3D Part Similarity Comparison based on Levels of Detail in Negative Feature Decomposition Using Artificial Neural Network

Han-Chung Cheng¹, Chih-Hsing Chu², Eric Wang³ and Yong-Se Kim⁴

¹National Tsing Hua University, <u>d937802@oz.nthu.edu.tw</u>
²National Tsing Hua University, <u>chchu@ie.nthu.edu.tw</u>
³Sungkyunkwan University, <u>ewang@skku.edu</u>
⁴Sungkyunkwan University, <u>yskim@skku.edu</u>

ABSTRACT

Duplicate designs consume a significant amount of company resources during product development. Search for similar parts for a given query part, which facilitates design reuse, is crucial to avoiding this problem. Previous studies have only compared parts on a complete scale, not on a partial scale. This paper proposes a novel scheme which incorporates the concept of LOD (Levels of Detail) into 3D part comparison in order to assess partial similarity. Different LOD variants are generated from negative feature decomposition of a solid model. A human comparison behavior model (HCBM), mainly consisting of a back-propagation artificial neural network (ANN), is established by training with the result of a similarity ranking experiment. It combines the dissimilarity value at each LOD based on a modified D2 distribution. Test examples show that the proposed scheme is effective in 3D part search with LODs.

Keywords: Similarity assessment, levels of detail (LOD), negative feature, feature recognition, part search, design retrieval.

1. INTRODUCTION

Product development plays an important role in competition among modern enterprises. Previous studies have identified that lowering costs, improving quality and shortening the time for product development are the key factors in the development process [1]. Design is a critical stage in the development of any new product. It generally determines 80% of the total development costs, which can only be decreased by less than 25% after this stage [2]. Over 75% of design activities involve design modification, variant design or case-based design [3]. Engineers usually adapt existing products or technologies to fulfill new requirements, enhance functionalities, and/or lower costs. Enterprises also encourage design reuse in development projects for the same purpose. Many companies have introduced product data management (PDM) to manage product design data, but effective tools for search similar parts by the way of design intent are not yet available. Using a text-based search engine to search on file names and keywords does not always produce the expected results. Consequently, most of the time parts will be recreated by constructing different CAD files for the same function, making multiple copies of the same data. This thus raises cost and complicates the data management. Therefore, advanced CAD technologies should be adopted to manage part design data.

Most assessment schemes for 3D part similarity generate a *Shape Signature* from the CAD model, and distinguish different models based on the dissimilarity of the signature defined by a distance function [4-7]. *Shape Signature* is a high-level abstraction of 3D geometry that is unique to each model. However in practice, the comparison of two 3D parts often needs to be partially similar, and necessarily to be similar in every detail. A partial similarity, rather than a complete similarity may better satisfy part search for design reuse. Previous studies were lack of support to such partial-similarity search. To overcome this deficiency, this paper proposes a part comparison scheme that integrates the concept of LOD (Levels of Detail) in similarity assessment of 3D shapes. It generates different LOD variants of a 3D model from the negative feature decomposition process, which has a close correlation with design functions. The similarity assessment combines the dissimilarity value at each LOD using a modified D2 distribution. A back-propagation artificial neural network (ANN) is established to characterize human cognition in an experiment of ranking similar parts. The network trained with the experimental result, namely a Human Comparison Behavior Model (HCBM), serves as a centerpiece for similarity assessment. Test examples show that the proposed scheme performs well in discriminating 3D mechanical parts. It facilitates design reuse by offering 3D part search with partial similarity.

2. SIGNIFICANCES OF 3D PART SEARCH TECHNOLOGY

3D part search helps enhance the effectiveness of design reuse in product development [8]. Design reuse influences not only on CAD construction, but also production utilization, inventories, tooling, and other activities in a product lifecycle. Product data management that permits knowledge reuse would need a combination of text-based as well as content-based search. Taking on a wider context, strategic sourcing to find appropriate component providers becomes an important issue in the distributed environment [9]. Broadening the supplier base is a significant task for enterprises that generally outsource many components and/or customized designs to their suppliers [10]. Intelligent searching for geometrically as well as functionally similar parts would make huge numbers of parts and suppliers available electronically all over the world. Similarity assessment for 3D components is an enabling technology in supplier selection in this case [11]. In summary, 3D mechanical part search technology provides the following benefits for companies [12]:

