Robotics and Computer–Integrated Manufacturing 51 (2018) 103–111

Contents lists available at ScienceDirect

Robotics and Computer–Integrated Manufacturing

journal homepage: www.elsevier.com/locate/rcim

An efficient assembly retrieval method based on Hausdorff distance

Jie Zhang*, Jiazhen Pang, Jianfeng Yu, Pan Wang

Northwestern Polytechnical University, Xi'an, China

ARTICLE INFO

Keywords:
Assembly retrieval
Hausdorff distance
Assembly model
Shape information

ABSTRACT

An assembly provides significant knowledge in design reuse, since it can represent parts information and their relationships comprehensively. The retrieval of assembly has the practical value in accelerating design and enhancing efficiency. Our previous work addressed this issue by using a method based on earth mover's distance. However, it requires a longer, less favorable retrieval time to obtain the retrieval results. To solve this problem, we propose an efficient assembly retrieval method based on a modified Hausdorff distance (MHD). By employing shape distributions, each part is quantitatively described as a point and an assembly is turned into a set of points, then the MHD is performed to evaluate the dissimilarity between the generated point sets. The MHD could be used in overall retrieval of assembly, since it considers the contributions of similar and dissimilar parts. For the scenario of flexible retrieval, a directed Hausdorff distance is proposed, which only compares the local dissimilarity. Experiments are carried out to demonstrate the accuracy and efficiency of the MHD, compared with the earth mover's distance and the vector space model. Our study reveals that the proposed method can retrieve relevant assemblies efficiently on the premise of assuring the accuracy.

1. Introduction

With the application of computer aided design, enterprises have accumulated a large number of 3D digital product models in CAD library. These numerous design models convey important knowledge and experiences, which are worth reusing to improve design efficiency and quality.

In the engineering field, an assembly is one of the most common kinds of product models. Compared with a single mechanical part, an assembly is a collection of parts, which represents not only the shape information but also relationships between parts, with more worthy and significant embedded knowledge to be reused in the design [1]. Therefore, an efficient retrieval method to locate assembly models with reference value has significance in design.

The investigation in the enterprise shows that in the design of a new product, there are two scenarios for assembly retrieval: (1). When an assembly digital model has been developed, a designer wants to find assemblies with similar shape, then he could retrieve models with overall similarity to discover, analyze, and refer the reusable information of material, process, simulation, etc. (2). When a designer wants to know which assemblies contain certain existing parts or components, then he could query the flexible combination of parts or components to retrieve the assemblies which include the similar combination. Then the designer could study the application of the parts or components, and learn the design experience and knowledge from the retrieval results. We define two modes of retrieval in the above scenarios as overall retrieval and flexible retrieval, and they satisfy the most common requirements for model resource location.

Lately, some assembly retrieval methods have been carried out [2–8] which mainly focus on expressing part shapes, connection relationships, and high-level semantic information synthetically. Nevertheless, it is difficult to normalize these diverse kinds of information, and the more kinds of information are constrained, the fewer scenarios of retrieval could be applied on, which may lead to loss of useful retrieval results. The part-shape information is a basic unit of representation for structure, and we intend to propose a unified method for above mentioned scenarios based on part shapes. In the research of assembly retrieval based on part-shape information, Wang et al. [2] proposed an assembly retrieval method based on shape distributions and earth mover's distance (EMD). It turned a part model into a point in space, and an assembly model into a set of spatial points, finally achieving assembly matching through an EMD algorithm. However, it spends much time computing assembly similarity through the EMD algorithm in practical application. When the model library covers more assemblies, a longer, less favorable retrieval time is to be expected. In this paper, we analyze the retrieval method and propose an efficient assembly retrieval approach based on a Hausdorff distance.

* Corresponding author.
E-mail addresses: zhangjie@nwpu.edu.cn (J. Zhang), pangjiazhen@mail.nwpu.edu.cn (J. Pang), yujf@nwpu.edu.cn (J. Yu), wpda2003@mail.nwpu.edu.cn (P. Wang).

https://doi.org/10.1016/j.rcim.2017.11.012
Received 21 March 2017; Received in revised form 13 November 2017; Accepted 14 November 2017
0736-5845/

Taking the method by Wang et al., this research changes the assembly into a point set revolving around shape distribution in space, and accordingly, we transform assembly matching into point set matching. We then use a modified Hausdorff distance to match point sets. It is an effective many-to-many matching algorithm which can achieve a matching result in an abbreviated time. This method can shorten retrieval time and meet retrieval demand while still obtaining accurate retrieval results. In addition, this method considers shape information of parts, although the connection and other information are not considered, its feature of relaxed constraints can support flexible retrieval as well as overall retrieval.

The second section reviews related work. Section 3 briefly introduces the approach to turning an assembly into a point set. Then in Section 4, we demonstrate a method of computing assembly similarity based on a modified Hausdorff distance. We discuss the experiments in Section 5 and end with our conclusions in Section 6.

