Springer

KSME

Journal of Mechanical Science and Technology 31 (12) (2017) 5627~5638
www.springerlink.com/content/1738-494x(Print)/1976-3824(Online)
DOI 10.1007/s12206-017-1103-3

Shape distribution-based approach to comparing 3D CAD assembly models †

Hyungki Kim[1], Moohyun Cha[1] and Duhwan Mun[2,*]

[1]*Mechanical Systems Safety Research Division, Korea Institute of Machinery and Materials, Daejeon, Korea*
[2]*Dept. of Precision Mechanical Engineering, Kyungpook National University, Gyeongsangbuk-do, Korea*

(Manuscript Received July 9, 2017; Revised August 3, 2017; Accepted August 3, 2017)

Abstract

As a key technology for three-dimensional (3D) model retrieval from model databases, the ability to perform comparisons of 3D models is important. In particular, during the product design phase, 3D CAD models should be robustly retrieved for design reuse or machining process planning. However, typical retrieval methods that are based on model names, codes, or annotations are limited in terms of their usability and robustness. Therefore, there are increasing requirements for shape-based 3D model retrieval techniques, as well as the ability to perform 3D model comparisons. In this paper, we propose a shape distribution-based 3D CAD assembly comparison method that has the ability to identify dissimilarities in 3D CAD assembly models by differentiating between assembly relationships and part-shape dissimilarities. This is different from existing methods, which are limited to only shape dissimilarities. We present experimental results for various test cases by comparing our proposed method and existing methods. Based on our experiment results, we found that our method can enable a comprehensive comparison of 3D CAD assembly models including assembly relationship dissimilarities, part-shape dissimilarities, and overall model dissimilarities.

Keywords: 3D CAD assembly model; Assembly relationship dissimilarity; Shape distribution-based comparison; Part-shape dissimilarity

1. Introduction

With recent advancements in computing power, as well as modeling and visualization techniques, 3D shapes have been widely used in many applications including product design, engineering simulation, virtual prototyping, entertainment, and cultural heritage [1-7]. This trend has led to a very large number of three-dimensional (3D) shape models.

Shapes can be interpreted in various ways. Kendall [8] defined a shape as "all the geometrical information that remains when location, scale, and rotational effects are filtered out from an object". Based on this definition, we define a 3D Computer-aided design (CAD) model comparison as "quantitatively evaluating similarity among 3D shape data contained in 3D CAD models".

Obtaining a similarity measurement of 3D shape data has been an area of research in disciplines such as computer vision, mechanical engineering, artifact searching, molecular biology, and chemistry [9]. Whereas the computer vision and graphics fields focus on shape matching, mechanical engineering and CAD fields have studied techniques involving comparisons of 3D CAD models using domain knowledge such as manufacturing processes and product information, as well as 3D shapes. Typical applications of similarity measurements in the fields of mechanical engineering and CAD are the retrieval of a desired 3D CAD model from a model database for design reuse [10], mass customization [11], and machining process planning [12].

The simplest way to compare 3D CAD models is to retrieve model names, model numbers, and annotations that are assigned to 3D CAD models, and to evaluate their similarities based on pre-defined evaluation metrics with those data [13]. However, this approach has the following limitations. Annotations regarding design and manufacturing are not attached to all 3D CAD models. Moreover, part names and codes are determined differently depending on the naming conventions employed. In addition, users would not know the exact part name or part code. Therefore, there is a need for a technique that evaluates similarities in 3D CAD models using 3D shape data because 3D shape data are always present in a 3D CAD model.

3D CAD models are classified into 3D CAD assembly models and part models. 3D CAD part models contain the shape information of a part or component. Meanwhile, 3D CAD assembly models are composed of 3D CAD part models. They contain only structural information, i.e., the position and orientation of 3D CAD part models and assembly constraints for different 3D CAD part models. Because product design is

*Corresponding author. Tel.: +82 54 530 1271, Fax.: +82 54 530 1278
E-mail address: dhmun@knu.ac.kr
This paper was presented at the ICMDT2017, Ramada Plaza Jeju Hotel, Jeju, Korea, April 19-22, 2017. Recommended by Guest Editor Dong-Gyu Ahn.

typically modeled in the form of 3D CAD assembly models, it is important to consider this application when comparing 3D CAD models.

In this paper, we propose a new shape-distribution-based holistic method for the comparison of 3D CAD assembly models. This method discriminates part-shape dissimilarities and assembly relationship dissimilarities, and it also evaluates two types of dissimilarity based on the shape-distribution-based method. The proposed method has two distinguishing technical features. First, it independently evaluates part-shape dissimilarities and assembly relationship dissimilarities, and produces an overall dissimilarity of 3D CAD assembly models. Second, it has very high applicability, and can be applied to most 3D CAD models (Feature-based, boundary representation (B-rep), and mesh models). Further, it does not require many types of data to perform the comparison.

The remainder of this paper is organized as follows. In Sec. 2, we review previous studies that were performed. In Sec. 3, we introduce shape-distribution-based dissimilarity measurements and propose improvement that can be made. In Sec. 4, we discuss the application of the proposed method in the comparison of 3D CAD assembly models. In Sec. 5, we then demonstrate the proposed method by performing experiments using test cases, while in Sec. 6, we present the Conclusion and future works.

2. Review of related studies and literature

In order for a comparison to be made, each 3D model has to be identified using a shape descriptor that provides a compact overall description of the 3D shape [14]. Similarity comparison methods can classify shapes into six categories according to the data used to generate the shape descriptor. These are global feature-based, manufacturing feature recognition-based, graph-based, 3D object recognition-based, histogram-based, and product information-based methods [9].

Paquet et al. [15] proposed a compact way of representing the coarse shape, scale, and composition properties of an object for 3D-shape matching. Ramesh et al. [16] presented a machining feature-based similarity comparison method for the retrieval of mechanical parts. El-Mehalawi and Miller [12] developed a representation scheme that created the attributed graph to compare CAD models of engineering parts. Horn [17] reported that Extended Gaussian images (EGI) are useful for representing the shapes of surfaces, and are thus appropriate for object recognition.

Hermann et al. [18] proposed a plan-based design similarity measure to explicitly exploit process-plan similarities and to improve the variant process planning approach. Rodríguez and Egenhofer [19] developed a method for semantic similarity measurements that was conducted at two levels: Measuring the similarities between elements themselves, and measuring the similarities of neighborhoods of the elements. Alizon et al. [20] proposed a knowledge-reuse procedure for part-similarity measurement. Mun et al. [21] enhanced the method proposed by Alizon et al. using ontology and multi-criteria decision-making techniques to improve the accuracy of the comparison results.

Deshmukh et al. [22] proposed the assembly comparison method based on assembly notation, part type, part characteristics, mating condition and joint orientation relationship. The Comparison method proposed by Chen et al. [23] uses hierarchical assembly structure, assembly interface composed of three layers and geometry of a part comprising an assembly. Both Refs. [22, 23] requires detailed information of assembly relationship to retrieve the intended assembly model in the database.

Histogram-based methods randomly select sample points from a 3D model and extract characteristics from those points to create a histogram or a distribution that represents the frequency of occurrence of each characteristic. Osada et al. [24] presented a method that represents the signature of an object as a shape distribution for the comparison of 3D mesh models. This method is simple and robust because it does not require pose registration, feature correspondence, or model fitting. Ohbuchi et al. [25] enhanced the method proposed by Osada et al. by introducing the angle-distance and absolute angle-distance histograms, which are calculated based on the D2 shape function. Tangelder and Veltkamp [26] used as shape descriptors weighted point sets that can be viewed as 3D probability distributions. Hwang et al. [11] presented shape-similarity measurements using Ray distance-to-surface (RDS) and Normalized ray distance-to-surface (NRDS). Several CAD model databases including Engineering shape benchmark [27] provide 3D CAD models that can be used as test cases for evaluating and comparing performances of various shape retrieval methods. They typically serve 3D CAD part models in the form of a mesh model.

