Computers & Graphics 37 (2013) 460–472

Contents lists available at SciVerse ScienceDirect

Computers & Graphics

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

Special Section on 3D Object Retrieval

Data-aware 3D partitioning for generic shape retrieval ☆, ☆☆

Ivan Sipiran [a,*], Benjamin Bustos [a], Tobias Schreck [b]

[a] KDW+PRISMA Research Group, Department of Computer Science, University of Chile, Chile
[b] Visual Analytics Group, Department of Computer and Information Science, University of Konstanz, Germany

ARTICLE INFO

Article history:
Received 24 October 2012
Received in revised form
4 April 2013
Accepted 5 April 2013
Available online 1 May 2013

Keywords:
Mesh partitioning
Optimization matching
Shape retrieval

ABSTRACT

In this paper, we present a new approach for generic 3D shape retrieval based on a mesh partitioning scheme. Our method combines a mesh global description and mesh partition descriptions to represent a 3D shape. The partitioning is useful because it helps us to extract additional information in a more local sense. Thus, part descriptions can mitigate the semantic gap imposed by global description methods. We propose to find spatial agglomerations of local features to generate mesh partitions. Hence, the definition of a distance function is stated as an optimization problem to find the best match between two shape representations. We show that mesh partitions are representative and therefore it helps to improve the effectiveness in retrieval tasks. We present exhaustive experimentation using the SHREC'09 Generic Shape Retrieval Benchmark.

1. Introduction

Three-dimensional objects are a valuable resource in many fields such as engineering and medicine. They can represent the shape of a real object in a suitable way in order to be used by computers. The versatility of this representation has resulted in an increasing interest of the scientific community in several related topics. For instance: shape analysis, shape processing, modeling applications, and so on. In addition, it is currently possible to find massive and publicly available 3D data. For example the Google Sketchup collection, for which its use is becoming a common practice. For these reasons, the search for efficient and effective tools for this kind of data is imperative in order to support future applications.

In particular, the content-based similarity search of 3D objects has received much attention in recent years. This can be performed without relying on additional information for searching, only using the provided shapes. Additionally, many fields (for example medicine [1,2], CAD/CAM [3], etc.) have benefited from the large amount of approaches proposed to overcome the problem of 3D matching. Nevertheless, the problem remains challenging and it is far from being completely solved. Moreover, part of the problem resides in the possibility of defining a suitable similarity measure between 3D models.

In this paper, we consider the problem of generic shape retrieval. A common approach to facing this problem is to compute an intermediate representation (feature vectors or graphs, for instance) and subsequently defining the similarity of two objects as the similarity of their representations. In this direction, there are methods that exploit the visual similarity, the statistical properties of 3D measures, or the possibility of defining transform functions on the data, just to name a few. However, one of the most critical problems is the semantic gap. That is, the intermediate representation may not be able to capture all the needed information of a shape and therefore the effectiveness of searching may be seriously affected.

A previous study by Bustos et al. [4] showed that some features could well represent certain classes of objects and furthermore, some features could be complimentary in representing a shape. This is because algorithms cover only a part of the possible spectrum of characteristics such as shape, silhouette, or intrinsic properties. Thus, a natural extension of classic approaches was the combination of features for improving the effectiveness of retrieval. Approaches in this direction have been previously presented by Bustos et al. [5], Vranic [6], and Papadakis et al. [7], all of them with promising results. However, the semantic gap is still latent in this approach as any possible combination of features could not represent important characteristics to discriminate between objects.

A more recent approach is the combination of global and part-based information. The idea is to combine features extracted from an entire object with features extracted from parts of an object. Some techniques have been presented so far by Li and Johan [8], Bustos et al. [9], and Schreck et al. [10]. All these techniques share a common aspect: the part-based features come from a fixed partitioning of the objects. Although it was possible to improve the effectiveness with respect to using only global features, the fixed

☆☆To comment on this article, please join the discussion on the Collage Authoring Environment Google Group https://groups.google.com/group/collage-authoring-environment.

* Corresponding author. Tel.: +56 2 29784972.
E-mail addresses: iasipiranm@gmail.com, isipiran@dcc.uchile.cl (I. Sipiran), bebustos@dcc.uchile.cl (B. Bustos), tobias.schreck@uni-konstanz.de (T. Schreck).

0097-8493/$ - see front matter © 2013 The Authors. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cag.2013.04.002

Fig. 1. Two globally dissimilar chairs. Note that the chair at right is taller than the left one. Nevertheless, it is possible to find similarities between their parts, which can be exploited to improve the similarity measure between the two objects.

partitioning limits the possibility of having truly distinctive parts. This opens up a question on how to define a new kind of partitioning dependent on the shape information.

