Bag-of-Feature-Graphs: A New Paradigm for Non-rigid Shape Retrieval

Tingbo Hou   Xiaohua Hou   Ming Zhong   Hong Qin

Department of Computer Science, Stony Brook University
{thou, xihou, mzhong, qin}@cs.stonybrook.edu

Abstract

This paper advocates a new paradigm, called bag-of-feature-graphs (BoFG), for non-rigid shape retrieval. It represents a shape by constructing graphs among its features, which significantly reduces the number of points involved in computation. Given a vocabulary of geometric words, for each word the BoFG builds a graph that records spatial information of features, weighted by their similarities to this word. This eliminates unlikely points in a word category, during shape comparison. Feature graphs are governed by their affinity matrices of weighted heat kernels, whose eigenvalues form a concise shape descriptor. Evaluations of the proposed method are conducted via quantitative measurements. The results demonstrate that the BoFG has competitive precisions w.r.t. state-of-the-art methods, and is much faster to compute.

1. Introduction and related work

Non-rigid 3D shape retrieval is a challenging problem in computer vision. It has received considerable attention in recent years. One of the most-recent techniques is based on an informative shape representation in graphics: the heat kernel signature (HKS) [8]. Unlike conventional mesh and point-cloud representations, the HKS characterizes the shape up to isometry, making it ideal for non-rigid shape comparison. It has been rapidly applied to the state-of-the-art shape google [1, 3, 7]. Given a vocabulary of geometric words, the shape google computes frequencies of words over the entire shape, which costs a lot of computation.

The aforementioned challenges hinge upon representation paradigms. Recently, bag-of-words (BoW) methods prevail in shape retrieval, coincident with the trend in image retrieval. It can be traced back to the previous work of shape topics [6]. In [10], a part-based representation was utilized by partitioning the model into subparts. In [9], a descriptor was integrated into a BoW approach, which is an indexed collection of closed curves on the 3D surface. In [5], uniform sampling and local spectral descriptor were adopted for partial shape retrieval. The shape google, originally proposed by Ovsjanikov et al. [7], employs the HKS as a shape descriptor, and computes frequencies of words in a vocabulary. The HKS is a concise and informative representation, which preserves all information about the intrinsic geometry of the shape. Later, a scale-invariant heat kernel signature (SI-HKS) [3] was proposed to solve scale changes for this approach. In [1, 7], the shape google also introduced the spatially-sensitive bag-of-words (SS-BoW) by looking at frequencies of word pairs, with encoded spatial relations.

In this paper, we are motivated by the urgent need for a concise and spatially-informative representation for shape comparison and retrieval. The main contribution of this paper is a new paradigm, called bag-of-feature-graphs (BoFG). The key idea is to construct graphs of features on the shape. Given a vocabulary of geometric words, corresponding to each word we build a graph that records spatial information between features, weighted by their similarities to this word. Specific characteristics of the BoFG include:

- It is concise by significantly reducing the number of points involved in representation, and thus, is fast to compute.

- It explicitly records spatial information among features.

- It is representative, since features are salient points containing important information of the shape.

- Graphs have different dominating features associated with corresponding words. This greatly improves the accuracy of shape comparison by eliminating unlikely word-distributions.

We adopt the HKS for feature descriptor (though other spatially-sensitive feature descriptors are also applicable), and heat kernel matrices for graph representation. The heat kernel, intrinsically relevant to the partial differential equation and random walks, is invariant to iso-
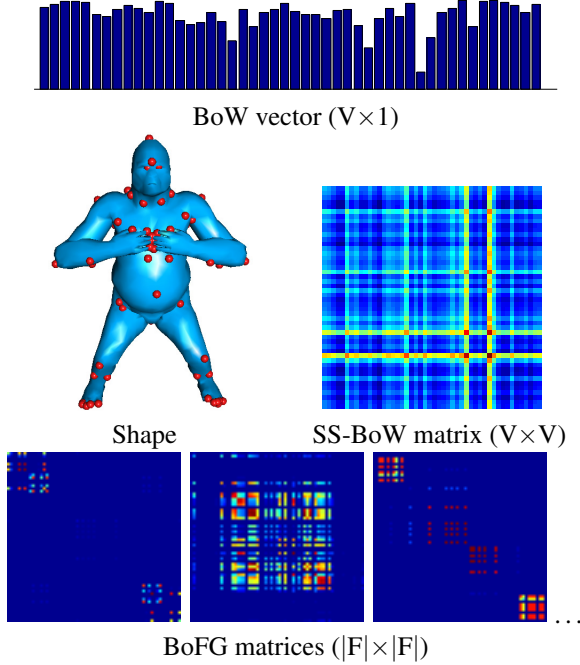
BoW vector (V×1)