The shape distribution-based method proposed by Osada et al. often results in a poor similarity comparison result when the shapes to be compared have similar overall shape but different shape characteristics [28]. Furthermore, as 3D CAD models become more complex, shape distributions tend to a bell-shaped normal distribution regardless of the level of detail in the shapes. In order to solve this problem, Ip et al. [28] proposed a method to create four shape distributions according to each of the point-pair categories, namely the All, In, Out and Mixed point pairs, and to use them as shape descriptors of a 3D CAD model. Cheng et al. [10] developed a new similarity measurement scheme that integrates the shape distribution and negative feature decomposition. Similarly, Chu and Hsu [29] adopted an integrated approach that uses the following multiple shape signatures: feature adjacency graphs, topological graphs, and shape distributions.

3. Shape distribution-based methods

In this study, we propose a new method for the comparison of 3D CAD assembly models based on shape distributions proposed in Refs. [24, 25]. In this section, we therefore briefly

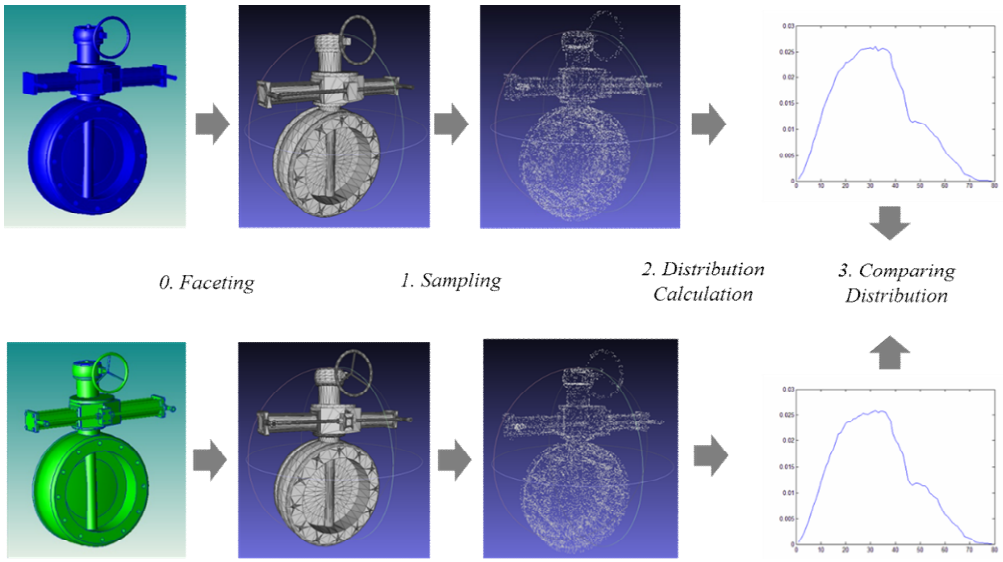

Fig. 1. Comparison procedure for shape distribution-based method.

introduce shape distribution-based comparison methods.

Shape distribution-based comparison has an advantage in that translation and rotation-invariant comparisons are possible, and do not require higher level information, including manufacturing features and processes. The dissimilarities of shapes are metrically derived from the distance between the Probability density functions (PDFs), which represent the probability of values obtained from points sampled from the surface of 3D models. In Ref. [24], five types of shape functions (A3, D1, D2, D3 and D4) were proposed to calculate values for the PDF.

The D2 shape function, which is the Euclidean distance from two points, was adopted in Refs. [24, 25], as well as other related studies because of its simplicity and robustness regardless of the shapes of the models to be compared. Fig. 1 illustrates the shape distribution-based comparison method. Using the sampling and distribution calculation process, the shape of a 3D model is represented as a single PDF. To compare two PDFs, we calculate the metric distance between two PDFs. The resulting distances between the two PDFs represent the dissimilarity between those two 3D models. As shown in Fig. 1, an additional faceting process is required before the sampling process when an input 3D model is not represented as a mesh model, such as Constructive solid geometry (CSG), B-rep, and feature-based models.

Some studies have proposed methods for improving the description power of the PDF by separating point pairs using topological information. This is done by decomposing volumes and by combining the two types of shape functions. Of those methods, combining two types of shape functions results in a two-dimensional (2D) PDF that enables a more precise shape comparison, while maintaining the advantage of generality because higher level information is not required. In Ref. [25], the mutual Angle-distance (AD) shape function is proposed to combine the Euclidean distance (D2) and the dot product of normal vectors of the point pair, resulting in a 2D PDF. Moreover, to overcome the issue of normal flip, another shape function called mutual Absolute-angle distance (AAD), which takes the absolute value of the dot product of normal vectors, was also proposed.

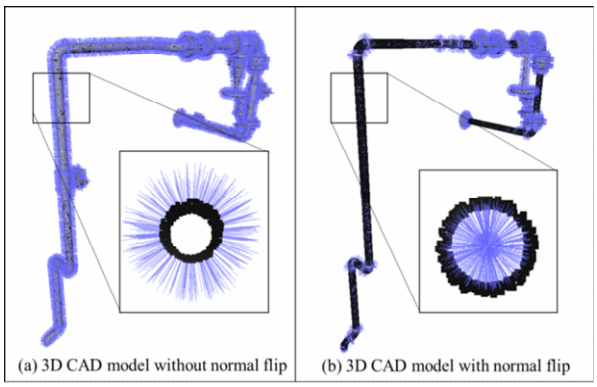

Fig. 2. Normal flip problem in 3D CAD models.

The normal flip problem is frequently observed when 3D models are retrieved from a public 3D model database in which the model quality is not guaranteed. However, the same problem also frequently occurs when a 3D CAD model generated from a 3D CAD system is converted into another format, as shown in Fig. 2(b). In this figure, the normal direction illustrated by the blue line is the opposite for two 3D CAD models. Therefore, we decided to apply the AAD shape function to robustly compare 3D CAD models.

4. Shape distribution-based comparison of 3D CAD assembly models

4.1 Introduction of 3D CAD assembly models

The assembly model is a model that consists of individual

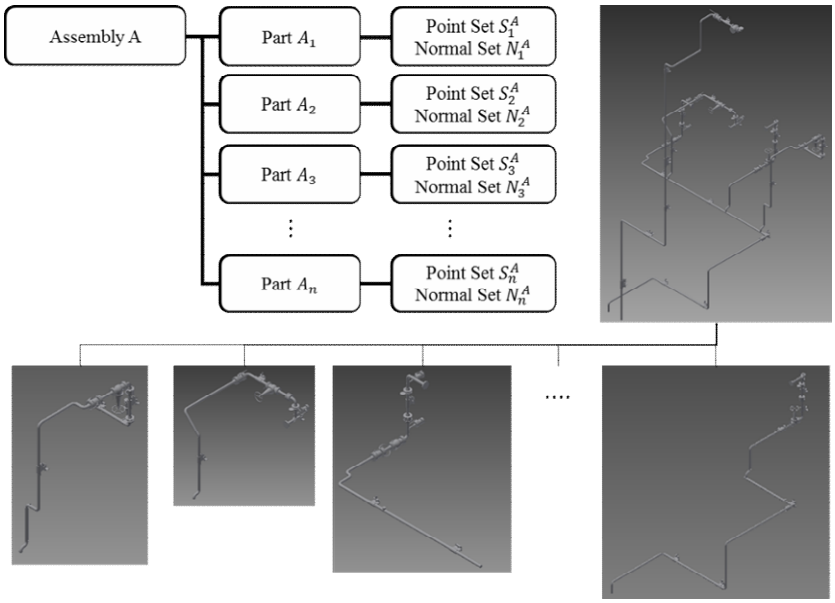

Fig. 3. Components of a 3D CAD assembly model.

part models, as shown in Fig. 3. The assembly model contains assembly information such as the positions and orientations of each part, as well as assembly constraints that exist between parts. Generally, products are designed primarily as 3D CAD assembly models. In addition to part models, assembly models are frequently reused when a new product is designed. Therefore, methods that compare 3D CAD assembly models are essential to ensure the efficiency of the reuse process. Shape comparisons are closely related to 3D model retrieval technology because we typically obtain similar models automatically by comparing a query 3D CAD assembly model with candidate 3D CAD assembly models that are in a database.

