complete knowledge item can be formed from step 3 to Step 5. A complete knowledge packet can be formed from step 2 to Step 5, it. Through the iterations of the five steps, the embedding process of model knowledge can be achieved, The process will form multiple knowledge packets for different model levels and the explanations in which they are.

3 3D Model search method based on design intention and model features

In the 3D model search method, which combines the design intention and the model feature, the input content is divided into two parts, as shown in Figure 2. One part is the model which need to match, its essence is to describe the geometric requirements of the search through the model; the other part is the design intention description. Through the input of these two search requirements, the search results can be controlled by the two aspects, so that it can meet the more design requirements.

Figure 3 shows the basic flow of the search method. For the two input parts of the search needs are dealt respectively. The features sequence is extracted from the input model, and the natural language description of the design intention is used to construct the word sequence by using the word segmentation method.

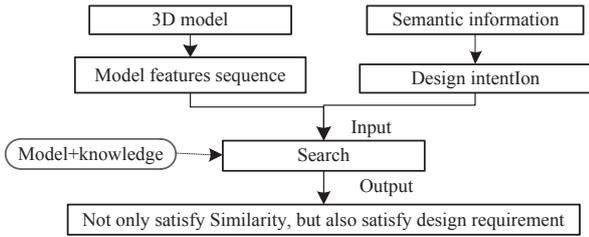


Figure 2. Ideas of Design intent and model characteristics combined with the model to retrieve

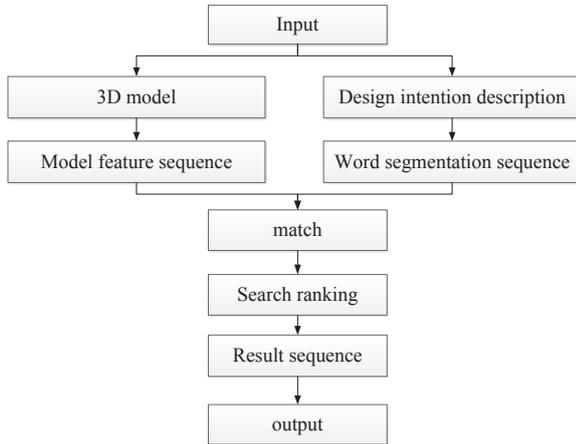


Figure 3. Flow of model search algorithm

Step 1: Input of Design Knowledge and 3D Model

The input content of the MBD model search is divided into two parts by using the method proposed by the paper. One is the models that need to be matched. The essence is to describe the geometric requirements of the search through the model, which is defined as 3D model and the other part is the design knowledge description, design knowledge is the interpretation for product geometry information and size function information, including the designer's product design requirements, design standards and so on.

Step 2: Model Matching Based on Design Knowledge Segmentation

The present invention adopts the method of keyword presence to match the model in the model library with the search request. Since the primary purpose of the design is to meet the design requirements and achieve the required functionality, the design intent is used for matching as the only control factor. On the other hand, for the characteristics of the model, the types of features are limited, and their combinations are relatively limited, so there is little effect on the matching results. Sequence of the segmentation $W=[w_1, w_2, \dots, w_i, \dots, w_n]$ and the knowledge focused index are searched, as long as there is any member of W , the model is added to the model set. The specific algorithm flow is shown in Figure 4.

Step 3: Geometric similarity calculation based on the longest common sub-sequence ratio

Step3.1: generation of Model feature sequence F

A set of feature sequences are arranged based on the modeling order. Here f_{t+1} is the characteristic that is generated after f_t in the modeling process.

$$F = [f_1, f_2, \dots, f_t, f_{t+1}, \dots, f_n] \quad (1)$$

The feature sequence is a set of feature sets arranged in time. The feature set itself can partially define the geometric information of the model, and the more accurate definition of the model geometry can be achieved by adding the time factor.

Step3.2: Calculate the longest common subsequence Fc

The longest common subsequence F_c of the characteristic sequence F and the input model i feature sequence F_i after matching, n_c is the length of F_c , N_i is the length of F_i , the ratio of the common subsequence of F to F_i is:

$$R_f = \frac{n_c}{N_i} \quad (2)$$

In particular, the number of sequences is n , and if a sequence T is a subsequence for each of the sequences, then T is referred to as a common subsequence of the sequence. If T is the longest of all the subsequences of the sequence, it is called the longest common subsequence.

