Ontology Expression of MBD Model and Research on Model Retrieval Technology

Yu Yong, Gu Li  
School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, China

Yin Pu  
Institute of Telecommunication Satellite, China Academy of Space Technology, China

ABSTRACT: An MBD model knowledge representation and management method is proposed to meet the reuse requirement of MBD history models data designed in the product full three dimensional digitalization development, and a new model retrieval algorithm is verified considering geometry and non-geometry information. Firstly, we construct ontology structure based on the normative framework of MBD dataset and the purpose of model retrieval, divide it into geometry feature layer and engineering note layer in detail. Secondly, the similarity comparison method of two layers was given based on the ontology structure. Finally, MBD model retrieval is completed based on geometry and non-geometry information. The proposed method uses the ontology theory as its basis, achieving structural and semantic expression and storage of MBD dataset, thus is capable of achieving the reuse of MBD model information by similar model retrieval based on both geometry and non-geometry information.

1 INSTRUCTION

With the technological progress of Manufacturing technology and digital modeling widely adopted, large-scale equipment manufacture represented by aerospace is phasing in full three dimensional digitalization design and manufacturing methods of product based on MBD (Model Based Definition)[1]. Full three dimensional digitalization design is being carried out in full swing, full three dimensional design standards based on CATIA, UG and Pro/E are constantly improving and has a higher level of application especially in aerospace field, its gradually getting through the data chain of digitalization design and manufacturing. In real-world application, there are more and more MBD models designed by enterprises, it is becoming a problem to be solved urgently how to reuse the enterprise existing resources effectively, reduce development period of new product and lower the cost and raise modeling efficiency[2]. research and implementation of Ontology modeling and retrieval technology of MBD model is just a solution to the problem.

In order to find the right three dimensional CAD model for reuse, many scholars at home and abroad did a lot of researches on three dimensional CAD model retrieval algorithm. Li at al. firstly used CAD model to build Feature Dependency Directed Acyclic Graph, making it possible to search CAD models completely or partly[3]. Zethaban at al. put forward an intelligent CAD model retrieval algorithm based on group technology, using shape codes as description of CAD model and made it possible to retrieve CAD model with GT codes of the user as input[4]. Wang Hongshen at al. transformed partial structure retrieval of CAD model into similarity measurement of component faces of each model, achieving retrieval of partial structure of three dimensional CAD model[5]. Zhang Shuzhen at al. proposed a design resource algorithm of integrated product information models and achieved retrieval by taking into account function and attribute information of CAD model[6]. Firstly, these retrieval methods mainly consider traditional CAD models, and most of them only consider geometrical information of CAD models, some methods consider non-geometrical information as prerequisites, but geometrical information and non-geometrical information are correlative in MBD model, so they failed to clearly reflect reuse value of non-geometrical information; at the same time, these methods gave structural representations of traditional CAD model, but they didn’t aim at MBD model integrated with geometrical and non-geometrical information, failing to store and organize these structural information formally and easily to use. In order to make better use of MBD model resource, we adopt ontology technology and take OWL language as semantic descriptors of geometric feature information and non-geometric information in MBD model, then store history MBD models’ data as MBD history model data ontology knowledge database[7], so we can acquire all
geometric and non-geometric information in every MBD model by searching ontology knowledge database, finally we can retrieve models that meet users’ requirements by using matching algorithm of geometric and non-geometric information.

2. BASIC CONCEPTION

2.1 Model based definition

MBD means expressing full product definition information by using integrated three dimensional entity model, adding dimension, tolerance and process information originally defined in 2D engineering drawings into three dimensional model formally. A full defined MBD dataset should include 3D geometric information, design reference, dimension, tolerance and process information, as shown in Fig. 1:

Different from traditional 3D CAD model, MBD model describes dimension and tolerance information of some particular geometric features by adding three dimensional notes information on them; it also contains engineering information, including many process, manufacturing and checking information during product development process. Fig. 2 shows a MBD model that belongs to MBD model history database of some astronomic enterprise:

2.2 Ontology technology

Ontology originates from philosophy and aims at describing semantic information hidden in objective entity in the real world \[8\]. OWL is a kind of ontology description language which is recommended by W3C, it contains ontology elements as follows:

1) Ontology label: It describes ontology itself, including note, label, version information and ontology importing statements.