However, existing shape distribution-based methods have limitations with regards to their ability to compare 3D CAD assembly models. Because they do not consider the assembly relationship, the comparison process for assembly models is the same as that for part models. In other words, the assembly model is considered to be a single shape. Therefore, existing shape distribution-based methods provide insufficient information regarding dissimilarities that exist between assembly models because they cannot separately evaluate dissimilarities in the part shapes and dissimilarities in the assembly relationship between assembly models.

4.2 Holistic approach to the comparison of 3D CAD assembly models

In order to solve the problem, we developed a holistic approach to 3D CAD assembly comparison based on shape distributions. In our proposed approach, assembly retrieval method proposed by Deshmukh et al. [22] and Chen et al. [23] can be utilized to retrieve assembly models in the database with assembly structures similar to query assembly models using assembly name, annotation, part name and type. In this step, assembly structure only includes hierarchy between assembly and parts. Then part shape similarity and assembly relationship similarity between query and database assembly models are calculated by the proposed shape distribution-based comparison method. In the proposed method, assembly relationship represents position and orientation of part models comprising assembly model.

We focus on the limitation and the improvement strategy of the shape distribution-based comparison method. The limitation of the shape distribution-based comparison method is that the method is not able to reflect assembly relationship. Compared with Deshmukh et al. [22] and Chen et al. [23], the proposed method can compare assembly relationships of two assembly models in case the direct assembly structure information is not available. Our similarity measurement approach requires only geometrical information of assembly models.

Our goal is to derive separate dissimilarity values for the shapes of parts and the assembly relationship. 3D CAD assembly and part models have different data structure from each other [30-32]. Moreover, they are recorded in different files. In this context, dissimilarity should be separately measured for assembly relationship and part shapes of an assembly model. In the proposed method, the dissimilarity between two assembly models, D_{assy}, is defined in Eq. (1). In Eq. (1), D_a represents the dissimilarity of the assembly relationship between the assembly models, while D_p represents the dissimilarity of the part shapes between assembly models.

$$D_{assy} = \left[D_a, D_p \right]. \tag{1}$$

We can take advantage of separated similarity values for part shape and assembly relationship as follows. Well-known application of similarity measure such as 3D model retrieval

system can provide multidisciplinary retrieval results using separated similarity values. The user may request assembly model that shares similar part models assembled in different assembly relationship from query assembly model, which is the case of high D_a value and low D_p. Note that value of D_a and D_p is inversely proportional to the similarity between two models. Similarly, the condition of low D_a value and high D_p value can be assigned to retrieve assembly models with similar assembly relationship and dissimilar part shapes from query assembly model. The proposed approach requires no detailed assembly information extracted from 3D models in similarity measuring of part shape and assembly relationship. So the abovementioned advantage is available to the wide range of 3D model database.

4.2.1 Preprocessing

To compare assembly models, we obtain a point dataset S_i and a corresponding normal vector set N_i from each 3D CAD part model in the preprocessing step. We assume that the position set S_i and N_i of generated points is coordinates described in an assembly coordinate system. We applied the same method as that employed in Ref. [24] to obtain the point sampling set from the 3D surface model, which is represented as a triangular mesh. The sampling equation is shown in Eq. (2).

$$s = \left(1-\sqrt{r_1}\right)v_1 + \sqrt{r_1}\left(1-r_2\right)v_2 + \sqrt{r_1}\left(r_2 v_3\right). \tag{2}$$

In Eq. (2), s represents the coordinates of a randomly sampled point in the target mesh, v_1, v_2 and v_3 represent the vertex coordinates that are used to construct a target triangular mesh face, and r_1 and r_2 are random values between 0 and 1. By repeatedly selecting the target mesh and generating point samples using Eq. (2), we can obtain the intended number of point samples. The selection of the target mesh is adaptively performed using the weighted random sampling, where the weight of each mesh face is proportional to the area of the mesh face. As a result, point samples are uniformly distributed along the entire surface of the 3D model. The corresponding normal vectors for each point sample are also stored as-is, if the source model has face normal information, while the cross product of two vertex pairs is stored if the face normal is missing in the source model.

4.2.2 Dissimilarity measurement process for assembly relationship

Based on the fact that the assembly relationship can be obtained from the relative position and orientation of parts comprising an assembly, we propose a method to measure the dissimilarities of the assembly relationship based on shape distributions. Fig. 4 illustrates the proposed dissimilarity measurement process for the assembly relationship. Let two target assembly models for comparison–usually in retrieval case, one assembly model from database and the other from user input- be assembly A and assembly B. Based on our method explained in Sec. 4.2, we first found assembly models with similar assembly structures using the approach proposed in Refs. [22, 23] before shape distribution-based comparison. Therefore, we can assume that the number of comprising part of assembly A and assembly B are identical, and correspondence between parts among two assembly models are known. Let the number of comprising part of assemblies A and B be M, so that assembly A has $\{A_1, A_2, \ldots, A_M\}$ and B has $\{B_1, B_2, \ldots, B_M\}$ as its comprising part set. Without loss of generality and for the convenience of explanation, let part models with identical index be corresponding part models. Through the preprocessing described in Sec. 4.2.1, point set and normal vector set (S_m^A,N_m^A) and (S_m^B,N_m^B) –where m = *1…M*; of part A_m and B_m are obtained. The point set and normal vector set are used as input for the next step, which is divided into the sampling stage and dissimilarity comparison stage. In the sampling stage, we calculate the function values (*AADs*) for the assembly relationship for target assembly models assembly A and assembly B. In the dissimilarity computation stage, we convert the calculated values into PDFs, after which we calculate the distances between two PDFs to determine the metric dissimilarity of the assembly relationship D_a.

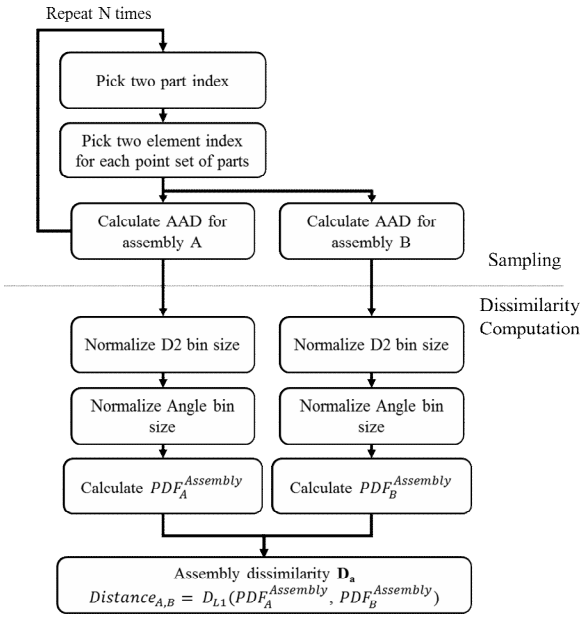

Fig. 4. Dissimilarity measurement process for assembly relationship.