We believe that the use of local features can enhance the use of global features in shape retrieval. That is, we are trying to mitigate the effect of the semantic gap. For instance, a common fact is having two objects with different appearance in the same class. Obviously, a global feature could differ in those objects. However, in a local sense, it is still possible to find correspondences between parts, so we can take advantage of this fact to improve the similarity measure (see Fig. 1). Therefore, the discriminative power of local features combined with global features could help to improve the effectiveness in the similarity search.

In this paper, we propose a shape retrieval method using a data-aware partitioning algorithm. Our idea is to exploit the local characteristics of objects to determine discriminative parts. Thus, each object is represented by its global feature and a set of features extracted from parts. The partitioning method relies on finding robust local features (namely keypoints) on the object's surface and subsequently determining the parts where there is a high concentration of keypoints (for instance, a human shape commonly has many features located in hands, feet, and head). Beyond techniques which made use of the bag of features approach to aggregate local descriptors for retrieval, our method is the first attempt in combining global and local features found in a data-adaptive way for generic shape retrieval.

Our main contribution is three-fold:

- We propose a model partitioning algorithm based on local features. Regions on the surface with high concentration of local features will be selected as parts.
- We combine the global feature with features obtained from parts and define a combined distance to assess the similarity. The distance between global features is performed as usual. The distance between sets of parts is stated as an optimization problem. In addition, we propose a geometrical consistency criterion which can be formulated within the same optimization problem.
- We evaluate our approach using a well-known, established benchmark dataset and appropriate performance measures.

Our approach is a generic, simple framework by which global and local descriptors can be combined in a data-adaptive way. The approach is able to provide on average, an improvement over the retrieval effectiveness of state of the art global descriptors. A careful, systematic analysis of the results is performed to assess in detail the magnitude of the improvement, relating it with global-only methods, and identifying classes of models for which the method is particularly effective. We test our approach using a state-of-the-art local interest point detector with desirable properties in combination with two robust view-based descriptors. Our approach is flexible in that it can accommodate further, possibly application-specific, object segmentation and description schemes, if needed.

Our paper is organized as follows. Section 2 briefly presents the state of the art in generic shape retrieval. Section 3 describes our partitioning algorithm based on local features. Section 4 is devoted to the matching methods and the definition of our similarity measure. Section 5 describes our experiments and presents the discussion of our results. Finally, Section 6 draws the conclusions.

2. Related work

The interest in 3D model retrieval has resulted in a large amount of proposed techniques to overcome the problem. One of the most studied approaches is to convert a 3D model into a more convenient representation for comparison, for example feature vectors. Then, the comparison can be done by defining a distance between those representations. For generic shape retrieval, this approach has received attention due to the efficiency of computing distances between vectors. In this section, we provide a brief description of the state of the art related to descriptors for generic shape retrieval and possible combinations to improve the performance. For a comprehensive study, surveys by Bustos et al. [11] and Tangelder and Veltkamp [12] are an excellent resource.

Classic methods for 3D shape retrieval can be classified into three groups: view-based, histogram-based, and transform-based. This classification is based on how a feature is extracted from the shape. View-based methods transform a 3D shape into a set of 2D views and subsequently we can apply image techniques to describe the obtained views. For example, the Depth Buffer method [13] computes six views corresponding to the six faces of the bounding cube of an object. Each view stores the projected distances from the object to the projection plane. Then, each view is represented by Fourier coefficients and the final vector is the concatenation of the six obtained views. Another example is the PANORAMA descriptor [14], which computes three views taken from the lateral faces of cylinders oriented according to the coordinate axes. Similar to the Depth Buffer, each lateral face encodes the distance from the object to the face. Then, Fourier and Wavelets coefficients are extracted from each view, which form the final descriptor.

Histogram-based methods summarize shape properties in order to use them as features. For instance, Shape Distribution [15] is a method that computes several geometric properties (distances between pairs of surface points, angles between three random surface points, etc). The method consists of sampling a large amount of points on the shape surface and subsequently measuring some property. Each value obtained for the chosen property is accumulated in a histogram. Thus, the histogram represents an approximation of the distribution of the property and it is expected to be distinctive for each object.

Transform-based methods consist of converting the geometric information by using some mathematical transformation prior to the feature extraction. The goal of applying a transformation is to enhance some information which is not evident in the Euclidean space. In particular, in 3D model retrieval, there is an interest for spherical harmonics to extract features from shapes. Vranic proposed the ray-based descriptor [13] by using a spherical function which is able to capture the behavior of the rays starting in the origin and the intersections with the shape. Similarly, Kazhdan et al. [16] used spherical harmonics frequencies along with the Gaussian Euclidean Distance Transform in a volume representation of a shape.

An interesting and new approach is the combination of different descriptors to improve the performance of individual descriptors. The basic idea is that different descriptors could extract complementary features and their combination could lead to improvements. Bustos et al. [5] proposed to dynamically