2) Class and Individual: An owl class is a special type of resource, it describes resource collection with common features; OWL individuals represent an instance of some class.

3) Property: Property play as predicate in individual statement, containing two types: Object property joins individual with individual; Data property joins individual with data.

4) Annotation and Datatype.

In certain fields and application range, we can define owl class and owl property and achieve creating domain ontology according to certain knowledge.
3. MBD MODEL ONTOLOGY MODELING

MBD model is different from traditional three dimensional CAD model, it has not only geometric feature but also 3D dimension information and engineering notes information[9]. All the information mentioned above should be completely expressed when constructing domain ontology structure. At the same time, with the purpose of MBD model information reuse, the process of constructing ontology Class and Property information of ontology should be done according to the reuse information. We refer to the MBD dataset normative definition [10] requirements put forward by Boeing Company during the process of building the skeleton structure of ontology Class(Design Reference and Other Information is not considered in this paper), conclude MBD-Model as parent class, Geometry-feature class and Engineering-notes class as its child class. The former child class is built to describe entity models’ dimension and tolerance information, the latter is to describe engineering notes information.

3.1 Geometric feature class

In consideration of retaining initial design intent and design history of designers and getting better efficiency of transmitting design parameters, we choose geometry features as descriptors for expression of geometric information. MBD models mentioned in this paper are all built in CATIA, so we refer to the definition and classification in CATIA system when building child classes of Geometry-feature.

3.1.1 Sketch based features

Sketch-based features make up a large proportion in models built in CATIA system, we focused on five common features when constructing ontology classes, including Pad, Pocket, Shaft, Stiffener and Hole class, as shown in Fig. 3:

As for Pad and Pocket class, we will give detailed description of Pad class as an example. Based on whether the sketch is a regular graphic, Pad can be divided into two child classes namely Regular-Pad (pads with regular sketches) and Irregular-Pad (pads with irregular sketches); Regular-Pad is divided into Rectangle-Sketch-Pad (pads with rectangle sketches), Circle-Sketch-Pad (pads with circle sketches) and Hexagon-Sketch-Pad (pads with hexagon sketches), their data property is shown in Fig. 4:

As for Irregular-Pad class, geometric feature information can’t be described with common property, so we adopt D2 shape descriptors to represent shape information of features [11] (we use n-dimensional vector to record shape descriptor line) and Oriented Bounding Box (OBB) to describe geometric dimension information of features. Geometric information of features combined with shape and geometric dimension information of features is shown in Fig. 5:

As for Shaft and Stiffener class, we can also use D2 shape descriptors and OBB to describe geometric information.

As for Hole class, it can be divided into Simple-Hole and Complex-Hole class. Simple-Hole has two data properties, namely has-Holedepth and has-Holediameter; Counterbored-Hole is elaborated as an example of Complex-Hole class which is shown in Fig. 6:
3.1.2 Dress up features

Dress up features belong to another important kind of feature type in CATIA system, we mainly extract three common feature types, namely Edge Fillet, Chamfer and Draft Angle (Draft Angle class is described by D2 shape descriptor and OBB), classes and their data properties are shown in Fig. 7:

![Dress-Up Features Diagram](image)

\[\begin{align*}
\text{Edge Fillet} & \quad \text{Chamfer} & \quad \text{Draft Angle} \\
\text{has-Filletradius} & \quad \text{has-Chamerangle} & \quad \text{Chamer lenght} \\
\text{has-Chamerangle} & \quad \text{Chamer angle} & \\
\text{Fillet radius} & \quad \text{Chamer lenght} & \\
\end{align*}\]

Figure 7. Dress-UP feature ontology type and data property.