Shape   SS-BoW matrix (V×V)

BoFG matrices (|F|×|F|)

Figure 1. Representations of a give shape.

metric deformations and resilient to noise. Various experiments are conducted to evaluate the performance of our method. Quantitative measurements imply that our method has competitive results in comparison with some state-of-the-art methods, and is much faster to compute.

2. BoFG for shape retrieval

2.1. Shape-google revisit

We first revisit the shape google originally introduced by [7]. It utilize a HKS-based BoW. The HKS is defined as the amount of heat transferred from a point x to itself at time t: $h_t(x, x)$, with

$$h_t(x, y) = \sum_{l=0}^{\infty} e^{-\lambda_l t} \phi_l(x) \phi_l(y), \qquad (1)$$

where λ_l and ϕ_l are the l-th eigenvalue and eigenfunction of the Laplace-Beltrami operator. The HKS descriptor $K(x)$ is a vector of HKS probed at different values of t. Let $\mathbf{W} = \{W_1, \ldots, W_V\}$ be a vocabulary of geometric words with size V. The words $\{W_i\}$ are representative HKS vectors in the descriptor space clustered by the k-means algorithm. For each point x, the shape google computes its word distribution $\Theta(x) = [\theta_1(x), \ldots, \theta_V(x)]^T$. The similarity of x

and word W_i is given by

$$\theta_i(x) = c(x) e^{-\frac{\|K(x) - W_i\|^2}{2\sigma^2}}, \qquad (2)$$

where σ is a parameter, and $c(x)$ is a constant for normalization. The BoW descriptor of a surface M is computed by integrating word similarities over the entire shape

$$f(M) = \int_M \Theta(x) d\mu(x), \qquad (3)$$

where $\mu(x)$ denotes the surface area of x. As shown in Fig. 1, the BoW descriptor is a V×1 vector that measures the frequencies of words appearing on the shape. The shape google also introduced a SS-BoW descriptor, given by

$$F(M) = \int_{M \times M} \Theta(x) \Theta^T(y) h_t(x, y) d\mu(x) d\mu(y). \qquad (4)$$

As shown in Fig. 1, it is a V×V matrix that measures frequencies of word pairs. Assume the time complexity for computing a HKS descriptor is O(D). For a shape with N points, the time complexity of BoW is O(ND), and SS-BoW is O(N²D) which is quadratic to N.

2.2. Bag of feature graphs

To reduce the complexity of the shape google, one needs to reduce the number of points involved in representing the shape. A straightforward solution is to select feature points, which keep most information of the shape geometry. Because of the multi-scale property, HKS features contain geometry information ranging from points in small scales to the entire shape in large scales. However, one concern is that a reduced number of points may not be sufficient to faithfully represent the shape. Therefore, instead of counting word frequencies, we construct graphs on detected features, giving rise to a bag-of-feature-graphs (BoFG) paradigm. The graphs encode spatial relations between features, which contain much more geometry information in representing the shape.

We adopt weighted heat kernel matrices to capture global structures of graphs. Specifically, for a shape M with feature set F, only points $x \in F$ are involved in computing word distributions $\Theta(x)$, which reduces much computation. Features are vector-quantized by a fuzzy classification, which assigns $\theta_i(x)$ portion of similarity to word W_i in the distribution of feature x. The distribution $\Theta(x)$ is computed by Eq. (2) with σ set as a quarter of the average distance of words in the vocabulary. This fuzzy classification reduces ambiguities in graph comparison, and also avoids misclassification in a hard quantization. For a geometric word W_i,
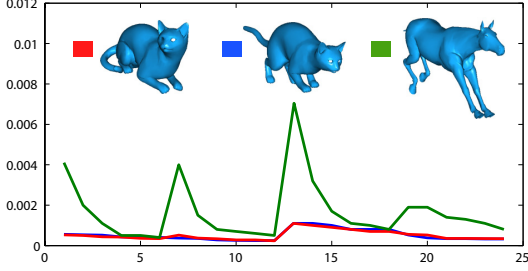
Figure 2. Some non-rigid shapes and their BoFG descriptors.