4.2.2.1 Sampling stage for assembly relationship

In Fig. 4, N is the target number of samples. Then, we repeatedly performed the sampling process described below N times. First, we selected two indices of parts from 1 to M with uniformly distributed random sampling, where M is the number of parts comprising an assembly. Our method considers each part in the assembly to have an equal weight in the assembly relationship. We refer to the selected indices as i and j; part models A_i and A_j are selected from assembly A, while part models B_i and B_j are selected from assembly B. In the next

step, we select a single point from each of the four point sets S_i^A, S_j^A, S_i^B, S_j^B of part model A_i, A_j, B_i and B_j. Therefore, two point pairs are selected from assemblies A and B, where those points are belong to the corresponding parts between the two assemblies. Using the coordinates and normal vectors of one point pair for each assembly, we calculate the AAD shape function proposed in Ref. [25]. As we described earlier, the AAD shape function is for a 2D PDF, where the Euclidean distance and the magnitude of the dot product of the two vectors form two axes. We denote calculated value as *AAD*, that is, the 2-tuple value of *D2* and the *Angle*, as shown in Eq. (3).

$$AAD_p = D2_p,\ Angle_p \text{ p} \in [1, N] \tag{3}$$

where N is the number of intended samples.

D_2^p indicates the Euclidean distance of the point pair, and $Angle_p$ indicates the magnitude of the dot product of the corresponding normal vector pair. Note that the range of each $Angle_p$ is [0,1]. By performing N repetitions, we can obtain $2N$ number of *AADs* for assemblies A and B. Intuitively, obtained *AAD* value expresses difference of position and orientation of part i and j in an assembly model. As we described in Sec. 4.2.1, set of point position and normal vector is expressed in an assembly coordinate system. Thus, the distance between two points sampled in different part models encodes the distance between two part models, and the angle between two vectors encodes the orientation between two parts as distribution.

4.2.2.2 Dissimilarity computation stage for assembly relationship

From N *AADs*, a 2D PDF is generated for each assembly; two axes of the 2D PDF are the *D2* distribution and *Angle* distribution. Because the ranges of the *D2* and *Angle* values differ according to the model, a normalization process is required. In particular, for the *D2* value, the scale difference realized by directly comparing the models affects the range of the *D2* value. Therefore, a scale invariant comparison is enabled after normalization of the *D2* distribution and *Angle* distribution.

In existing shape distribution-based studies, each axis of a PDF is divided into a predefined number of bins for the normalization. In order to equalize the number of bins in each PDF, the interval length of the bin should be defined. To determine the interval length of the bin, methods using the maximum value of *D2* or the mean value of *D2* were proposed. We applied the normalization method based on maximum value. Number of bins in *D2* and *Angle* are fixed; thus intervals for each bin is decided by dividing maximum value of *D2* and *Angle* of each model by the number of bins. Normalization of PDFs allows scale-invariant shape comparison. In case of *Angle*, the result value of each sample is strictly limited in the range [0, 1] so it is reasonable to fix the number of bins. However, *D2* value can vary according to the model scale. If overall scales of two models differ, dissimilarity can be amplified due to the bias if adaptive normalization method is applied such as the method proposed in Ref. [10]. Therefore we decided to adopt maximum-based normalization method with fixed number of bins.

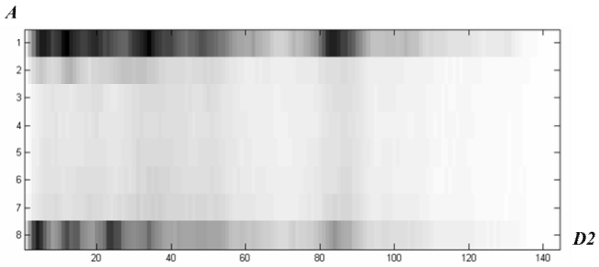

Fig. 5. Illustrated 2D PDF.

After normalization, the generation of the 2D PDF is straightforward. Fig. 5 shows the illustrated 2D PDF. Two axes of the 2D PDF are the *D2* distribution axis and the *Angle* distribution axis; the *D2* axis is divided into n^{D2} bins, and the *Angle* axis is divided into n^A bins. The brighter part indicates a low-probability bin, and the darker part indicates a high-probability bin. The probability of each bin is the number of *AADs* that belong to the bin divided by N, which is the total number of *AADs* for an assembly.

The final dissimilarity value of the assembly relationship is metrically calculated by the distance between the two PDFs of the models. In related studies, experiments with various distance metrics were performed including the Minkowski L_n norm, the earth mover's distance, and elastic matching. From our results and those reported from related studies, the type of distance metric does not significantly affect the resulting dissimilarity value. We applied the Minkowski L_1 norm, as in Eq. (4), because it is the simplest and fastest distance metric.

$$Distance_{L1}(X, Y) = \sum_{p=1}^{n^{D2}} \sum_{q=1}^{n^A} \left| x_{p,q} - y_{p,q} \right|. \tag{4}$$

In Eq. (4), X and Y indicate the 2D PDFs, where $x_{i,j}$ and $y_{i,j}$ are the probabilities for the p-th and q-th bins in the *D2* axis and *Angle* axis, respectively. n^{D2} and n^A indicate the number of bins in each axis. The resulting distance represents the dissimilarities in the assembly relationship D_a between two assembly models.

4.2.3 Dissimilarity measurement process for part shape

For dissimilarities of part shapes D_p, a similar process can be applied. The process diagram for the comparison of the part shapes is shown in Fig. 6. Unlike the process for the assembly relationship, only the point set and normal vector set for each individual part are used in the sampling process, while the dissimilarity computation process is similar to that of the assembly relationship.

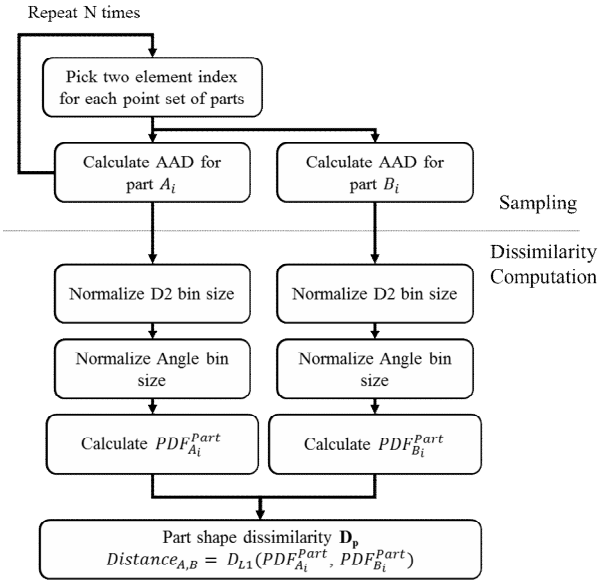

Fig. 6. Dissimilarity measurement process for part shape.

5. Experiments

5.1 Test models for experiments

To test and validate our proposed method, we performed several experiments where we measured the dissimilarities between the test 3D CAD assembly models. Four base assembly models *A*, *B*, *C* and *D* provided by the company "*D*", one of the largest shipbuilding companies in the world, were prepared for experiments.

As shown in Figs. 7 and 8, assembly models *A* and *B* consist of 5 part models whereas assembly models *C* and *D* consist of 7 part models. Assembly models *A* and *C* were created using Aveva Marine. Assembly models *B* and *D* created using Intergraph SmartMarine 3D were remodelled versions of assembly models *A* and *C,* respectively. Each assembly model represents the 3D design of a unit object for an offshore plant. Each part model represents the 3D design of a branch object. A unit object consists of branch objects.

Assembly models *A* and *C* (Assembly models *B* and *D*) have different level of detail of their parts, as shown in Fig. 9; assembly models *B* and *D* have simpler part shapes than assembly models *A* and *C*. A flipped normal was observed in several part models of assembly models *B* and *D*; surfaces with a flipped normal are rendered in dark color in Figs. 7 and 8. Scales of the original assembly models (*A* and *C*) and the remodelled assembly models (*B* and *D*) are different; the remodelled models are approximately 1/1000 of the original models.

From the four base assembly models, ten variation assembly models were generated by changing assembly relationship of assembly models, as shown in Figs. 10 and 11. The variations of assembly relationship are differences in orientation, position, position and orientation. Test assembly models *B-Base* and *C-Base* are the same as assembly models *B* and *C*.