- Avoiding duplicate design: engineers would like to upload a query model; search through existing designs in the company legacy data; obtain similar parts, and modify them for the current design task, therefore save hidden costs for companies.
- Facilitate knowledge management of product lifecycle: intelligent component search promotes transfer of
 knowledge fragments, which maximizes the value of knowledge management in modern enterprises. It provides a
 good approach to disseminating knowledge within and among organizations.
- *Expedite B2B E-Commerce*: design retrieval allows a designer to search for potential suppliers and technology providers over the Internet. A supplier that generates "similar" parts to the intended use would be potential candidates for collaboration in outsourcing and engineering.



Fig. 1: Computation procedure for search similar 3D parts.

3. 3D PART SIMILARITY ASSESSEMENT WITH LOD

3.1 Present State of Similarity Comparison Method

Most previous approaches generate a "shape signature" from the CAD model, and compare different signatures according to a pre-defined measure function. They can be classified into six categories based on the CAD information that makes up the shape signature: feature-based similarity, topology-based similarity, histogram-based similarity (e.g., shape distribution, shape statistics), graph-based similarity (e.g., skeletal graph, reeb graph, aspect graph), product information-based similarity (e.g., section image, group technology) and Octree-based similarity [4-5]. The choice of

shape signature and similarity function can influence discrimination capability of similarity assessment [13-14]. These assessment schemes include all the geometric information in the shape signature, and thus cannot compare partial similarity.

This work develops a 3D shape similarity assessment scheme integrated with the LOD concept. Fig. 1 shows the search procedure based on the scheme. First, the user uploads a query part or classified sets of negative features that may represent the current design intention. Different feature-based LOD models are generated by negative feature decomposition. A shape signature based on a modified D2 distribution is then created at each level from the distance sets of all the features belonging to it. Every component to be assessed is then retrieved in sequence from the database, along with its shape signature of each level, and compared with the query model. The user is allowed to specify a set of comparison attributes. We then compute the dissimilarity value between the query/candidate parts at each level and combine them using an artificial neural network.



Fig. 2: Different feature-based LOD models of 3D part.



Fig. 3: The D2 shape histograms for different LOD models.

3.2 Feature-Based LOD Model

Volumetric decomposition [15] provides hierarchical approximation capability based on the feature of a part, and generates this part by with a Boolean operation. Convex decomposition works well for polyhedral objects, and provides decomposition methods for form feature and negative feature based on convex decomposition [16-18]. Previous studies have also noted the importance of manufacturing feature recognition in automatic process planning [16]. Therefore, a LOD model is defined as a representation of 3D part with different levels of features in this paper (see fig. 2). It corresponds to different groupings of the hierarchy in negative feature decomposition [15], which consists of one positive base feature (stock) and other negative removal volumes (machining volumes).

3.3 D2-Based Similarity Assessment

The shape distribution of a 3D object is generated based on a chosen shape function such as A3, D1, D2, D3 and D4. The D2 function is regarded as a better discrimination criterion among them [19]. Thus we adopt D2 distribution as the major shape signature in the complete/partial similarity assessment. However, the part comparison using regular D2 cannot distinguish certain geometries, e.g. Fig. 3(a) indicates that two different parts consisting of negative features have a similar D2 distribution [20]. The reason is that the distances taken in regular D2 generation do not distinguish among in-distance, out-distance, and mixed distance, which fails to differentiate minor shape discrepancies or severe feature interactions. In contrast, Figs. 3(b) and 3(c) show the distributions of the same parts at lower LODs. Therefore, we will improve the discerning capability of D2 distribution in two ways. First, the generation of D2 distribution is modified to differentiate in/out/mixed distances. In addition, the comparison results at all LODs are taken into account in similarity assessment, which characterizes the model construction process to a certain degree.

We develop a four-step algorithm for similarity assessment complying with the above ideas. Instead of comparing the dissimilarity of two complete parts, it calculates the dissimilarity between models at each LOD level using the modified D2.