Figure 3. Time performance (in seconds) of three descriptors on two shapes.

we construct a matrix G_i, whose entry $G_i(x, y)$ with $(x, y) \in F \times F$ is computed by,

$$G_i(x, y) = \theta_i(x)\theta_i(y)h_t(x, y). \qquad (5)$$

It is the heat kernel between x and y weighted by their similarities to the geometric word W_i.

The matrix set $\mathbf{G}(M) = \{G_1, \ldots, G_V\}$ comprises a BoFG representation of the shape M. As shown in the bottom row of Fig. 1, matrices characterize spatial information of features assigned to different word categories. The near-zero entries in a matrix indicate they are hardly classified to this category, and therefore, not considered in this graph. It contains all the geometric information of features in a multi-scale way, which faithfully characterizes the shape. The computation complexity for this matrix representation is O($|$F$|^2$D), as the computed heat kernels can be shared by all matrices. Considering the size of feature set is always much less than the total number of points on the shape, the BoFG is much faster than the shape google.

2.3 Shape retrieval

The mechanism of shape retrieval is to build concise BoFG descriptors of shape models in a database in a off-line process, and retrieve related shapes for a query one by the approximate nearest neighbor (ANN) search. The BoFG descriptor consists of significant eigenvalues of BoFG matrices. Each G_i is a real symmetric matrix, whose eigenvalues are all real and eigenvectors are perpendicular to each other. We choose its six largest eigenvalues denoted as $S_i(M)$, which contributes to a 6V×1 vector $[S_1(M), \ldots, S_V(M)]^T$ as a concise descriptor. This reduces the dimension of the matrix by multi-dimensional scaling (MDS) [2]. Fig. 2 shows some non-rigid shapes and their BoFG descriptors. The deformed cat-models have very similar BoFG descriptors, while the horse-model has a quite different one.
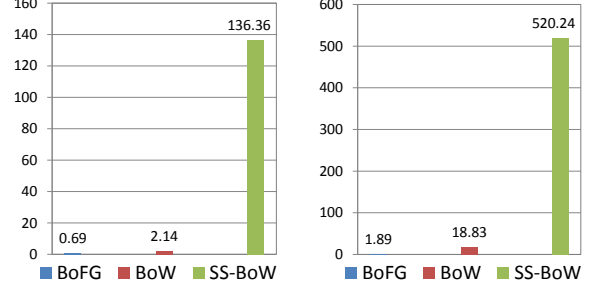
It projects the matrix to its main directions with coordinates leaving in $S_i(M)$, which are stable to a small amount of outliers. Then, we define the similarity distance between two shapes M_1 and M_2 as

$$d(M_1, M_2) = \sum_{i=1}^{V} \|S_i(M_1) - S_i(M_2)\|_2. \qquad (6)$$

The above distance is based on one-scale heat kernels, which can be easily extended to multi-scale by averaging distances of heat kernels at different values of t.

3. Experimental results

We conducted various experiments of shape retrieval to evaluate the proposed method. The test database includes non-rigid shapes of the TOSCA [1] dataset as positives, and shapes from the Princeton Shape Benchmark [2] as negatives. The TOSCA database contains 12 classes of a total 148 non-rigid shapes.

First, we compare the time performance of computing a descriptor of BoW, SS-BoW, and BoFG. For a query shape, one needs to compute its descriptor first to intimate the retrieval. Fig. 3 shows the time performances of three descriptors on a shape with 3k vertices (Left) and another one with 30k vertices (Right). The feature numbers involved in BoFG for two shapes are 42 and 98, respectively. The time for computing Laplace-Beltrami eigenfunctions are excluded, since it is shared by all three methods. By reducing the number of points used in computation, the BoFG significantly improves the time performance of computing shape descriptors. The improvement is more significant when the ratio of points to features is greater.