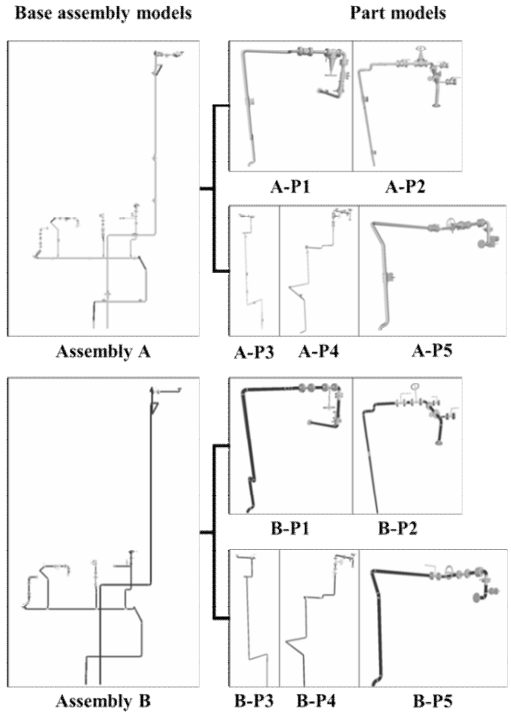

Fig. 7. 3D CAD base assembly models *A* and *B* prepared for experiments.

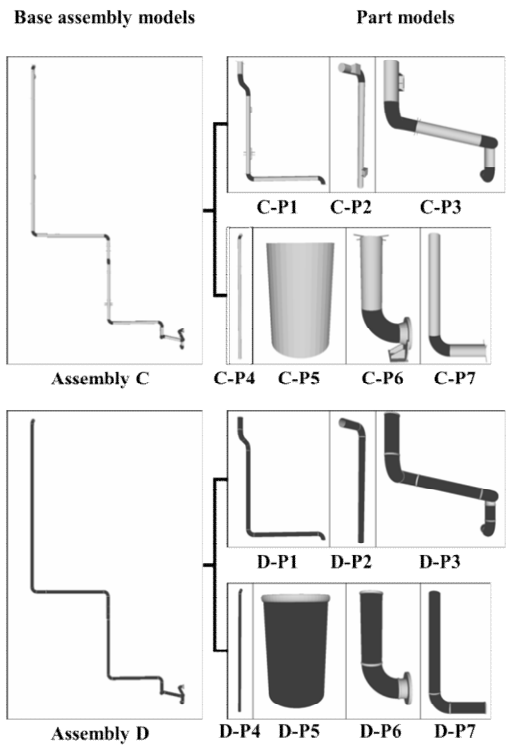

Fig. 8. 3D CAD base assembly models *C* and *D* prepared for experiments.

Test assembly models *A-Base* and *D-Base* are the same as assembly models *A* and *D*. Test assembly models *A-Orien* and

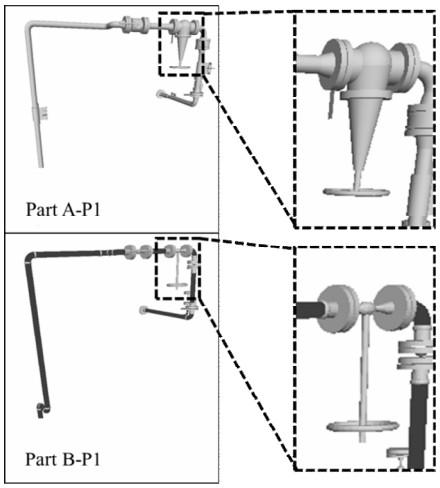

Fig. 9. Differences in two part models between assembly models.

D-Orien contain a part (Part models *A-P5* and *D-P7*) that has a different orientation compared to test assembly models *A-Base* and *D-Base*. Test assembly models *A-Pos* and *D-Pos* contain a part (Part models *A-P1* and *D-P7*) that has a different position compared to test assembly models *A-Base* and *D-Base*. Finally, test assembly models *A-Pos_Orien* and *D-Pos_Orien* contains a part (Part models *A-P5* and *D-P7*) that has a different position and orientation compared to test assembly models *A-Base* and *D-Base*.

5.2 Test cases and experiment settings

We defined five test cases and conducted experiments to measure the dissimilarity between assemblies for the five test cases. The test cases are summarized in Table 1. In the first test case, we measured the dissimilarity between test assembly models *A-Base* and *B-Base,* and test assembly models *D-Base* and *C-Base;* two assemblies to be compared have an identical assembly relationship but different shapes for each part model. Test cases 2-4 measure the dissimilarity of test assembly models *A-Base* and *D-Base* with test assembly models *A-Orien*, *A-Pos* and *A-Pos_Orien* and *D-Orien*, *D-Pos* and *D-Pos_Orien*, respectively. In these cases, the shapes of the parts comprising an assembly are identical, but the position and orientation of the parts are different according to the changes in the assembly relationship. Finally, in test case 5, we measured the dissimilarity of test assembly models *A-Orien* and *B-Base* with test assembly models *D-Orien* and *C-Base*; two assembly models to be compared have different part shapes as well as different assembly relationship.

The parameters determined for the experiments that measure the dissimilarity between test assemblies are as follows. Parameters are selected based on the Refs. [24, 25] with consideration of performance and computation cost. In case of sample number(*N*), higher number results in accurate similarity with higher computational cost. Target 3D models of our research are 3D CAD models with relatively complex shapes. Thus, we set a higher sample number similar to Ref. [24].

Table 1. Five test cases for experiments.

	Dissimilarity between		*Difference*
	Assembly A and B	*Assembly C and D*	
Test case 1	A-Base & B-Base	D-Base & C-Base	Shape of parts
Test case 2	A-Base & A-Orien	D-Base & D-Orien	Assembly relation (By orientation)
Test case 3	A-Base & A-Pos	D-Base & D-Pos	Assembly relation (By position)
Test case 4	A-Base & A-Pos_Orien	D-Base & D-Pos_Orien	Assembly relation (By position and orientation)
Test case 5	A-Orien & B-Base	D-Orien & C-Base	Shape of parts and assembly relation (By orientation)

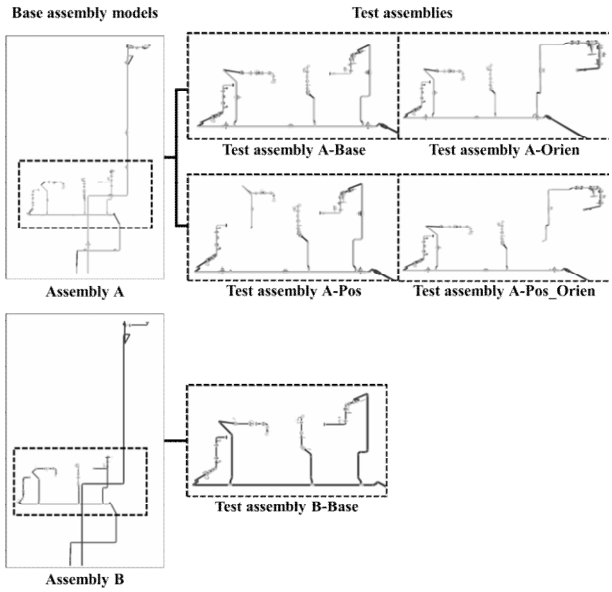

Fig. 10. Generation of five variation assembly models from two base assembly models A and B for experiments.