Generate the modified D2 distribution at each level

Generate a set of distances between two random points for every feature of each level, then compute D2 distributions by the distance sets of all features for each level. Such a D2 distribution only utilizes in-distances within features. This step can be described as follows:

- Suppose $L = \{LOD_1, ..., LOD_i, ..., LOD_n\}$ denotes a set of LOD models at *n* levels and the feature set of LOD_i is given by $F^i = \{f_1, ..., f_n, f_n\}^i$.
- Triangulate each f_j in F^i into facets. Assume d_j represents a set of distances between two random points

in the facets. All the distance sets of LOD_i are given by $D^i = \{d_1, ..., d_m\}^i$ and generated as:

- a. Assume that the triangular facets of f_j are written as $T = \{t_1, t_2, ..., t_k\}$. Calculate the surface area of each triangular facet; estimate the total area (TA) by adding the facet area, and log the accumulated TA corresponding to each facet.
- b. Generate a random value within [0,TA] and identify the corresponding facet based on the log value.
- c. Generate two random values r_1 , r_2 within [0,1]; then calculate a random point p with the three vertices P_1 , P_2 , and P_3 of the facet and r_1 , r_2 according to:

$$p = (1 - \sqrt{r_1})p_1 + \sqrt{r_1}(1 - r_2)p_2 + \sqrt{r_1}r_2p_3$$
(1)

- d. Calculate d_j , which consists of a set of distances between two points randomly generated from f_j .
- Generate the modified D2 distribution ($h_i(bin, \delta)$) of LOD_i as:
 - a. Derive the maximum distance (D_MAX) from D^i and bin-width is given by $\frac{D_MAX}{bin}$ number.
 - b. Set the corresponding bin_i for every distance value of D^i , $j = 1,...,bin_number$.
 - c. Generate a shape distribution $h_i(bin, \delta)$ based on occurrence probability δ_j of distance in each bin_j ,
 - where *bin_number* is the user input parameter.
- Define the level D2 set of LOD model as $H = \{h_1, \dots, h_n\}$.

<u>Calculate the dissimilarity value of each level</u>

- Since the level number of each LOD model is different, the dissimilarity value is calculated as follows:
- Determine the level maximum (max_n) of the query model and all candidate models.
- Derive $H^Q = \{h_1, ..., h_i, ..., h_q\}^Q$ for the query model, and $H^C = \{h_1, ..., h_i, ..., h_c\}^C$ for a candidate model.
- For each h_i , we need to distinguish four conditions in calculating the dissimilarity value of each level between the query and candidate parts.
 - a. If $\max_{n > q}$ and $\max_{n > c}$:
 - > If h_i exist in both query and candidate models, the dissimilarity value (DS_i) is defined as.

$$DS(H^{1}, H^{2}) = \sum_{i=1}^{bin} |\delta_{i}^{1} - \delta_{i}^{2}|$$
(2)

- > If h_i does not exist in either of them, then $DS_i = 1$.
- > If h_i does not exist in both, then $DS_i = 0$.

b. If $\max_n q = q$, $\max_n c = c$:

- > If h_i exists in both query and candidate models, then calculate DS_i based on (2).
- > If h_i does not exist in the candidate model, then $DS_i = 1$.
- c. If $\max_{n > q} \max_{n = c} c$:
 - > If h_i exists in both query and candidate models, then calculate DS_i based on (2).
 - > If h_i does not exist in the candidate model, then $DS_i = 1$.
- d. If $\max_n = q = c$, then calculate DS_i according to (2).
- Assume the value set of one candidate model is $L = \{DS_1, ..., DS_{i_1}, ..., DS_{max_n}\}$; the dissimilarity value of all candidate models is $DisSet = \{L^1, ..., L^k, ..., L^p\}$, where *P* denotes the number of candidate parts.
- <u>Combine the dissimilarity values of all LOD levels</u>

User-driven search: the user is allowed to specify individual preferences in searching similarity part such as complete/partial similarity and weights of each LOD.

- Specify the number of LODs (L_n) in similarity assessment and select between complete/partial comparison.
- Input weights for each level within [0,1] and transform normalized weight (w_i) according to (3):

$$w_i = \frac{R_i}{\sum_{i=1}^{L-n} R_i}$$
(3)

where R_i denotes the weighting factor of LOD_i .

- Obtain $L^k = \{DS_1, ..., DS_i, ..., DS_{L_n}\}$ from *DisSet* and combine them into *DisVaule^k* as follows:

$$DisValue^{k} = \sum_{i=1}^{L_{n}} w_{i} DS_{i}$$
(4)

For a given set of p candidate parts, we compute the corresponding DisVaule^k and determine their similarity ranks with respect to the query part.