3.1.3 Relationship definition

Description of feature information itself is not enough, topological relation between features should also be described. We choose feature dependence relationship to express topological structure between features\[^3\], define has-Dependence object property to express dependence relationship between two features, it is shown in Fig. 8 how to transform feature dependence relationship between feature \( F_2 \) and \( F_1 \) into ontology structure:

![Feature-dependence relationship ontology representation](image)

Figure 8. Feature-dependence relationship ontology representation

Finally, ontology modeling of geometry information in MBD model is done, then we can acquire feature information which is needed to construct ontology structure by CATIA secondary development, instantiating features to individuals in ontology structure, building relationships between individuals and accomplishing ontology expression of models. Fig. 9 and Fig. 10 show how to express geometric information of a MBD model by geometry feature ontology structure.

![MBD model geometry information](image)

Figure 9. MBD model geometry information.

3.2 Engineering notes class

As for engineering notes of MBD dataset, we constructed Engineering-notes as parent class, then construct Part-notes, Standard-notes, Material-notes and Annotation-notes class as its child classes according to MBD dataset structure. Part-notes class is taken as an example.

Part notes mainly describe product definition information which is necessary during technology planning, including heat treatment, part final treatment and part mark etc. The usual Definition of part notes is shown in chart 1:

![Table 1. Engineering note item](image)

<table>
<thead>
<tr>
<th>Code</th>
<th>Note Content</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SN00074</td>
<td>After molding, products need nondestructive flaw detection and provide the test report</td>
<td><a href="http://standards.web.boeing.com/hlgw.cgi?app=BAC&amp;spec=BAC5969">http://standards.web.boeing.com/hlgw.cgi?app=BAC&amp;spec=BAC5969</a></td>
</tr>
</tbody>
</table>

According to the structure in chart 1, we define engineering notes code as individual names, has-Content and has-Url as data properties, the ontology structure is shown in Fig. 11:

![Figure 11. Engineering note structure](image)

In conclusion, we have accomplished ontology modeling of MBD models and expressed geometry and non-geometry information in MBD dataset.
structurally [12], the final ontology class structure is shown in Fig. 12:

Figure 12. MBD model full ontology structure.

Making use of CATIA secondary development, we can extract geometry features and non-geometry engineering notes information of a history MBD Model, then we instantiate these information as individuals of ontology, reinvest ontology individuals with properties and construct relationships between them. Finally, we can achieve building MBD history model data ontology knowledge database by means of traversing data of all history models.

4. RETRIEVAL ALGORITHM

On the basis of accomplishment of MBD history model data ontology knowledge database, we can transform comparison problems of similarities between different MBD models into the problems of similarity evaluation between MBD model individuals in ontology database[13]. As mentioned above, every individual has geometry feature layer and non-geometry layer, so we give similarity evaluation ways of two layers respectively.

4.1 Comparison of geometry information layer

The ontology storage structure of a model individual’s geometry feature layer describes model’s feature dependency directed acyclic graph, all the feature properties and feature dependency relationships respectively describe node and edge information of a graph structure, so we can use graph match algorithm to calculate similarities of geometry information layer. We changed a kind of inexact adjacent graph matching of CAD model surface properties which is used to retrieve CAD models raised by Tao Songqiao [14], because graph nodes and graph edges are all changed, so we come up with new comparison methods of node properties and edge properties.

4.1.1 Geometry feature node information comparison method

When comparing node properties of two geometric features[15], we mainly consider the following factors:

1) Geometric feature type

Firstly geometric feature type is considered, if these two geometric features are of different type, their similarity is 0 and there is no need to compare the following factors.

2) Geometry dimension

Geometry dimension information of geometry features is expressed by OBB, so we can suppose there are two geometry features whose OBB length, width, height are respectively \(a_1\), \(a_2\), \(a_3\) and \(b_1\), \(b_2\), \(b_3\), their geometry dimension similarity is recorded as \(S_Q\):

\[
S_Q = \frac{1}{3} \sum_{i=1}^{3} \left(1 - \frac{|a_i - b_i|}{\max(a_i, b_i)}\right)
\]  

(1)