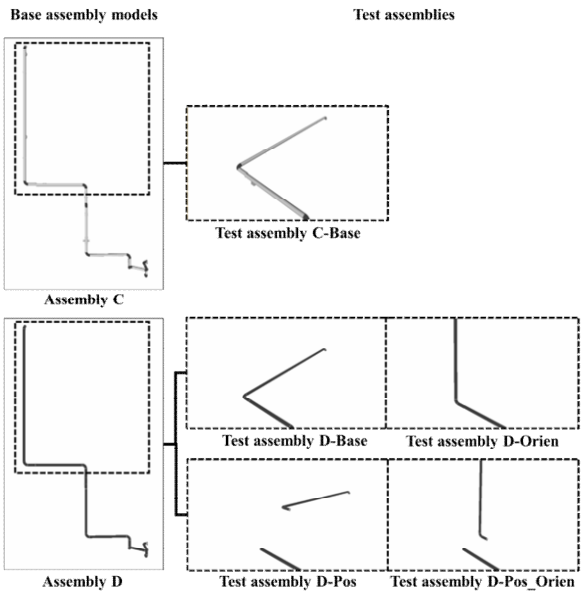

Fig. 11. Generation of five variation assembly models from two base assembly models C and D for experiments.

Table 2. Measured dissimilarities of part models in test assembly models *A-Base*, *A-Orien*, *A-Pos*, *A-Pos_Orien* and *B-Base*.

		Test assembly A-Base					*Test assembly A-Orien*				
		A-P1	A-P2	A-P3	A-P4	A-P5	A-P1	A-P2	A-P3	A-P4	A-P5
		Test case 2									
Test assembly A-Orien	A-P1	**0.026817**	0.295488	0.486723	0.376917	0.369459					
	A-P2	0.298677	**0.026478**	0.430237	0.398092	0.326891					
	A-P3	0.485092	0.428581	**0.025003**	0.433872	0.48316					
	A-P4	0.376133	0.392714	0.433826	**0.024851**	0.346939					
	A-P5	0.376165	0.323179	0.485088	0.351265	**0.029676**					
		Test case 3									
Test assembly A-Pos	A-P1	**0.025391**	0.289764	0.477659	0.381561	0.373371					
	A-P2	0.295834	**0.036888**	0.427795	0.394276	0.323132					
	A-P3	0.479391	0.427839	**0.032553**	0.433369	0.489456					
	A-P4	0.374956	0.39447	0.428421	**0.029833**	0.350002					
	A-P5	0.366816	0.322149	0.484486	0.350161	**0.029121**					
		Test case 4									
Test assembly A-Pos_Orien	A-P1	**0.027803**	0.295343	0.480801	0.376528	0.374069					
	A-P2	0.293098	**0.029657**	0.423988	0.394936	0.325615					
	A-P3	0.483814	0.427023	**0.029243**	0.430727	0.487049					
	A-P4	0.374504	0.394852	0.43396	**0.02669**	0.349745					
	A-P5	0.370993	0.323441	0.488291	0.347511	**0.026791**					
		Test case 1					**Test case 5**				
Test Assembly B-Base	B-P1	**0.296032**	0.393471	0.551767	0.445578	0.422449	**0.292643**	0.400885	0.556381	0.444731	0.421358
	B-P2	0.371014	**0.296165**	0.5431	0.519394	0.453199	0.372257	**0.307947**	0.550291	0.520166	0.452797
	B-P3	0.494551	0.453352	**0.331026**	0.418934	0.498888	0.498011	0.45639	**0.341726**	0.420874	0.501524
	B-P4	0.394224	0.391735	0.457687	**0.233049**	0.392269	0.394012	0.394629	0.46364	**0.238201**	0.392899
	B-P5	0.499901	0.432144	0.557121	0.524916	**0.338299**	0.490332	0.431601	0.558573	0.518276	**0.338028**

Table 3. Measured dissimilarities of part models in test assembly models *C-Base*, *D-Base*, *D-Orien*, *D-Pos,* and *D-Pos_Orien*.

		Test assembly D-Base							*Test assembly D-Orien*						
		D-P1	D-P2	D-P3	D-P4	D-P5	D-P6	D-P7	D-P1	D-P2	D-P3	D-P4	D-P5	D-P6	D-P7
		Test case 2													
Test assembly D-Orien	D-P1	**0.036255**	0.255255	0.489885	0.448797	0.847824	0.548336	0.374199							
	D-P2	0.244972	**0.040691**	0.351061	0.304085	0.806921	0.455463	0.210308							
	D-P3	0.477859	0.342482	**0.0396**	0.287466	0.786642	0.438213	0.270668							
	D-P4	0.436972	0.309111	0.299484	**0.034746**	0.738026	0.508183	0.308432							
	D-P5	0.84481	0.804663	0.799662	0.740088	**0.029299**	0.821268	0.792379							
	D-P6	0.536682	0.445564	0.425251	0.503349	0.805624	**0.055714**	0.365398							
	D-P7	0.367668	0.218443	0.276628	0.308695	0.789169	0.378481	**0.035553**							
		Test case 3													
Test assembly D-Pos	D-P1	**0.034435**	0.254848	0.48736	0.445215	0.849613	0.545902	0.371561							
	D-P2	0.244972	**0.040691**	0.351061	0.304085	0.806921	0.455463	0.210308							
	D-P3	0.477859	0.342482	**0.0396**	0.287466	0.786642	0.438213	0.270668							
	D-P4	0.436972	0.309111	0.299484	**0.034746**	0.738026	0.508183	0.308432							
	D-P5	0.84481	0.804663	0.799662	0.740088	**0.029299**	0.821268	0.792379							
	D-P6	0.536682	0.445564	0.425251	0.503349	0.805624	**0.055714**	0.365398							
	D-P7	0.366352	0.215647	0.281975	0.314924	0.793846	0.372335	**0.037478**							
		Test case 4													
Test assembly D-Pos_Orien	D-P1	**0.034277**	0.254211	0.48699	0.443724	0.846342	0.547413	0.372574							
	D-P2	0.244972	**0.040691**	0.351061	0.304085	0.806921	0.455463	0.210308							
	D-P3	0.477859	0.342482	**0.0396**	0.287466	0.786642	0.438213	0.270668							
	D-P4	0.436972	0.309111	0.299484	**0.034746**	0.738026	0.508183	0.308432							
	D-P5	0.84481	0.804663	0.799662	0.740088	**0.029299**	0.821268	0.792379							
	D-P6	0.536682	0.445564	0.425251	0.503349	0.805624	**0.055714**	0.365398							
	D-P7	0.371189	0.217941	0.279247	0.308893	0.789028	0.373638	**0.033623**							
		Test case 1							**Test case 5**						
Test assembly C-Base	C-P1	**0.24568**	0.315859	0.251966	0.421801	0.844261	0.476021	0.299902	**0.240917**	0.318323	0.253466	0.424998	0.846302	0.476318	0.30349
	C-P2	0.311922	**0.220812**	0.308796	0.231359	0.929188	0.435896	0.195723	0.310671	**0.217331**	0.305712	0.231744	0.924604	0.434362	0.191208
	C-P3	0.275146	0.389082	**0.206182**	0.461411	0.851469	0.420271	0.328337	0.271072	0.381275	**0.195032**	0.456207	0.854916	0.42009	0.319519
	C-P4	0.426134	0.388559	0.465536	**0.051003**	1.019928	0.565662	0.336924	0.437571	0.404972	0.473419	**0.052821**	1.028971	0.570995	0.344419
	C-P5	0.768791	0.824884	0.775309	0.843287	**0.666128**	0.843185	0.750645	0.776104	0.823517	0.781803	0.846401	**0.662598**	0.848377	0.753239
	C-P6	0.413349	0.484543	0.454355	0.531414	0.897959	**0.24185**	0.415058	0.40365	0.486822	0.442074	0.523279	0.892553	**0.221672**	0.394123
	C-P7	0.275715	0.309711	0.260309	0.352152	0.903734	0.377239	**0.105598**	0.272039	0.318861	0.262136	0.353225	0.897839	0.375015	**0.10601**

We conducted experiments with different parameter settings for the number of bins in *D2* with {32, 64, 128} and *Angle* with {4, 8, 16} according to the Ref. [25]. In these parameter settings for the number of bins, we observed no significant changes such as changes in order of similarities between test cases. Therefore we decided to report the experimental results from parameter setting 64 and 8 in the number of histogram bin for *D2* and *Angle*, respectively.

