Available online at www.sciencedirect.com
ScienceDirect
Computer-Aided Design 39 (2007) 398–407
COMPUTER-AIDED DESIGN
www.elsevier.com/locate/cad

A spectral approach to shape-based retrieval of articulated 3D models

Varun Jain, Hao Zhang*

GrUVi Lab, School of Computing Science, Simon Fraser University, Burnaby, BC, Canada

Received 1 September 2006; accepted 16 February 2007

Abstract

We present an approach for robust shape retrieval from databases containing articulated 3D models. Each shape is represented by the eigenvectors of an appropriately defined affinity matrix, forming a spectral embedding which achieves normalization against rigid-body transformations, uniform scaling, and shape articulation (i.e., bending). Retrieval is performed in the spectral domain using global shape descriptors. On the McGill database of articulated 3D shapes, the spectral approach leads to an absolute improvement in retrieval performance for both the spherical harmonic and the light field shape descriptors. The best retrieval results are obtained using a simple and novel eigenvalue-based descriptor we propose.

Keywords: 3D shape retrieval; Bending invariance; Geodesic distance; Graph distance; Shape descriptor; Spectral embedding

1. Introduction

In recent years, there has been a tremendous advance in 3D model acquisition technology and a large number of 3D models have become available on the web, e.g., [1–3], or through other means. The problem of indexing and retrieval of 3D shapes [3–19] has become as important, both in practice and in terms of research interest, as that of the indexing and retrieval of image or textual data. Consider a database of 3D shapes represented in the form of triangle meshes. Note that a liberal use of the term mesh is adopted in this paper: the mesh can be non-manifold, open or closed, disconnected, or simply a triangle soup [10]. Given a query shape, a shape retrieval algorithm seeks to return shapes from the database that belong to the same class as the query, where the classification is done by humans. The shapes returned are typically ordered by their decreasing visual similarity to the query shape.

Since the cognitive process of object recognition by humans is not yet completely understood, we are still incapable of proving theoretically that one particular shape retrieval algorithm is the best. In practice, several benchmark data sets and their associated performance evaluations [2,3,5,18] are available to empirically measure the quality of existing shape retrieval algorithms. The most comprehensive comparative study of 3D shape retrieval algorithms to date is due to Shilane et al. [3], based on the now well-known Princeton Shape Benchmark.

A variety of shape retrieval algorithms have been proposed, e.g., see [14] for a recent survey. Typically, each shape is characterized by a shape descriptor. An appropriately defined similarity distance between the descriptors sorts the retrieved models. Commonly used quality criteria for shape descriptors include invariance to rigid-body transformations, scaling, bending and moderate stretching, robustness against noise and data degeneracy, and storage and computational costs. The discriminative power of a shape descriptor and its associated similarity distance is most often judged by plotting the precision-recall (PR) curve [3] generated from a benchmark database, but other evaluation criteria also exist [15].

Most state-of-the-art descriptors, including the twelve compared by Shilane et al. [3] on the Princeton Shape Benchmark, are designed to be invariant to only rigid-body transformations and uniform scaling. Hence, it is no surprise that they do not perform well when applied to shapes having non-rigid transformations such as bending or stretching, which are obviously harder to handle due to their non-linearity and increased degrees of freedom. In this paper, we propose a technique to render a descriptor invariant to bending, hence enhancing its performance over databases that contain

* Corresponding author. Tel.: +1 604 268 6843; fax: +1 604 291 3045.
E-mail addresses: vjain@cs.sfu.ca (V. Jain), haoz@cs.sfu.ca (H. Zhang).

0010-4485/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cad.2007.02.009

articulated shapes. Our experiments will thus be conducted on the McGill database of articulated 3D models [2].

Given a shape represented as a triangle mesh, which may be disconnected, we first convert it into a connected weighted graph. Shortest graph distances between pairs of nodes, mimicking geodesic distances over the shape's surface, provide an intrinsic characterization of the shape structure. We filter these distances appropriately to remove the effect of scaling and then compute a low-dimensional spectral embedding of the shape to obtain invariance to bending and rigid-body transformations. The spectral embeddings are given by appropriately scaled eigenvectors of the matrix of filtered distances. The corresponding eigenvalues can then be used to derive a simple and novel shape descriptor that is shown to work effectively on the articulated models in the McGill database.