Human Comparison Behavior Model (HCBM): this model characterizes the intelligence of human comparison in determining $L^k = \{DS_1, ..., DS_i, ..., DS_{L_n}\}$ from *DisSet*; then generates one similarity value (*DisVaule^k*) from the model as:

$$DisValue^{k} = HCBM(DS_{1},...,DS_{i},...,DS_{max})$$
(5)

3.4 Human Comparison Behavior Model (HCBM) Using Artificial Neural Network

Artificial neural networks (ANN) construct complex system models without the need of explicit descriptions of the system behavior or rules. A network consists of processing elements and connections. Each processing element has a single output signal that fans out along connections to every other processing element. An MLFF (Multi-Layer Feed forward) network using arbitrary squashing functions can approximate most functions of interest to any desired degree of accuracy [21]. Any continuous mapping based on back-propagation can be approximately represented by multi-layer networks with sigmoid output functions [22]. This study adopts back-propagation network (BPN) for HCBM by combining the dissimilarity value of each level. The corresponding experiment comprises the following steps.

- <u>Produce test 3D parts</u> According to the limitations of negative feature decomposition, we generate 32 test parts and randomly select one query part from them.
- <u>Acquire BPN input data</u>

Utilize the similarity assessment algorithm to obtain input values of BPN as $DisSet = \{L^1, ..., L^{32}\}$. Since the maximum LOD number is 6 in all the parts, BPN has six input nodes, and the dissimilarity value set of each candidate part is $L^i = \{DS_1, ..., DS_6\}$.

• Obtain the target value of BPN with a part-comparison experiment

Sequence values are obtained by a series of part similarity comparison experiments in which eight engineering graduate students participate. These sequence values are manually determined by their individual perception on the similarity between the query part and all the candidate parts. They are then transformed into a dissimilarity values as:

Dissimilarity Value =
$$1 - [(32 - \text{Sequence Value}) \times (1/32)]$$
 (6)

where the sequence value 1 indicates that the query and candidate parts are the most similar. The target value is derived as the average of eight dissimilarity values for each candidate part. The BPN has only single output, corresponding to the average dissimilarity value.

Train BPN to characterize human comparison behavior

There are many parameters (e.g. learning rate, momentum, network architecture, training/testing ratio, epochs, and input sequence) need to be determined through the training process. The procedure for finding a proper combination of these parameters is as follows:

- Select the error function
 The function estimates the error between the output and the target. The mean square error (MSE) is used as the performance function during the training process.
- Choose the activation function
 Since the similarity assessment is a continuous mapping with a value within [0,1], a binary sigmoid function is chosen as the activation function of BPN. The input transforms the output based on:

$$f(x) = \frac{1}{1 + e^{-x}}$$
(7)

Determine the momentum (μ)
Obtain five values for μ by decreasing μ in a step of 0.1 from 0.9 to 0.5, with other parameters fixed. The MSE of the testing set becomes minimal at μ=0.9. We next obtain ten values by decreasing μ in a step of 0.01 from 0.95 to 0.9 to determine whether MSE begins to rise when μ > 0.9 or when μ < 0.9.
Decide the learning rate (LR)

Find the condition when the MSE starts to increase when LR>0.1, and again obtain nine values incrementing in a step of 0.01 from 0.01 to 0.1. The MSE of the testing set becomes minimal in the case of LR=0.05.

Confirm training/testing number

The number of test parts is 32. The training/testing ratio is chosen as 5:3 for multifold cross-validation. The testing set contains J5, J11, J12, J14, J15, J16, and J26 (see implementation for these models). The validation set include J6, J13, J17, J22, and J23, which are randomly chosen from the testing set. The training/testing are used in the training process of BPN and the validation set verify the result of testing unknown parts.

Judge the epochs

A MSE figure helps identify that epochs=30000 is the stop point during the training under the parameter combination as μ =0.9, LR=0.05 and network architecture 6-4-1. The MSE of the testing set becomes minimal at this point.