3) Geometry shape

As for regular-sketch feature, we propose there are two features whose sketch length and sketch width are respectively \(L_1, W_1\) and \(L_2, W_2\) (as for sketch described by radius \(R\), we count \(L=W=R\)), stretch dimension are respectively \(D_1\) and \(D_2\), so we can calculate shape similarity \(S_D\) by using cosine similarity:

\[
S_D = \frac{L_1L_2 + W_1W_2 + D_1D_2}{\sqrt{L_1^2 + W_1^2 + D_1^2} \times \sqrt{L_2^2 + W_2^2 + D_2^2}}
\]  

(2)

As for irregular-sketch feature, feature shape information is expressed by D2 shape descriptors, so we propose that there two features whose shape distribution line are respectively \(H_1\) and \(H_2\): \(H_1=\{h_{11}, h_{12}, ..., h_{1n}\}\) and \(H_2=\{h_{21}, h_{22}, ..., h_{2m}\}\). The final shape similarity \(S_D\) is calculated by EMD (Earth Mover’s Distance), EMD is an effective way of calculating distance between two vectors.

\[
S_D = 1 - EMD(H_1, H_2)
\]  

(3)

4) Dimension and tolerance

We suppose that there are two features whose tolerance class are respectively \(IT_1\) and \(IT_2\), surface roundness are respectively \(R_1\) and \(R_2\):

\[
S_T = \frac{1}{2} \left(1 - \frac{|IT_1 - IT_2|}{\max(IT_1, IT_2)}\right) + \frac{1}{2} \left(1 - \frac{|R_1 - R_2|}{\max(R_1, R_2)}\right)
\]  

(4)

Finally, \(Q_1\) and \(Q_2\) represent two compared features, the whole similarity is \(S_{sim}\), calculation formulas is as follows:

\[
S_{sim} = \begin{cases} 
0, & Q_1 \neq Q_2; \\
\partial_1 \times S_Q + \partial_2 \times S_D + \partial_3 \times S_T, & Q_1 = Q_2; 
\end{cases}
\]  

(5)
Where $\partial_1$, $\partial_2$ and $\partial_3$ represent weight of different comparison factor.

4.1.2 Geometry feature edge property comparison
We propose that there are two geometry features namely A and B, edge property is $L(A,B)$ if A directs to B, calculation formulas is as follows:

\[
L(A,B) = \begin{cases} 
0, & \text{none} \\
1, & B \text{ has } \text{dependence } A \\
-1, & A \text{ has dependence } B 
\end{cases}
\]  

(6)

None means there isn’t has-dependence relationship between A and B.

4.1.3 Graph match of geometry information layer
By adopting the comparison method of graph node and graph edge mentioned above and changing the graph match algorithm, we can calculate geometry information layer similarity between two compared model individuals, counted as $G_{sim}$

4.2 Non-geometry engineering notes information comparison
Comparison of non-geometry engineering notes belongs to short-text comparison problem. Short texts in this paper—MBD engineering notes, they face special domain, include much semantic information, so we treat it in the following way of thinking:

There are many engineering notes in an MBD model, but representation degree of a single note on the whole model semantic information is very limited, so we choose to combine all engineering notes into a span of text, then it is possible to compare engineering notes of two compared MBD models by comparing engineering notes texts of MBD models.

Semantic information of a single note can be expressed by basic terminologies of machine manufacturing and unique terminologies related to product manufacturing and inspections of the company. So semantic information of an MBD model can be expressed by keywords of engineering notes text, realization procedures are presented as follows:

1) Refer to basic terminologies of manufacturing technology in GB/T 4863-2008(T 4863-2008 basic terminology of manufacturing technology) and also collected company unique terminology included in MBD models, then we use them to construct key-word dictionary.

2) Import key-word dictionary as user dictionary into ICTCLAS.

3) Traversing all model individuals’ engineering notes in ontology dataset, we can build engineering notes texts of all MBD models and use them as engineering notes corpus.

4) As for the two compared models A and B, we extract key words by using ICTCLAS added with new tokenization lexicon, then adopted TF-IDF algorithm to calculate TF-IDF weight of all key words in model A and B by using the generated corpus. Then we extract twenty key words from large to small, respectively record them as $A\{a_1,a_2,a_3…a_{20}\}$ and $B\{b_1,b_2,b_3…b_{20}\}$.

