Computers & Graphics 34 (2010) 545–555

Contents lists available at ScienceDirect

Computers & Graphics

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

Technical Section

Automatic discovery of common design structures in CAD models

Lujie Ma *, Zhengdong Huang, Yanwei Wang

CAD Center, School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, PR China

ARTICLE INFO

Article history:
Received 4 September 2009
Received in revised form
27 May 2010
Accepted 7 June 2010

Keywords:
Design reuse
Common design structure
Shape analysis
Frequent subgraph discovery
Data mining

ABSTRACT

This paper presents an approach for extracting common design structures from a set of B-rep models. Here, a B-rep model is first transformed into a face adjacency graph (FAG), and then each node of an FAG is mapped to a point in a two-dimensional plane after representing face shape characteristics with two coordinates. Thus, the common design structures are just the frequently appearing subgraphs of FAGs drawn in a plane. In the area of data mining, the apriori-based graph mining (AGM) is a well-known algorithm for solving the problem of frequent subgraph discovery, but its efficiency is still low in processing large graphs like the FAGs of CAD models. In this research, we develop a novel algorithm that improves AGM in two aspects. First, the exact subgraph-isomorphism checking is replaced by comparing the shape descriptors composed from the point coordinates corresponding to the nodes of the subgraphs in question. Second, a new approach for generating frequent subgraph candidates is adopted, which allows large frequent subgraphs to be found in fewer iterations. Experiments show that the proposed method is efficient and can produce a reasonable result.

1. Introduction

Enterprises in industry have a tradition of exploiting the commonality of products to improve design efficiency and consolidate manufacturing process, which is reflected in the existence of such a wide variety of applications of techniques like case-based design, product platforms, product families, and group technology. However, they usually identify the commonalities manually based on their past experience. The automation of product commonality discovery has obvious advantages over the manual method, like its higher efficiency and ability to take into account more detailed information. Since the CAD model includes kernel information on products, here we attempt to investigate the problem of automatic common design structure discovery (CDSD) from a set of solid models.

In general, a common design structure is a local structure that frequently appears in multiple models; it is also called a frequent substructure of models. Therefore, CDSD is a problem of frequent substructure discovery and can be formally described as follows. Given a set of 3D models $\boldsymbol{M}=\{m_1, m_2, m_3, \ldots\}$ and a partial model m', if m' appears f_i times as a substructure of m_i and $\sum_{i=1}^{|M|} f_i \geq \zeta$ for a given threshold value ζ, then m' is called a common design structure of $\boldsymbol{M}$. This mathematical formation is quite similar to that of frequent subgraph discovery in data mining [1]; however, CDSD is more difficult than general frequent subgraph discovery problems because it is not only related to frequent substructure discovery, but also to shape an analysis of 3D models, especially local shape analysis.

In the area of computer-aided design (CAD), 3D geometric model retrieval [2] has been widely researched in the last decade; the model retrieval uses shape analysis technique to discover useful or re-usable information in a large sample of models. Although both CDSD and model retrieval require shape comparison, their objectives are completely different; the goal in CDSD is to identify local structures shared by multiple models, while model retrieval seeks to find models with a given structure. In shape comparison, various graphs like reeb graphs [3] and shock graphs [4] have been used because the graph is an adequate form for the description of local structures. However, these graphs focus more on the representation of the topology of 3D models, and their geometry description capability is relatively poor. In contrast, face adjacency graphs (FAG) are easier to create and are suitable to describe the topological relations as well as the geometric shapes of CAD models. Therefore, in this research, we use FAGs for model description, but we extend their node attributes to capture the local shape characteristics around a face, not only that of a single face corresponding to a graph node.

In addition to part model retrieval, assembly model retrieval also receives attentions from researchers recently [5]. To achieve more effective assembly retrieval than traditional text-based methods, content-based assembly retrieval not only needs to compare the shapes of parts in assembly, but also needs to compare the mating and joint relationships of assemblies. In this research, we restrict the CDSD study to part models, but it is of significance to extend the topic to cover assembly models just like extending part search to assembly search.

* Corresponding author: Tel.: +86 02787543973.
E-mail address: mljyahu@yahoo.com.cn (L. Ma).

0097-8493/$ - see front matter
doi:10.1016/j.cag.2010.06.002

In the area of data mining, the apriori-based graph mining (AGM) [6] is a well-known algorithm for solving the problem of frequent subgraph discovery, but its efficiency in processing large graphs like the FAGs of CAD models is still low. Here, we develop a novel frequent subgraph discovery algorithm. Compared with an AGM algorithm, the proposed algorithm improves the computation efficiency in two aspects. First, the exact subgraph-isomorphism checking is replaced by comparing the shape descriptors composed from the geometric information of faces corresponding to the nodes of the subgraphs in question. Second, a new approach for generating frequent subgraph candidates is adopted, which allows large frequent subgraphs to be found in fewer iterations.

The rest of this paper is organized as follows. First, a brief review of the related literature is given in Section 2, and a method of solid description using FAG is introduced in Section 3. Then, a detailed CDSD algorithm is developed in Section 4. Following this, the implementation method as well as the results obtained are presented in Section 5. Finally, the paper concludes with some discussion and conclusions.