- Construct the network architecture
- Determine a proper training-parameter combination according to the previous steps. The hidden layer of BPN is single in the architecture-parameter and according to (8) and (9) the maximum and minimum of hidden nodes are 13 and 4. Among the five network architectures from [13,4], the MSE of the testing set is lowest in the 6-6-1 architecture. The 6-7-1 architecture raises the MSE.

$$H = \frac{M+N}{2} \tag{8}$$
$$H = 2N+1 \tag{9}$$

where H denotes the number of hidden nodes, N represents the number of input nodes, and M is the number of output nodes. A large number of hidden nodes can lower the MSE of the training set in the

result, but the MSE of testing set also rises when H>6. The best parameter combination are thus chosen as μ =0.9, LR=0.05, epochs=30000 and 6-6-1 architecture.

• <u>Integrate HCBM into the similarity assessment algorithm</u> The trained BPN serves as the underlying mechanism of the HCBM. The major model parameters include network architecture (6-6-1), sigmoid function, and weights of all connections. The network is integrated into the algorithm described in Section 3.1 to reflect the behavior of human similarity cognition in 3D part search.







Fig. 5: Partially similar parts with respect to a query model with user-input weights.

4. IMPLEMENTATION

The search scheme proposed by this paper was implemented using ACIS C++ library [24]. All 32 test models are converted into the SAT format prior to experiment. All the test parts (see Fig. 6) have gone through negative feature decomposition, among which Parts J01-J15 are the ones previously published in [25]. All the results are listed in a decreasing order of part similarity. Fig. 4 shows the result of complete part comparison with the weights of 1, 0.6, 0.4, 0.2, 0.1 and 0.1 chosen by the user for each LOD. Fig. 5 illustrates the results of partial similarity based on different LOD models (2 and 3 LODs respectively) of a second query part. The weights used in the first test remain the same in this test. The hierarchical structure of the query model in NFD is also shown. Note that the retrieved first five models are different in both cases. Fig. 6 shows the result of complete part comparison produced by human with a third query model. Five validation models are randomly chosen and listed at the bottom. Fig. 7 depicts the rank generated from the HCBM with the same query model. For the validation samples (unknown to the HCBM), the similarity rank of Part J17 is different from that of the human comparison shown in Fig. 6. In addition, the first four parts have been successfully identified among the top five. Tab. 1 shows the search results with different query models. Only the best

three models are included. All the identical models have been identified from the database. The above results illustrate that the network has captured the human discerning capability to a large extent. They also validate that the proposed scheme can search for parts that are similar to the ones being queried.

5. CONCLUSION

Automatic 3D part search facilitates design reuse in new product development. This requires part comparison which is based on partial similarity and captures design intent. Previous studies have provided a variety of methods for assessing 3D shape similarity, but do not yet address partial similarity assessment. This study proposes a novel scheme of 3D part comparison that integrates the concept of LOD into similarity assessment process. LOD variants correspond to different subtrees in negative feature decomposition of a solid model. Part comparison incorporates the dissimilarity value of each level computed with a modified D2 shape distribution of the corresponding LOD model and weighted by user inputs. In addition, a human cognition experiment is conducted to determine the similarity rank of a set of test models with respect to a query part. A Human Cognition Behavior Model based on back-propagation artificial neural network is trained with the experimental result. It then tests on another set of validation models and thus verifies its discriminating capability of 3D shapes. Future work based on this study is as follows. First, additive volumes need to be considered in part comparison with form feature decomposition. Moreover, a part search engine of practical use should provide GoogleTM-like user interfaces for more complex query with Boolean operations applied to feature attributes (e.g. form, kind, parameters).