5) Combining all elements in vector A and B into a new set, calculate each text word frequency for each word in the new set, record the result as word frequency vector of A and B: $C_A\{C_{A1}, C_{A2}, C_{A3}…C_{A40}\}$ and $C_B\{C_{B1}, C_{B2}, C_{B3}…C_{B40}\}$.

6) Calculate cosine similarity between these two vectors $C_A$ and $C_B$, record the result as $E_{sim}$ use it as similarity evaluation criterion of engineering notes content of two compared models.

Finally, we can combine similarity calculation of engineering notes and the former geometry information similarity, the final similarity calculation formula is as follows:

\[
Sim = \alpha \times G_{sim} + \beta \times E_{sim}
\]  

(8)

Where $\alpha$ and $\beta$ respectively represent weight of geometry and non-geometry information.

5. APPLICATION AND VERIFICATION

All data come from some aerospace enterprise, we use MBD history models of this enterprise as data source, construct ontology structure in Protégé, achieving importing 232 CATIA MBD models into ontology structure by using Jena (Ontology SDK).

In our experiments, we experimented the MBD model in Fig. 13 and distributed weight as follows:

1) $\alpha = 1.0$ $\beta = 0.0$ (only concern about geometry information)

2) $\alpha = 0.5$ $\beta = 0.5$ (concern about geometry and non-geometry information in the same weight)

3) $\alpha = 0.25$ $\beta = 0.75$ (concern more about non-geometry information)
The final retrieval result is shown in table 2:

<table>
<thead>
<tr>
<th>Retrieval Model</th>
<th>Weight</th>
<th>Retrieval Result Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>New-01</td>
<td>$\alpha = 1.0$, $\beta = 0.0$</td>
<td>XXX-7_21-05, XXX-7_21-09, XXX-7_21-10, XXX-7_11-03</td>
</tr>
<tr>
<td>New-01</td>
<td>$\alpha = 0.5$, $\beta = 0.5$</td>
<td>XXX-7_21-09, XXX-7_21-10, XXX-7_21-05, XXX-7_11-03</td>
</tr>
<tr>
<td>New-01</td>
<td>$\alpha = 0.25$, $\beta = 0.0$</td>
<td>XXX-7_21-09, XXX-21-015, XXX_03-012, XXX_03-008</td>
</tr>
</tbody>
</table>

It can be seen from chart 2 that similar models can be retrieved according to similarity degree of geometry feature information and topology between features when we only consider geometry information similarity ($\alpha = 1.0$, $\beta = 0.0$). If geometry information and non-geometry is of same weight ($\alpha = 0.5$, $\beta = 0.5$), XXX-7-21-09 has the highest similarity with the retrieved model, its engineering notes information is shown in Fig. 14:

![Figure 14. Engineering notes information.](image-url)
Similar with the retrieved model, XXX-7-21-09 has engineering notes that describe product weight, non-destructive inspection, undated dimension tolerance and material selection.

If $\alpha = 0.25$, $\beta = 0.75$, there are new retrieval results in range 3, 4, 5, these models has much difference with the retrieved model, but there are many similarities in engineering notes information which is of much retrieval value.

6. CONCLUSION

This paper is aimed at MBD model retrieval, we express geometry and non-geometry information in MBD model in a structured and semantic way by using ontology technology and transform MBD history modeling data into off-line MBD history modeling ontology knowledge database so we can achieve multi-granularity retrieval. Also we propose similarity algorithm of geometry layer and non-geometry layer and have achieved MBD model retrieval considering geometry information and non-geometry information at the same time by adopting weighted calculation of two layers, it is shown in experiments that:

1) By transforming MBD history modeling data into ontology knowledge database, it will become very easy to acquire all geometry feature information and non-geometry engineering notes information, managing knowledge effectively and laying the foundation of knowledge retrieval and reasoning for modeling process of companies.

2) Different from traditional model retrieval, we consider geometry and non-geometry information and the relationship between them comprehensively, achieving retrieval of MBD models with similar geometry semantic, design semantic or manufacturing semantic information according to needs of the company. In this way, reuse value of MBD model is greatly improved and it will greatly improve modeling quality and efficiency for design departments.

REFERENCES