In summary, from each 3D CAD part model, 10000 points and normal vectors are generated for the point set and normal vector set. The sampling number, *N*, is 1024 times 1024 for PDF generation. As mentioned in Sec. 4, the numbers of bins for the *D2* and *Angle* axis are fixed as 64 and 8 respectively in our experiments. To compare our results with those of an existing study, we also developed a comparison method proposed in Ref. [25], while the same parameter is used for the point set, normal vector set, sampling number, and number of bins.

5.3 Experiment results

Table 2 shows the measured dissimilarities between two part models included in test assembly models *A-Base*, *A-Orien*, *A-Pos*, *A-Pos_Orien* and *B-Base* used in the test cases 1-5. In Table 2, the column-wise lowest dissimilarity values are shown in bold text. Because assembly models (Test assembly models *A-Base*, *A-Orien*, *A-Pos* and *A-Pos_Orien*) have same part shapes in test cases 2-4, the resulting dissimilarity of two part models corresponding to each other is close to zero. Because assembly models with different part shapes are compared in test cases 1 and 5 (Test assembly models *A-Base* versus *B-Base*, test assembly models *A-Orien* versus *B-Base*), the resulting dissimilarity of two part models corresponding to each other is higher compared with other test cases. However, in all the test cases, a part model have the lowest dissimilarity with its corresponding part model among all part models to be compared.

Table 3 shows the measured dissimilarities between two part models included in test assembly models *D-Base*, *D-Orien*, *D-Pos*, *D-Pos_Orien* and *C-Base* used in the test cases 1-5. In Table 3, the column-wise lowest dissimilarity values are shown in bold text. Table 3 shows the same trend as Table 2.

The dissimilarity values measured by our proposed method and that of Ohbuchi et al. [25] for ten test cases are summarized in Tables 4 and 5. In these tables, the dissimilarity of the part shape in the proposed method is determined by averaging the dissimilarities of comprising parts; in each test case, the dissimilarity of the part shape is measured between each part of a test assembly and its corresponding part of another test assembly. In this table, we measured the dissimilarity values using the method proposed by Ohbuchi et al. [25] by treating an assembly model as a single shape (Single part).

In the case of the part-shape dissimilarity, test cases 1 and 5 show high dissimilarities, while test cases 2-4 show lower dissimilarities. In the case of the assembly relationship dissimilarity, test case 1 shows the minimum dissimilarity value because two assembly models in this test case have the same assembly relationship. From these results, we found that the proposed method efficiently measures the dissimilarity between part shapes and the dissimilarity in the assembly relationship.

When we compare our proposed method with the existing method [25], the main advantage of our method is that it provides separate dissimilarity values for the shapes of parts and the assembly relationship, and thus provides more detailed information on the dissimilarity when we compare 3D CAD assembly models. For example, dissimilarities measured by the method proposed by Ohbuchi et al. [25] are similar in test cases 1 and 4; test cases 1 and 4 show a 0.03 and 0.01 difference in the dissimilarity in Tables 4 and 5, respectively. However, there are very different reasons for the occurrence of the dissimilarity between the test assemblies in the two test cases. In test case 1, the dissimilarity is due primarily to the difference in the part shapes, while in test case 3, it is primarily due to the difference in the assembly relationship. Existing distribution-based methods cannot distinguish between test cases 1 and 3, while the proposed method can do so.

The second advantage of the proposed method is that it is simple and robust by virtue of adopting a distribution-based method for the comparison of 3D CAD assembly models. The proposed method can be applied to almost all types of 3D CAD models, whether they are feature-based, CSG, B-rep, or mesh models. In addition, the proposed method does not require any higher level information for the comparison operation to be performed.

In calculating D_p and D_a of test cases, 143 and 200 milliseconds elapsed in average. We observed that point indexing for each part model consumed additional time when calculating D_a. In application such as shape retrieval system, PDFs corresponding to the 3D models in the database are generated once as an off-line process before search operation. In on-line process when query model input is given, PDF for query model is generated and distance calculation between PDFs in the database and query model is performed to measure similarity. In our experiments, distance calculation between two PDFs is done in 0.01 millisecond. Therefore rapid retrieval process is possible with shape distribution-based comparison method.

However, the proposed method has the following limitations. In the proposed method, dissimilarity results of the assembly relationship and part shape are not completely independent from each other. In the results of test case 1 shown in Tables 4 and 5, the dissimilarity of the assembly relationship was measured not as close to zero but as 0.134392 and 0.172878 even though two assembly models had the same assembly relationship. This is because points and normal vectors used for the assembly relationship were also sampled from each of the parts, which have different shapes in this test case. We can observe the same effect by comparing the results of test case 2 and test case 5 in Table 4. In these test cases, difference in assembly relationship are the same between two

Table 4. Measured dissimilarities in five test cases for models *A-Base*, *A-Orien*, *A-Pos*, *A-Pos_Orien* and *B-Base* obtained using the proposed method and Ohbuchi's method.

	Test assemblies	Dissimilarity		
		Proposed Method		Ohbuchi et al. [5]
		Part shape (D_p)	Assembly relationship (D_a)	Assembly overall
Test case 1	A-Base and B-Base	0.298914	0.134392	0.23547
Test case 2	A-Base and A-Orien	0.026565	0.237392	0.164288
Test case 3	A-Base and A-Pos	0.030757	0.267134	0.176373
Test case 4	A-Base and A-Pos_Orien	0.028037	0.287998	0.205187
Test case 5	A-Orien and B-Base	0.303709	0.303616	0.326914

Table 5. Measured dissimilarities in five test cases for models *D-Base*, *D-Orien*, *D-Pos*, *D-Pos_Orien* and *C-Base* obtained using the proposed method and Ohbuchi's method.

	Test assemblies	Dissimilarity		
		Proposed Method		Ohbuchi et al. [5]
		Part shape (D_p)	Assembly relationship (D_a)	Assembly overall
Test case 1	D-Base and C-Base	0.248179	0.172878	0.1209383
Test case 2	D-Base and D-Orien	0.038837	0.235651	0.142684
Test case 3	D-Base and D-Pos	0.038852	0.257597	0.1274614
Test case 4	D-Base and D-Pos_Orien	0.038279	0.297234	0.1330366
Test case 5	D-Orien and C-Base	0.242340	0.209667	0.1712484

assembly models to be compared (test assembly model *A-Base* versus *A-Orien* and test assembly model *B-Base* versus *A-Orien*). It means that assembly relationship dissimilarities calculated in these test cases are similar to each other. However, resulting dissimilarity of assembly relationships differ between test cases 2 and 5 because the difference of part shape affects dissimilarity measure of assembly relationship in test case 5. Nevertheless, the relative differences in the dissimilarity for the assembly relationship among the test cases remain valid; assembly relationship dissimilarity value calculated between assembly models with the same assembly relationship is lower than that between assembly models with the different assembly relationship. This guarantees the suitability of the proposed method for engineering applications such as 3D CAD assembly retrieval.

6. Conclusion

In this paper, we proposed a new shape distribution-based holistic method for the comparison of 3D CAD assembly models. Contrary to existing methods, this method discriminates part-shape dissimilarities and assembly relationship dissimilarities, and evaluates two types of dissimilarities based on the shape distribution-based method. The separation of the part-shape dissimilarity and the assembly relationship dissimilarity enables the analysis of detailed dissimilarity characteristics between assembly models.

To test and validate our proposed method, we performed several experiments where we measured dissimilarities between test 3D CAD assembly models. From the experiments, we concluded that our method can provide an improved description of dissimilarities in assembly models compared with existing methods that process an entire assembly model as a single shape (Single part).

However, the proposed method also has two limitations. First, there is a dependency between the part-shape dissimilarity and the assembly relationship dissimilarity. Second, intuitive differences in the assembly relationship are not directly reflected by the dissimilarity value of the assembly relationship. In future study, we will aim to find solutions for these limitations.

