Information Sciences 320 (2015) 429–442

Contents lists available at ScienceDirect

Information Sciences

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

Graph-based characteristic view set extraction and matching for 3D model retrieval

Anan Liu, Zhongyang Wang, Weizhi Nie *, Yuting Su

School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China

ARTICLE INFO

Article history:
Received 18 June 2014
Received in revised form 15 April 2015
Accepted 18 April 2015
Available online 9 May 2015

Keywords:
3D model retrieval
Characteristic view
Graph clustering
Graph matching
Rayleigh quotient maximization

ABSTRACT

In recent times, multi-view representation of the 3D model has led to extensive research in view-based methods for 3D model retrieval. However, most approaches focus on feature extraction from 2D images while ignoring the spatial information of the 3D model. In order to improve the effectiveness of view-based methods on 3D model retrieval, this paper proposes a novel method for characteristic view extraction and similarity measurement. First, the graph clustering method is used for view grouping and the random-walk algorithm is applied to adaptively update the weight of each view. The spatial information of the 3D object is utilized to construct a view-graph model, thus enabling each characteristic view to represent the discriminative visual feature in terms of specific spatial context. Next, by considering the view set as a graph model, the similarity measurement of two models can be converted into a graph matching problem. This problem is solved by mathematically formulating it as a Rayleigh quotient maximization with affinity constraints for similarity measurement. Extensive comparison experiments were conducted on the popular ETH, NTU, PSB, and MV-RED 3D model datasets. The results demonstrate the superiority of the proposed method.

1. Introduction

With the development of computer vision and digital collection equipment [56,28], 3D models play an important role in many applications, such as medical industry, movie production, architecture design, and so on [37,38,60]. The popularization of 3D model has created an urgent demand for effective 3D model retrieval systems. The effective 3D model retrieval algorithm can speed up design or research processes, thus, content-based 3D model retrieval methods have become a hot topic in recent years [11,13].

Many methods have been proposed to handle 3D model retrieval problem [1,4,46,21]. Ankerst et al. [2] proposed an optimal selection of 2D views from a 3D model, which focuses on numerical characteristics obtained from the 3D model representative features. Shih et al. [44] proposed the Elevation Descriptor (ED) feature, which is invariant to translation and scaling of 3D models. However, it is not suitable for the 3D model that consists of a set of 2D images. Ansary et al. [3] proposed a Bayesian model-based method for 3D object search, which utilizes X-means [9] to select characteristic views and applies the Bayesian model to compute the similarity between different models. Gao et al. [23] proposed a general framework for 3D object retrieval to avoid the restriction of camera array setting. The model is generated on the basis of the query

* Corresponding author.
E-mail address: truman.nie@gmail.com (W. Nie).

http://dx.doi.org/10.1016/j.ins.2015.04.042
0020-0255/

Gaussian models by combining the positive matching model and the negative matching model. The method can achieve good performance but has high computation complexity, which makes it unsuitable for the real-time retrieval system. In all these retrieval methods [52,23], spatial information, which could benefit characteristic views extraction, has not been considered. Further, the inner correlation of images in each 3D model has not been discovered yet.

In order to tackle this problem, we propose a novel graph-based characteristic view set extraction and matching method for 3D model retrieval. First, the graph clustering method is implemented to extract the characteristic view set (CVS) of each model. Specifically, the characteristics of individual views (node of the graph) and the spatial information between pairwise views (edge of the graph) are leveraged to construct the graph for model representation. We formulate the energy function for graph clustering and solve it with the graph cut algorithm. Second, the view set of individual model can be grouped into several clusters and CVS can be extracted. The random-walk algorithm is further applied to adaptively update the weight for each characteristic view. Finally, the graph matching method is proposed to measure the similarity between pairwise CVSs by taking advantage of both visual similarity and spatial context. We formulate the objective function of 3D model matching as the Rayleigh quotient maximization with affinity constraints. The framework of the proposed method is shown in Fig. 1.

The main contributions of this paper are summarized as follows:

- Graph clustering for view clustering: Spatial context and visual similarity are used built graph model for representing 3D model. *Normalizedcut* is utilized to extract CVSs.
- Adaptively updating the weights of CVSs: We utilize the random-walk method to update the weights of individual 2D images for the similarity measurement between different models;
- Graph matching for similarity measurement between pairwise CVSs: The similarity between pairwise CVSs is measured in terms of visual similarity and spatial context, which is mathematically formulated as a the Rayleigh quotients maximization with affinity constraints;
- A new 3D model dataset: In this paper, we present a new 3D model dataset, MV-RED object dataset,[1] which has been utilized in the 3D Shape Retrieval Contest 2015 held in EUROGRAPHICS 2015. This dataset consists of 505 models and each model includes multi-view and multi-modal information. We will provide details of this dataset in Section 5.

This manuscript is organized as follows. In Section 2, related work is reviewed. In Section 3, the method for characteristic view extraction is detailed. Section 4 illustrates the similarity metric to compute the distance between the query model and the candidate model. Experimental results and discussion are provided in Section 5. Finally, the conclusion is stated in Section 6.

2. Related work

In this section, we will review the representative work in the field of 3D model retrieval [36,39,8,55]. 3D model retrieval methods are mainly classified into two categories [19,22]: (1) geometry-based techniques [50,42,26,34,10], (2) view-based techniques [9,54,59]. Early studies mainly focused on geometry-based techniques. This approach applied shape distribution, shape histogram, and other 3D spatial information to represent a 3D model. For example, Mihael et al. [2] introduced 3D shape histograms as intuitive feature vectors. The histograms were based on the partitioning of the space. Three methods for decomposition were proposed: the shell model, the sector model, and the spider web model. Vandeborre et al. [48] proposed the use of completed 3D information. The 3D objects were represented as mesh surfaces and 3D shape descriptors were used. With the wide spread of depth sensors, an increasing number of researchers utilized depth information to represent the geometry feature of an object. Zhao et al. [57] proposed a novel depth image-based method to handle action retrieval problem. Hu et al. [30] proposed an action tutor system that utilized an iterative learning method to learn the classifier for each object. Shape-based 3D model retrieval is also a popular research area. Hilaga et al. [29] developed a matching technique to compute the similarity between two models based on the shape matching of their multi-resolutional Reeb Graphs. Sundar et al. [49] compared 3D objects by applying graph matching techniques to match their skeletons. Additional related information is available in [43].

View-based techniques have attracted more attention in recent years because these techniques can benefit from numerous mature digital image processing algorithms. Further, in comparison with most geometry-based methods, view-based methods can significantly reduce the complexity of the algorithms. View-based 3D object retrieval is a special content-based image retrieval task because each 3D object is represented by a set of 2D view images that are recorded from different cameras. In [9], each object has 10 views, which are encoded by Zernike moments [31] and Fourier descriptors as features [5]. An interactive view selection method was introduced in [16], in which the query views were selected incrementally based on user relevance feedback. This is the first interactive method for representative view selection and it has shown acceptable performance in terms of precision and speed. Wang et al. [54] applied group sparse coding to represent the image set of each object, and utilized the reconstruction error to compute the similarity between the query model and the candidate model. The authors [21] proposed a novel 3D model descriptor, spatial structure circular descriptor (SSCD), that can convert the structure information of a 3D model into 2D images and applied traditional image processing methods for 3D

[1] http://media.tju.edu.cn/mvred/dataset1.html.