2. Related work

Common design structure discovery is related to local shape-matching and frequency subgraph discovery, both of which issues are widely studied in the areas of CAD and artificial intelligence. In this section, some typical algorithms are reviewed.

In local shape analysis, various forms of graphs are used for shape-matching, such as reeb graphs, shock graphs and feature relation graphs (FRG). The reeb graph of a geometric model is a skeleton consisting of connected points that represent collapsed contours defined with a continuous scalar function on the surface of the model. Hilaga et al. [7] developed a kind of multi-resolutional reeb graph to capture the global structures of 3D models in computer graphics, which was later applied to local structure search in mechanical CAD models by Silvia [3] et al. and Bespalov et al. [8]. Shock graphs are usually generated through a skeletonization process, which could be distance transformation, thinning, or Voronoi-based medial axis extraction. Gao et al. [4] and Sundar et al. [9] have studied methods of searching 3D models by means of Shock graphs, which are also suitable for local structure matching. Although these graphs are adequate for structure representation, the time cost of graph generation is high and their abstract nature prevents them from giving concrete shape representations of common structures when they are used in CDSD. Feature relation graphs (FRG) represents another type of graph-based method for analyzing model shapes. Some researchers [10–12] have proposed part similarity assessment methods based on invariants of feature types and attributes. However, for a more complete assessment, feature relations should be taken into account. Huang et al. [13] have presented a type of an FRG based on feature stacking, parallel, and intersection relationships. The major disadvantage of shape-matching based on an FRG is due to the fact that the FRG of a part is not unique.

Besides shape analysis, another area closely related to CDSD is frequent subgraph discovery in data mining. The core of frequent subgraph discovery is to generate frequent subgraph candidates. According to the candidate generation approach, the methods of frequent subgraph discovery can be classified into node-expansion-based and edge-expansion-based methods. The apriori-based graph mining (AGM) [6] and graph-based substructure pattern mining (gSpan) [14] are two well-known algorithms for the two kinds of methods. To discover all of the frequent subgraphs exactly, subgraph-isomorphism matching is used in these algorithms. Obviously, subgraph-isomorphism matching is an NP problem, whose efficiency is unacceptable in practical applications. While the efficiency of the edge-expansion method is slightly better than that of node-expansion method, both of them remain time-consuming because they rely on exact subgraph-isomorphism checking. To overcome this problem, this research compares subgraphs through a kind of graph invariant with high differentiation capability. Another shortcoming of the above-mentioned methods is that they expand frequent subgraphs with small step sizes, which requires more iterations to find large frequent subgraphs. Taking into account this issue, this research adopts a subgraph merging method to expand frequent subgraphs. With these two improvements, the efficiency of our approach can meet the requirements of CDSD for large CAD model databases.

3. Model description

Selecting a proper representation for CAD models and developing an effective model description method are two fundamental problems in common design structure discovery. Unlike a general 3D model represented by a triangular mesh in computer graphics, a CAD model has multiple kinds of representations like voxel representation, mesh representation, boundary representation (B-rep) and constructive solid geometry (CSG). Nevertheless, only B-rep and CSG are two popular representations used in commercial software. Since CSG does not represent a solid model uniquely, B-rep is usually preferred for the description and analysis of solid models [15–18]. Consequently, the algorithms for CDSD herein are also developed based on the B-rep model. Usually, a model description is a set of values, graphs or other abstract mathematical formulations that are extracted from the model representation to capture the major shape characteristics of solid model. To generate the shape characteristic descriptions of a solid model, various transformations are usually adopted to map the solid model in 3D Euclidean space E^3 to a specific description space S. Since the shapes of individual faces and their connectivity relations are both important to the description of a solid model, graph-based model description is suitable for CAD models. In this paper, face adjacency graphs (FAG) are utilized to express topologic relations in B-rep, and some attributes representing geometric shape characteristics, which are called shape parameters in this paper, are often attached to the FAG's nodes and arcs.

CAD models from enterprises usually have small features like fillet and chamfers, which may not be a major concern in shape comparison and design reuse, but usually cause unexpectedly large differences in the model description. To eliminate the sensitivity of the model description to these small features, many shape analysis algorithms usually remove this kind of features beforehand [10–12,16,17,19,20]. Here, we handle this problem in the same way and omit these features during the generation of FAGs.

In this section, we first introduce the geometric shape information of the B-rep model, and then present our approach to describing CAD models based on an FAG.

3.1. Geometric information on the B-rep model

The geometric information of the B-rep model is mainly face and edge information. In this paper, the descriptions of faces and edges are based on point classifications. Given a point on a face, its geometric characteristics can be described by the two principal curvatures $\kappa_{\min}$ and $\kappa_{\max}$ at the point. The value of a principal curvature can be positive (+), negative (−) or zero (0) [21] (see Fig. 1). Thus, two parameters $\omega=(\lambda_{\min},\lambda_{\max})$ can be used to represent different types of points. Here, $\omega=(\lambda_1,\lambda_2)$ is