Neurocomputing 168 (2015) 761–769

Contents lists available at ScienceDirect

Neurocomputing

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

A powerful 3D model classification mechanism based on fusing multi-graph

Biao Leng *, Changchun Du, Shuang Guo, Xiangyang Zhang, Zhang Xiong

School of Computer Science & Engineering, Beihang University, Beijing 100191, PR China

ARTICLE INFO

Article history:
Received 16 June 2014
Received in revised form
11 May 2015
Accepted 12 May 2015
Communicated by Zhi Yong Liu
Available online 22 May 2015

Keywords:
3D model classification
Graph fusion
Boost modified

ABSTRACT

Recently, integrating several feature descriptors to be a powerful one has become a hot issue in the field of 3D object understanding. The fusing mechanism is so crucial that can significantly affect the performance of 3D model classification. In this paper, a powerful model for 3D model classification, which can novelly integrate several graphs, is proposed. This mechanism is based on graph fusion and modifies each graph's weight in a boost manner. Each graph's weight in the fusion graph can be dynamically calculated according to its performance. Finally, a fusion graph is acquired to 3D model classification. We conduct the experiments on the publicly available 3D model databases: Princeton shape benchmark (PSB) and SHREC'09, and the experimental results demonstrate the powerful performance of the proposed method.

1. Introduction

The high-speed advance in graphical hardware and the popularity of Internet has led the widely applications of 3D models in many fields [4,7], such as movie production, architecture design, and medical industry. With the advent of mobile Internet [29], the scale of 3D model will remarkably increase. Therefore, high-effective technologies [9,21,22,32,37,40,58,65,68] for dealing with 3D models are becoming more and more important [17].

The field of 3D model understanding contains 3D model retrieval [4], 3D model cluster [64], 3D model classification [50], and so on. Research on extracting model features and measuring the distance between models are fundamental tasks. Until now, a lot of related methods have been proposed [24].

In the domain of extracting model descriptors, different levels of model features have been presented, including low-level features, such as machined features [27], surface distribution [4,47], volumetric information [15,57], view-based features [55,49,6,60,19,61,41,39] and some hybrid features [33,52,53,38].

Another issue, measuring the similarity between models, is closely related with descriptors of models. In general, there are two approaches to deal with the problem. The first one is to calculate the distance by descriptors [2,12], and the efficiency of these methods depends on the "reasonability" of descriptors. The other one is to acquire the semantic information, such as bag of words [16,18]. Gao et al. have imported probability concept to 3D model understanding [18]. Leordeanu et al. proposed a semi-supervised algorithm for learning the parameters that control the hypergraph matching model [42]. Leng et al. presented a viewpoints segmentation mechanism to select the best representing view [31].

Lots of methods for 3D model retrieval and classification consider one descriptor. But a single descriptor may be extracted to acquire just one property (such as discriminative or invariance), and furthermore, there must be some object information lost in the transform processing undoubtedly [54]. Therefore, just considering one descriptor is hard to capture more comprehensive features to describe the 3D model. There is a consensus in the field of 3D model understanding that integrating different features can always increase the effectiveness, and some papers have illustrated that fusing different descriptors can greatly improve the performance [11,60,62]. Recently, Yang et al. proposed a multi-graph learning method for model matching by integrating three directed structural models [66]. There are some different schemes for fusing various descriptors, and two significant mechanisms are shown in Fig. 1. This figure displays two kinds of fusion mechanisms. The early one fuses the descriptors in the beginning, and then acquires a "global feature" [11,60]. The other one is based on graph, and consists of three steps. At first, this scheme constructs graphs after obtaining distance between models. Then, it fuses the graphs by using a series of weights. In terms of the weights, they can be set equal, or other mechanism. At last, learning is conducted on the fused graph.

In this paper, a semi-supervised learning method based on fusing multi-graph is proposed for 3D object classification. First of all, a few model descriptors are extracted with different features.

* Corresponding author.
E-mail address: lengbiao@buaa.edu.cn (B. Leng).

http://dx.doi.org/10.1016/j.neucom.2015.05.048
0925-2312/

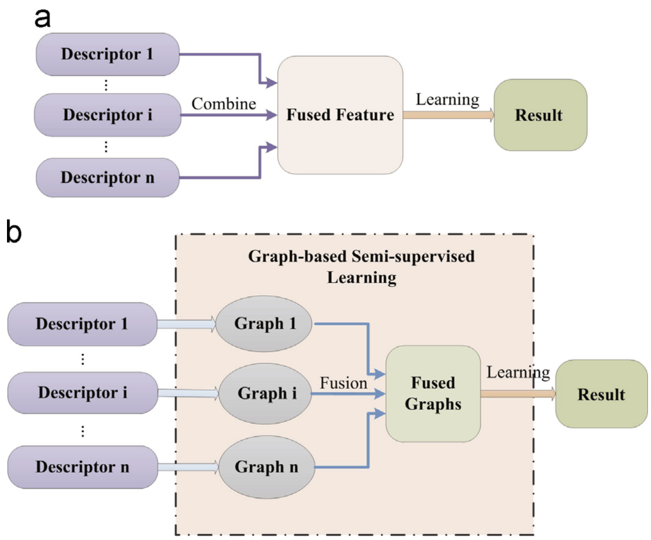

Fig. 1. Two schemes of fusion method: early fusion and fusion based on graph.