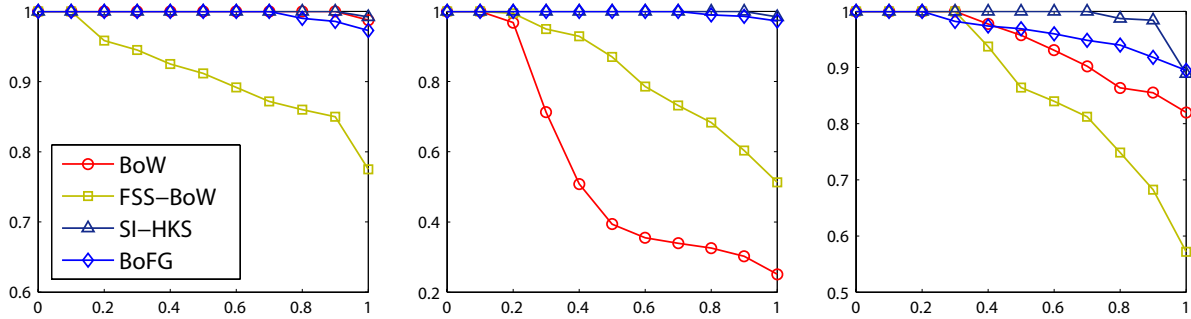The query shapes are obtained from the positives of the database. To test the methods under some challeng-

---

[1] http://tosca.cs.technion.ac.il/book/shrec.html

[2] http://shape.cs.princeton.edu/benchmark/

Figure 4. Precision-recall curves of evaluated methods, with categories of (from *Left* to *Right*) null, scale change, and hole.

ing cases, we apply transformations to the query shapes. This leads to categorized experiments, including null (no transformation), scale change (scaling vertex coordinates), and hole (topological change and missing information). For comparison purpose, we also evaluate some state-of-the-art methods that are similar to ours, including the BoW shape google, the SS-BoW shape google, and the SI-HKS. Since the SS-BoW runs extremely slow, we use over 100 features in its implementation, denoted as FSS-BoW. The three methods share a vocabulary with 48 words. For the BoFG, the vocabulary size depends on the diversity of models in the database, and the number of features usually identified on a shape. Here, we use about 30 to 50 features for one shape, and the vocabulary size is 4.

The methods are quantitatively evaluated by the *precision-recall* (PR) curve that is often adopted for evaluating retrieval performance [4]. It plots the trade-off between precision (ratio of the number of relevant shapes retrieved and the total number of shapes retrieved) and recall (ratio of the number of relevant shapes retrieved and the total number of existing relevant shapes that could be ideally retrieved). Fig. 4 plots the PR curves of evaluated methods, with categories of null, scale change, and hole. The BoFG has competitive results comparing with some state-of-the-art methods.

4. Conclusion

As a new and powerful paradigm for shape representation, the BoFG has demonstrated its feasibility, effectiveness, and efficiency through the above experiments. It offers a concise and faithful representation for shape comparison and retrieval. For the immediate future work, we plan to investigate the problem of graph comparison with heavy outliers. This will help us solve partial shape retrieval. We are also aiming to

improve the performance of our method, especially the efficiency for partial shape retrieval.

References

[1] A. M. Bronstein, M. M. Bronstein, L. J. Guibas, and M. Ovsjanikov. Shape google: Geometric words and expressions for invariant shape retrieval. *ACM TOG*, 2011.

[2] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Generalized multidimensional scaling: A framework for isometry-invariant partial surface matching. *PNAS*, 2006.

[3] M. M. Bronstein and I. Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *CVPR*, 2010.

[4] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs. A search engine for 3d models. *ACM TOG*, 2003.

[5] G. Lavoué. Bag of words and local spectral descriptor for 3d partial shape retrieval. In *3DOR*, 2011.

[6] Y. Liu, H. Zha, and H. Qin. Shape topics: A compact representation and new algorithms for 3d partial shape retrieval. In *CVPR*, 2006.

[7] M. Ovsjanikov, A. M. Bronstein, L. J. Guibas, and M. M. Bronstein. Shape google: a computer vision approach to invariant shape retrieval. In *NORDIA*, 2009.

[8] J. Sun, M. Ovsjanikov, and L. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *SGP*, 2009.

[9] H. Tabia, M. Daoudi, J. P. Vandeborre, and O. Colot. Local visual patch for 3d shape retrieval. In *3DOR*, 2010.

[10] R. Toldo, U. Castellani, and A. Fusiello. The bag of words approach for retrieval and categorization of 3d objects. *The Visual Computer*, 2010.