6. REFERENCES

- [1] Hoque, M.; Akter, M.; Monden, Y.: Concurrent engineering: A compromising approach to develop a feasible and customer-pleasing product, International Journal of Production Research, 43, 2005, 1607-1624.
- [2] Hacker, W.: Improving engineering design contributions of cognitive ergonomics, Ergonomics, 40(10), 1997, 1088-1096.
- [3] Wood, W. H.; Agogino, A. M.: A Case-based Conceptual Design Information Server for Concurrent Engineering, Computer-Aided Design, 28(5), 1996, 361-369.
- [4] Cardone, G. S. K.; Karnik M.: A Survey of Shape Similarity Assessment Algorithms for Product Design and Manufacturing, Journal of Computing and Information Science in Engineering, 3, 2003, 109-118.
- [5] Natraj, I.; Subramaniam, J.; Kuiyang, L.; Yagnanarayanan, K.; Karthik, R.: Three-dimensional Shape Searching: State-of-the-art Review and Future Trends, Computer-Aided Design, 37, 2005, 509-530.
- [6] Elinson, A.; Nau, D.S.; Regli, W. C.: Feature-based similarity assessment of solid models, ACM Fourth Symposium on Solid Modeling and Applications, 1999, 8-11.
- [7] Ramesh, M.; Yip-Hoi, D.; and Dutta, D.: Feature Based Shape Similarity Measurement for Retrieval of Mechanical Parts, ASME Journal of Computing and Information Science in Engineering, 1(3), 2001, 245–256.
- [8] Iyer, N.; Jayanti, S.; Lou, K.; Kalyanaraman, Y.; Ramani, K.: Shape-based searching for product lifecycle applications, Computer-Aided Design, 37, 2005, 1435–1446.
- [9] Chu, C. H.; Chang, C. J.; Cheng, H. C.: Empirical Studies on Inter-Organizational Collaborative Product Development in Asia Pacific Region, ASME Journal of Computing & Information Science in Engineering, 6(6), 2006, 179-187.
- [10] Yang, J.; Lai, F. J.: Harnessing value in knowledge acquisition and dissemination: strategic sourcing in product development, International Journal of Technology Management, 33(2-3), 2006, 299-317.
- [11] Tanskanen, K.: On-site component selection support using WWW, International conference on engineering design, 2, 1997, 243-248.
- [12] Chu, C. H.; Hsu, Y. C.: Similarity Assessment of 3D Mechanical Components for Design Reuse, Robotics & CIM, 22(4), 2006, 332-341.
- [13] El-Mehalawi, M.; Miller, R. A.: A Database System of Mechanical Components Based on Geometric and Topology Similarity. Part I: Representation, Computer-Aided Design, 35, 2003, 83-94.
- [14] El-Mehalawi, M.; Miller, R. A.: A Database System of Mechanical Components Based on Geometric and Topology Similarity. Part II: Indexing, Retrieval, Matching, and Similarity Assessment, Computer-Aided Design, 35, 2003, 95-105.
- [15] Han, J. H.; Pratt, M.; Regli, W. C.: Manufacturing feature recognition from solid models: A status report, IEEE Transactions on Robotics and Automation, 16(6), 2000, 782-796.
- [16] Kim, Y. S.; Wilde, D. J.: A Convergent Convex Decomposition of Polyhedral Objects, Journal of Mechanical Design, 114(3), 1992, 468-476.
- [17] Kim, Y. S.: Recognition of Form Features Using Convex Decomposition, Computer-Aided Design, 24(9), 1992,

461–476.

- [18] Waco, D.; Kim, Y. S.: Geometric Reasoning for Machining Features Using Convex Decomposition, Computer-Aided Design, 26(6), 1994, 477–489.
- [19] Osada, R.; Funkhouser, T.; Chazelle, B.; Dobkin, D.: Shape Distributions, ACM Transactions on Graphics, 21(4), 2002, 807-832.
- [20] Ip, C. Y.; Lapadat, D.; Sieger, L.; Regli, W. C.: Using Shape Distributions to Compare Solid Models, Seventh ACM/SIGGRAPH Symposium on Solid Modeling and Applications, Saarbrucken, Germany, 2002, 273-280.
- [21] Hornick, K., Stinchcombe, M. and White, H.: Multilayer Feedforward Networks are Universal Approximators, Neural Networks, 2, 1989, 359-366.
- [22] Funahashi, K.: On the Approximate Realization of Continuous Mappings by Neural Network, Neural Networks, 2, 1989, 183-192.



Fig. 6: Similarity sequence of human part comparison experiment.

- [23] http://www.mathworks.com/
- [24] <u>http://www.spatial.com/</u>
- [25] Kim, Y. S.; Jung, Y. H.; Kang, B. G.; Rho, H. M.: Feature-Based part similarity assessment method using convex decomposition, Proceedings of 23rd ASME Computers and Information in Engineering Conference, DETC2003/CIE-48184, Chicago, Sep. 2003.



Fig. 7: Similarity sequence of HCBM.



Tab. 1: Test results of HCBM corresponding to different query parts.