Then, corresponding graphs can be acquired by calculating the distance between models. Initial semantic information can be obtained by labeling samples manually. The following step is to classify 3D models by using the unsupervised features and labeled information. We propose a method to integrate several feature descriptors into a fused one, by adaptively adjusting each descriptor's weight. Then, a novel standard is presented to measure each descriptor's performance, and its weight is modified according to its performance. Finally, a fused graph is obtained, and the classification can be implemented based on the fusion graph.

The structure of this paper formulates the process of the proposed mechanism. A review of related work is presented in Section 2. In Section 3, some model descriptors are introduced, such as 2D polar-FFT, 2D Zernike moments, and 2D Krawtchouk moments. Section 4 formulates the classification method based on graph fusion. In Section 5, we conduct some experiments to test the proposed method. Conclusions are drawn in Section 6.

2. Related work

The study related with 3D model understanding is a relatively new and challenging work. Many research communities have devoted their efforts to this field. The study could be divided into three parts in general: (1) model descriptors; (2) 3D model semantic retrieval; and (3) 3D model classification.

2.1. Model descriptors

Up to now, many 3D model feature extraction methods have been proposed. Generally speaking, they can be divided into two types: low-level descriptors and view-based descriptors.

Low-level descriptors focus on capturing the low-level feature of models [35], such as geometric information, visual information, contour information, and so on. Depth buffer (Dbuffer) method [59] is used to characterize 3D models with six depth-buffer images. Silhouette [26] describes 3D models in terms of contour information, which can be obtained from canonical project. The light field descriptor (LFD) [9] is proposed to calculate 10 silhouettes obtained from the vertices of a dodecahedron over a hemisphere, and it has been tested to have better performance on the Princeton Shape Benchmark (PSB) [56]. But later, the descriptor DESIRE [60], combined from Ddepth, silhouettes, and ray-extents, is proved to be superior than LFD. Organizing the low level descriptor statistically is a generally handling way. Ferreira et al. [14] propose a part-in-whole matching method, which uses a collection-aware shape decomposition combined with a shape thesaurus and inverts indexes to retrieve 3D models using part-in-whole matching.

In the view-based 3D object retrieval framework, each object is described by a set of views and representative features. Adaptive views clustering [1] method can select the most characteristic views from 320 basic views. Then the Zernike features of 320 views are grouped into an adaptive number of clusters. Daras and Axenopoulos [11] propose the compact multi-view descriptor (CMVD) which can support Multimodal queries. This method is also a composite method combining several moments (2D Polar FFT moement, 2D Zernike moment, 2D krawtchouk moment). A novel 3D multi-view representation method bag-of-region-words [18] is proposed.

2.2. 3D model semantic retrieval

Matching is the process of determining how similar two objects are [58]. Therefore, the concentration of 3D model retrieval is to measure the similarity between objects. Based on the descriptors, distance between models is regarded as the similarity, and the methods of measuring distance include L1 distance, Euclid distance, Hausdorff distance [20], Cosine distance, and Earth Mover's distance [44].

In order to bridge the gap between low-level features and high-level semantic information, exploring method satisfying human cognition has become a hot issue in recent years [34,36]. Statistical methods and machine learning algorithms have been widely used for relevance feedback. Elad et al. [13] firstly apply machine learning concept to 3D model retrieval. The support vector machines (SVM) is used for the adaptation of measuring distance, bringing the "relevant" objects closer and pushing the "irrelevant" objects farther. Cox et al. [10] propose an approach called PicHunter, based on the Bayesian method, represents its uncertainty about the user's goal by using a probability distribution over possible goals. In order to acquire semantic information, some methods about relevance feedback have been explored in recent years, and they are classified into short-term and long-term methods [69]. Leifman et al. [30] present a novel relevance feedback algorithm that is based on supervised as well as unsupervised feature extraction techniques. Gao et al. [18] adopt a probabilistic graph model, in which they divide the captured views into several sets, and model them as a rise order in Markov Chain. The query comparison is then converted to be probabilistic analysis, which significantly alleviates the effects of transformation. Gao et al. [24] have proposed another multi-bipartite graph reinforcement model to re-rank all the views, with perfect experimental results.

2.3. 3D model classification

Compared with 3D model retrieval, fewer algorithms for 3D model classification have been brought forward. One of the important reasons is the "curse of dimensionality". Furthermore, it takes much effort to manually label the test dataset for classification. As a result, the desire of acquiring semantic information from a few labeled models has attracted extensive attentions. An algorithm has been proposed based on finding minimum cuts in graphs [5], which uses pairwise relationships among the examples in order to learn from both labeled and unlabeled data. Leifman et al. [30] adapt SVM for 3D model classification. A method is proposed to show how to learn a Mahanalobis distance metric for k-nearest neighbor classification by semi-definite programming [63]. Wang et al. [62] have brought forward a