2. Literature review

3D model retrieval is of longstanding interest to the design community and has been widely studied. From the research perspective, the essentials of a retrieval system are algorithms of model comparison, which largely depend on the type of shapes to be analyzed, and on the properties that are considered relevant in the problem at hand [9]. An important aspect of model retrieval is converting the shape into a calculable form, such as a feature vector or a relational data structure [10]. Therefore, many retrieval methods have been proposed, such as spherical harmonics [11,12], Poisson histogram [13], Reeb graph [14,15], heat kernel signature [16,17], face adjacent graph [18,19], etc. Some AI algorithms, such as deep learning [20–22], have also been employed in the retrieval of 3D shapes.

In engineering field, model retrieval technology has been widely employed, especially for retrieving a similar part model from a model library to accelerate the design process. To facilitate the reuse of parts, Liu and He [23] first voxelized a 3D model and divided the voxelized model into several subspaces, then calculated the entropy of each subspace to form a feature vector to represent the model when compared with other models. Bai et al. [24] represented the design-reusable subpart by all its local matching regions hierarchically, and presented three partial retrieval modes to support partial retrieval of 3D CAD models for design reuse. Cheng et al. [25] used a negative feature tree (NFT) to represent geometric primitives decomposed from the 3D parts, and weighted sum of similarity values on each level of NFTs was used to measure the similarity between two parts. In the research of hybrid information representation, You and Tsai [26] defined an attributed graph for model to capture the geometric and topological information, and a retrieval framework was presented to allow a user to acquire similar mechanical artifacts. Shapira et al. [27] proposed a context-HSDF framework to automatically find part analogies among 3D objects. The framework uses a shape-diameter function and semantic text to describe model information of a part to calculate similarity. Savelonas et al. [28] took a differential fast point feature histogram (dFPFH) as descriptor of the CAD model, and proposed a hybrid algorithm that integrated local information directly derived from local shape descriptors, as well as global shape information derived from Fisher vectors to compare models. In engineering applications, the model retrieval technology has been shown to satisfy the actual requirements of design reuse, and it has the potential to change the design mode by reusing existing designs.

Later, methods of assembly retrieval have been proposed, and the first task of assembly retrieval is to describe the assembly model. In assembly process planning, especially the assembly sequence planning, there are some kinds of descriptions of assemblies, such as the assembly relationship matrix [3], connector based structure [4,5], liaison graph [6], directed graph [7], and polychromatic sets based model [8]. However, these descriptions mainly focus on topological relationship in assembly, hardly providing a quantitative form of the part shape which is consistent with subjective feeling for deciding the similarity of assemblies. Therefore, they cannot be directly used in assembly retrieval.

Deshmukh et al. [29] directly introduced assembly retrieval, and provided a mating graph in which each node represented a part and each edge represented the mating relationship between two parts. The nodes and edges also carried attributes of the assembly, therefore, users could search assemblies with number of parts, size of assembly, overall shape characteristics, and other attributes in their system. Chen et al. [30] proposed a multilevel assembly descriptor to capture topological, geometric, semantic, and other kinds of information that varies from high-level to low-level. The similarity assessment method was also provided so that both whole and partial assembly assemblies could be retrieved with flexible queries. Both descriptors covered comprehensive information of assembly, and the corresponding retrieval methods ensured the retrieval results were reliable. However, both methods required much information that cannot be extracted from CAD models, such as semantic information. Therefore, some required information had to be added manually, which brought tedious and error-prone preprocessing of assemblies in model library.

Considering a retrieval of lightweight assembly models, Hu et al. [31] proposed a method that only considered geometric information of parts and composition information in assemblies. In their method, an assembly was represented as a vector through decomposing the assembly into a bag of parts, and a relaxed matching framework was presented to allow users to input fuzzy queries and obtain quantitative results. However, since this approach was based on a vector space model, the vector dimension would become high if model library contained many kinds of assemblies. If the library changed, all descriptions of the assemblies had to be reformulated. Wang et al. [2] represented an assembly as a point set, in which each point was constructed by quantitatively capturing a part shape. Then, an earth mover's distance (EMD) based method was proposed to measure the similarity between assemblies. This approach could obtain satisfactory retrieval results through considering the shape information of the parts. Unfortunately, it took the EMD method a relatively long time to solve the transportation problem and obtain the optimal solution.

In summary, assembly retrieval is a new research topic, and few studies have focused on it. This paper will consider the disadvantages of existing methods and propose an improved approach inspired by related works.

3. Quantitative description of assembly based on shape distributions

The basis and premise for achieving model retrieval is the quantitative description of a three-dimensional assembly model. An assembly quantitative description method based on shape distributions is put forth in [2]. Below we introduce such a method.

In shape distributions, a part's shape is described by a histogram, and different shapes are reflected in different histograms. By extracting the height of each bin in a histogram and representing the part in dimensional coordinates with a selected point, we transform a part model into a point in the multidimensional space, and we transform an assembly model into a point set in space.

The specific steps in the quantitative description of an assembly are as follows:

Step 1. Divide assembly model A into n part models, where A refers to a specific assembly model.

Step 2. Take random samples on surface of the single part model. Through measurement and statistics on distance between sampling points, turn a part into a point in a 1024-dimensional space to realize a quantitative description of the part model. The specific steps are as follows:

Step 2–1. Select a shape distribution function. According to [32], D2 calculates the distance between two points, and it is a good