Alternatively, any existing 3D shape descriptor can be applied to the 3D spectral embeddings, to improve upon their retrieval performance on articulated shapes. In this paper, we demonstrate using the McGill benchmark data set that this is indeed the case for the spherical harmonic shape (SHD) descriptor of Kazhdan and Rusinkiewicz [8] and the light field descriptor (LFD) of Chen et al. [5], two of the best-performing descriptors from the Princeton benchmark test [3]. As an added advantage, the spectral embeddings can also be used to obtain a meaningful correspondence between the vertices of two shapes [20], e.g., for cross parameterization [21] and other attribute transfer tasks. Finally, it is worth noting that with the aid of sub-sampling and interpolation via the Nyström method [22], the spectral embeddings and the eigenvalue-based shape descriptors are quite efficient to compute.

The rest of the paper is organized as follows. After briefly discussing previous work on 3D shape retrieval in Section 2, we describe the efficient construction of the bending-invariant spectral embeddings for a mesh, possibly with disconnected components and other degeneracies. In Section 4, we give a comparative study between various shape descriptors, including those derived from spectral embeddings and our new eigenvalue-based descriptor, for shape retrieval. Experimental results and discussions are presented in Section 5. Finally, we conclude in Section 6 and suggest possible future work.

2. Previous work

It is quite conceivable that a great deal of prior knowledge is incorporated into the process of human object recognition and classification, perhaps with subpart matching [7,23] playing an important role. In this paper however, we focus on purely shape-based approaches using global shape descriptors. At a high level, a 3D shape retrieval algorithm either works on the 3D models directly, e.g., [8,11], or relies on a set of projected images [5,6,10] taken from different views. Let us call these the object-space and the image-space approaches, respectively. The latter, e.g., LFD [5], has a more intuitive appeal to visual perception, and thus often results in better benchmark results for retrieval [3], but at the expense of a higher computational cost.

Many object-space shape descriptors construct one or a collection of spherical functions, capturing the geometric information in a 3D shape *extrinsically* [3]. These spherical functions represent the distribution of one or more quantities, e.g., curvatures [12], areas [4], surface normals [24], or distance from points on the shape to the center of mass [17]. The bins are typically parameterized by the sphere radius and angles. The spherical functions are, in most cases, efficient to compute and robust to geometric and topological noise, but they may be sensitive to the choice of the sphere center or bin structures. To align the bins for two shapes properly, these approaches require pre-normalization with respect to translation, rotation, and uniform, e.g., [4,17,24], or nonuniform scaling [25]. As an alternative, rotation-invariant measures computed from the spherical functions, e.g., the energy norm at various spherical harmonic frequencies [8], can be applied. However, non-rigid transformations cannot be handled by these approaches.

As one of the most important intrinsic geometric measures, surface curvature, as well as the principal curvature directions, has been utilized for shape characterization and retrieval [12, 18]. But these methods are sensitive to noise and non-rigid transformations such as bending. Another intrinsic approach uses shape distributions [11], where a histogram of pairwise distances between the vertices of a mesh define the shape descriptor. Other forms of statistics [9,11] can also be adopted, and bending-invariance can obviously be enforced if geodesic distances are used in this context. However, in terms of discriminative power, the shape distribution may be too coarse of a shape descriptor to compare favorably against its competitors, e.g., LFD and SHD [3].

The most common approach to handling shape articulation is via the use of skeletal graphs to model geometry or topology [13,19,26,27], followed by graph matching [28]. Some examples of such graphs include medial surfaces [19], multidimensional Reeb graphs [26], and shock graphs [29]. Another advantage of using skeletal representations is that they accommodate part matching [13]. However, the process of extracting the skeletons is typically quite complex and associated with the computational cost of a voxelization step [13,19]. Also, the subsequent graph matching step can be computationally expensive and sensitive to geometric or topological noises. Our approach also uses a graph-based intrinsic shape characterization, which is directly constructed on the input meshes. The spectral embeddings automatically normalize the shapes against rigid-body transformations, uniform scaling, and bending, and they are fast to compute. The resulting shape descriptors provide a more intuitive way of characterizing shapes, compared to shape distribution [11]. In addition, our spectral approach is quite flexible and allows for different choices of graph edge weights and distance computations, which can render the approach robust against topological noise.

The idea of using spectral embeddings for data analysis, mostly for clustering [30,31] and correspondence analysis [32, 20,33], is not new. Past work that is most relevant to ours is the use of bending-invariant shape signatures by Elad and Kimmel [34]. They work on manifold meshes and compute