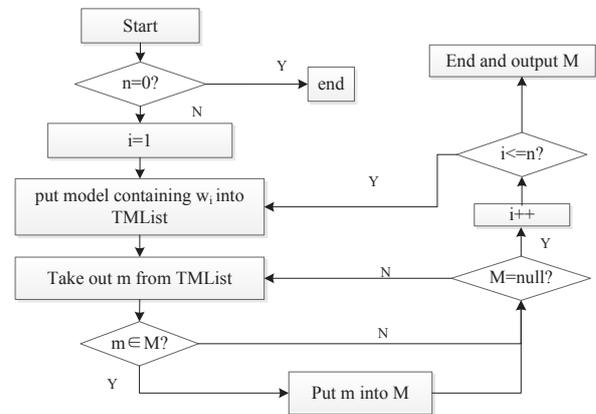


Figure 4. Flow of matching algorithm

Step 4: Calculating the Semantic Similarity of Basic Naive Bayesian

Step 4.1: Generate of the word segmentation sequence W of design knowledge

The word segmentation sequence is constructed by the order that each word appears in. Each word can be repeated, it is indicated that $W_i=W_j$ is allowed.

$$W=[w_1, w_2, \dots, w_i, \dots, w_j, \dots, w_n] \quad (3)$$

The design knowledge segmentation sequence is an ordered set of vocabularies in the order, a set of vocabularies is an effective description of the input design information entered by the user.

Step 4.2: Naive Bayesian method transformation

The Naive Bayesian method is generally used in the text classification, and its classification process is roughly through the training of multiple types of text to get a certain amount of training parameters related to the frequency of words, and finally classified text will be used for classification of word segmentation and word frequency statistics, The Bayesian basic formula is used to calculate the probability that the document belongs to each type, and the type with the maximum probability value is selected as the final classification result. Each of these types of training documents generally includes a lot.

The method applies the Bayesian in the classification transformation. Regarding each model as a type, the semantic information included in each model is as a training text of a type, where there is only one training text of each type, The semantic information of demand is used as the sorted document, and the similarity of the semantic information of each model is calculated and the ranking of the results is achieved according to the similarity of semantic similarity. The modified Bayesian method is compared with the traditional Bayesian method used for classification as shown in Fig 5.

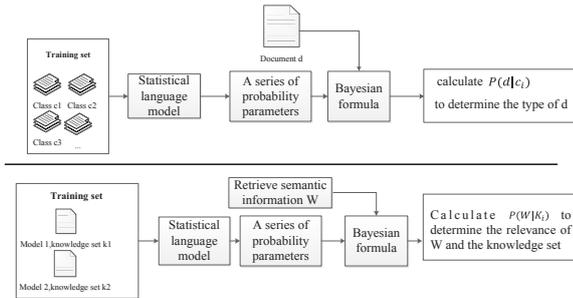


Figure 5. Naive Bayesian method for the comparison text classification and search

Step 4.3: Semantic similarity calculation based on Bayesian

The formula for calculating the semantic similarity using the Bayesian theorem is:

$$P(K_i | W) = \frac{P(W | K_i)P(K_i)}{P(W)} \quad (4)$$

Where K_i is the knowledge set corresponding to model i , W is the segmentation sequence after the design knowledge segmentation.

Since $P(W)$ is the same as for any K_i , so $P(K_i) = \frac{T_i}{\sum T_i}$ is used to calculate $P(K_i)$. Where T_i is the total number of words in K_i and $\sum T_i$ is the sum of the summations included in all sorting models.

By simplifying the simple assumptions of class conditions,

By simple assumptions in which the class conditions are independent, it's known:

$$P(W | K_i) = P([w_1, w_2, \dots, w_m] | K_i) = \prod_{j=1}^m P(w_j | K_i) \quad (5)$$

$$P(w_j | K_i) = \frac{t_j}{T_i} \approx \frac{t_j + 1}{T_i + |V|} \quad (6)$$

Where t_j is the number of occurrences of w_j in K_i , T_i is the total number of words in K_i , and $|V|$ is the total number of words in K . The semantic similarity R_w of the model can be calculated by formula (7) instead of $P(W|K_i)$.

$$P(W | K_i)P(K_i) = \frac{t_j + 1}{T_i + |V|} \cdot \frac{T_i}{\sum T_i} \quad (7)$$

Step 5: Relevance sequences based on design knowledge and model feature

Step 5.1: Normalization

Since the formulas R_f and R_w are used in two different calculation methods, the difference between the

two calculated values will be large. In order to make the two values be closer, R_f and R_w are normalized to the range of $[0,1]$. Finally, the normalization of the two values is achieved by formulas.

$$R_f^Z = \frac{R_f - R_f^{\min}}{R_f^{\max} - R_f^{\min}} \quad (8)$$

$$R_w^Z = \frac{R_w - R_w^{\min}}{R_w^{\max} - R_w^{\min}} \quad (9)$$

In this paper, we eventually adopts the method of human participation in the value of α and β , which are mainly used to balance the semantic similarity R_w and the geometric similarity R_f . The relationship is:

$$\alpha + \beta = 1 \quad (10)$$

Step 5.2: Selection of weight

The weight of the semantic similarity (R_f) is α and the weight of the model feature similarity (R_w) is β , semantic priority ($\alpha = 0.2, \beta = 0.8$), geometric priority ($\alpha = 0.8, \beta = 0.2$), the balance of semantic and geometric ($\alpha = 0.5, \beta = 0.5$), user-defined ($\alpha = x, \beta = 1 - x$).