Acknowledgment

This research was supported by the Industrial Core Technology Development Program (Project ID: 10048341) funded by the Ministry of Trade, Industry and Energy and by the Civil-Military Program (Project ID: CMP-16-01-KIST) of National Research Council of Science and Technology of the Korean government. The authors gratefully acknowledge their support.

References

[1] S. Yoo, K. Park and S. Y. Choi, The vulnerability assessment of ground combat vehicles using target functional modeling and FTA, *International Journal of Precision Engineering and Manufacturing*, 17 (5) (2016) 651-658.

[2] G. Li, Z. Wang and A. Kubo, The modeling approach of digital real tooth surfaces of hypoid gears based on non-geometric-feature segmentation and interpolation algorithm, *International Journal of Precision Engineering and Manufacturing*, 17 (3) (2016) 281-292.

[3] D. J. Yoo, New paradigms in cellular material design and fabrication, *International Journal of Precision Engineering and Manufacturing*, 16 (12) (2015) 2577-2589.

[4] Q. Meng, F. Li, L. Zhou, J. Li, Q. Ji and X. Yang, A rapid life cycle assessment method based on green features in supporting conceptual design, *International Journal of Precision Engineering and Manufacturing-Green Technology*, 2 (2) (2015) 189-196.

[5] P. Bo, G. Luo and K. Wang, A graph-based method for fitting planar B-spline curves with intersections, *Journal of Computational Design and Engineering*, 3 (1) (2016) 14-23.

[6] B. C. Kim and D. Mun, Enhanced volume decomposition minimizing overlapping volumes for the recognition of design features, *Journal of Mechanical Science and Technology*, 29 (12) (2015) 5289-5298.

[7] J. Li, D. Mun and S. Han, Profile-based feature representation method and its application in data exchange from mechanical CAD systems to ship CAD systems, *Journal of Mechanical Science and Technology*, 30 (12) (2016) 5641-5649.

[8] D. G. Kendall, The diffusion of shape, *Advances in Applied Probability*, 9 (3) (1977) 428-430.

[9] N. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman and K. Ramani, Three-dimensional shape searching: state-of-the-art review and future trends, *Computer-Aided Design*, 37 (2005) 509-530.

[10] H. Cheng, C. Lo, C. Chu and Y. Kim, Shape similarity

measurement for 3D mechanical part using D2 shape distribution and negative feature decomposition, *Computers in Industry*, 62 (3) (2011) 269-280.
[11] T. J. Hwang, K. Lee, H. Y. Oh and J. H. Jeong, Shape similarity measurement using ray distances for mass customization, *Proc. of ACM Symposium on Solid Modeling and Applications* (2004) 279-284.
[12] M. El-Mehalawi and R. A. Miller, A database system of mechanical components based on geometric and topological similarity. Part I: representation, *Computer-Aided Design*, 35 (1) (2003) 83-94.
[13] L. Zehtaban, O. Elazhary and D. Roller, A framework for similarity recognition of CAD models, *Journal of Computational Design and Engineering*, 3 (3) (2016) 274-285.
[14] J. W. Tangelder and R. C. Veltkamp, A survey of content based 3D shape retrieval methods, *Multimedia Tools and Applications*, 39 (3) (2008) 441-471.
[15] E. Paquet, M. Rioux, A. Murching, T. Naveen and A. Tabatabai, Description of shape information for 2-D and 3-D objects, *Signal Processing: Image Communication*, 16 (1) (2000) 103-122.
[16] M. Ramesh, D. Yip-Hoi and D. Dutta, Feature based shape similarity measurement for retrieval of mechanical parts, *Journal of Computing and Information Science in Engineering*, 1 (3) (2001) 245-256.
[17] B. K. P. Horn, Extended Gaussian images, *Proceedings of the IEEE*, 72 (12) (1984) 1671-1686.
[18] J. W. Herrmann and G. Singh, *Design similarity measures for process planning and design evaluation*, Maryland Univ. College Park Dept of Mechanical Engineering (1997).
[19] M. A. Rodríguez and M. J. Egenhofer, Determining semantic similarity among entity classes from different ontologies, *IEEE Transactions on Knowledge and Data Engineering*, 15 (2) (2003) 442-456.
[20] F. Alizon, S. B. Shooter and T. W. Simpson, Reuse of manufacturing knowledge to facilitate platform-based product realization, *Journal of Computing and Information Science in Engineering*, 6 (2) (2006) 170-178.
[21] D. Mun and K. Ramani, Knowledge-based part similarity measurement utilizing ontology and multi-criteria decision making technique, *Advanced Engineering Informatics*, 25 (2) (2011) 119-130.
[22] A. S. Deshmukh, A. G. Banerjee, S. K. Gupta and R. D. Sriram, Content-based assembly search: A step towards assembly reuse, *Computer-Aided Design*, 40 (2) (2008) 244-261.
[23] X. Chen, S. Gao, S. Guo and J. Bai, A flexible assembly retrieval approach for model reuse, *Computer-Aided Design*, 44 (6) (2012) 554-574.
[24] R. Osada, T. Funkhouse, B. Chazelle and D. Dobkin, Shape distributions, *ACM Transactions on Graphics*, 21 (4) (2002) 807-832.
[25] R. Ohbuchi, T. Minamitani and T. Takei, Shape-similarity search of 3D models by using enhanced shape functions, *International Journal of Computer Applications in Technology*, 23 (2-4) (2005) 70-85.
[26] J. W. Tangelder and R. C. Veltkamp, Polyhedral model retrieval using weighted point sets, *International Journal of Image and Graphics*, 3 (1) (2003) 209-229.
[27] S. Jayanti, Y. Kalyanaraman, N. Iyer and K. Ramani, Developing an engineering shape benchmark for 3D CAD models, *Computer-Aided Design*, 38 (9) (2006) 939-953.
[28] C. Y. Ip, D. Lapadat, L. Sieger and W. C. Regli, Using shape distributions to compare solid models, *Proc. of ACM Symposium on Solid Modeling and Applications* (2002) 273-280.
[29] C. H. Chu and Y. C. Hsu, Similarity assessment of 3D mechanical components for design reuse, *Robotics and Computer-Integrated Manufacturing*, 22 (4) (2006) 332-341.
[30] ISO, *ISO 10303-203:1994*, Industrial automation systems and integration — Product data representation and exchange — Part 203: Application protocol: Configuration controlled 3D designs of mechanical parts and assemblies, International Organization for Standardization (ISO), Geneva, Switzerland (1994).
[31] B. C. Kim, D. Mun and S. Han, Web Service with parallel processing capabilities for the retrieval of CAD assembly data, *Concurrent Engineering: Research and Applications*, 19 (1) (2011) 5-18.
[32] I. Kim, J. Lee, B. C. Kim, J. Hwang, C. H. Lim and D. Mun, Verification of product design using regulation knowledge-base and web service, *Journal of Mechanical Science and Technology*, 29 (12) (2015) 5113-5119.

Hyungki Kim is a Postdoctoral Researcher in the Department of System Reliability at Korea Institute of Machinery & Materials (KIMM), Daejeon, Korea. He received his B.S. degree from the Department of Mathematical Science at KAIST in 2009 and M.S. degree from the Department of Mechanical Engineering at KAIST in 2011 and the Ph.D. degree in Department of Mechanical Engineering at KAIST in 2015. His current research interests include computer-aided design, 3D urban modeling, computer vison, computer graphics and real-time visualization

Duhwan Mun is an Associate Professor of Precision Mechanical Engineering of Kyungpook National University. He received a B.S. in Mechanical Engineering from Korea University; an M.S. and Ph.D. in Mechanical Engineering from Korea Advanced Institute of Science and Technology (KAIST). His research interests include computer-aided design, industrial data standards for product data exchange, product lifecycle management, knowledge-based engineering, and virtual reality for engineering applications.