Step 5.3: Calculate Comprehensive correlation

The linear correlation is used to calculate the correlation between the retrieved MBD model and the input

$$R = \alpha \times R_f + \beta \times R_w \quad (11)$$

Step 5.4: Output MBD model whose comprehensive correlation are front.

4 Instance verification

The following example illustrates this method and compares it with the other two methods. Select "connect two parts" as the retrieve intent. Through the word segmentation, we can get segmentation series ["connect", "two", "part"]. The part in figure 6 is selected as the input of search geometry information, and the characteristic sequence is ["cylinder", "cylinder"].

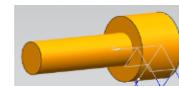


Figure 6 The part drawing

Retrieving the balance of semantic and geometric ($\alpha = 0.5, \beta = 0.5$) to get the results, as shown in Figure 7.

Method	1.18	1.12	0.94	0.55	0.40	0.38	0.31
The method of the paper							
Method based on semantic							
Method based on content							

Figure 7 the compared results after search

The following values under the model are the final comprehensive similarity values. Table 1 also gives a comparison of the results obtained by simple semantic search and content-based method search. By contrast, we can find that the method of simple semantic search can find the model to satisfy the corresponding search semantic information, but there are some models in the

result that the model is very different from the model geometry of the user. Therefore, these models are difficult to have practical meaning. The content-based approach can give a similarity of the search results model, but completely ignore the application of semantic information model, which makes a lot of models to meet the designer's geometric requirements, but it is difficult to directly meet the design requirements and can't be applied.

In comparison, the comprehensive search method based on design intention and model geometric information will be more satisfied with the user's functional requirements and geometric requirements, so as to be closer to the user's design needs. If it's applied in a larger model library, the advantages of the method will be more prominent.

5 Conclusion

This paper summarizes the shortcomings of the existing model search methods, and puts forward a new model search method for 3D models of embedded knowledge. The combination of design Knowledge and model Features can search the 3D model to meet the design functional requirements, also to meet the shape requirements of the model.

In the actual calculation, since the values of R_f and R_w vary greatly for different search requestment, the two values need to be normalized to control them between (0,1). The determination of the harmonic coefficients α and β needs further study.

Acknowledgements

This work has been supported by Project of National Key Technology R&D Program through approval No. 2015BAF17B01.

References

- 1 Lin Y J, Colello M, The International Journal of Advanced Manufacturing Technology, 23, (2004), 389-398.
- 2 Dias WPS, Knowledge-Based Systems, 20, (2007), 382-387.
- 3 Gómez-Pérez J.M., Erdmann M., Greaves M., et al, International Journal of Human-Computer Studies, 68, (2010), 641-668.
- 4 Quintana V., Rivest L., Pellerin R., et al, Computers in Industry, 61, (2010), 497-508.
- 5 Quintana V., Rivest L., Pellerin R., et al, Computers in Industry, 63, (2012), 79-90.
- 6 YU W-D, HSU J-Y, Automation in Construction, 31, (2013), 65-74.
- 7 BARRA V, BIASOTTI S, Pattern Recognition, 46, (2013), 2985-2999.
- 8 BAI J, GAO S, TANG W, et al, Computer-Aided Design, 42, (2010), 1069-1084.
- 9 Bano, Muneera, and N. Ikram, Knowledge Management for Service Oriented Requirements Engineering, (2011):494-499.
- 10 F Xuan, R.J. Howlett, Frontiers in Artificial Intelligence and Applications, 149, (2006).
- 11 H Wang, Ceglarek D, Assembly Automation, 29, (2009), 154-166.
- 12 Y Hou, J Liu, CAD&CG, 28, (2016), 1933-1939.

Table 1. similarity calculation of the model

List	Model	T_i	n_i	N_i	n_c	R_w^z	R_f^z
1		43	1	3	0	0.101448567	0
2		37	0	8	1	0.005492843	0.4375
3		59	1	5	1	0.065211437	0.7
4		67	1	12	0	0.050219444	0
5		41	0	9	2	0.096390827	0.777777778
6		54	1	13	2	0.075563228	0.538461538