Eurographics Workshop on 3D Object Retrieval (2012)
M. Spagnuolo, M. Bronstein, A. Bronstein, and A. Ferreira (Editors)

Sketch-based 3D Model Retrieval using Keyshapes for Global and Local Representation

J. M. Saavedra[1] and B. Bustos[1] and T. Schreck[2] and S. Yoon[3] and M. Scherer[4]

[1]KDW-PRISMA Research Group, Department of Computer Science, University of Chile, Chile
[2]Data Analysis and Visualization Group, University of Konstanz, Germany
[3]Yonsei Institute of Convergence Technology, Yonsei University, Korea
[4] GRIS, TU Darmstadt, Germany

Abstract

Since 3D models are becoming more popular, the need for effective methods capable of retrieving 3D models is becoming crucial. Current methods require an example 3D model as query. However, in many cases, such a query is not easy to get. An alternative is using a hand-drawn sketch as query. In this work, we present a new keyshape *based approach named HKO-KASD for retrieving 3D models using rough sketches as queries. Our approach comprises two general steps. First, a global descriptor is used to determine the appropriate viewpoint for each model. Second, we apply a local matching process to determine the final ranking for an input sketch. To this end, we present a local descriptor capable of working with sketch representations. The global descriptors as well as the local descriptors rely on a set of keyshapes precomputed from 2D representations of 3D models and from the query sketch as well. We evaluate our method using the first-tier precision and compare it with current approaches (HELO, STELA). Our results show a significant increase in precision for many classes of 3D models.*

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Applications—3D Model Retrieval

1. Introduction

In the last years there has been wide interest and progress on computer aided retrieval of multimedia data. The advances in this area have allowed users to look for a multimedia object in large repositories in a more efficient way. As advances in multimedia retrieval increase, new interesting applications come up. One of the current interesting applications is 3D model retrieval with impact extending from design to medical issues [BKSS07].

In the context of content-based 3D model retrieval users require a 3D model as an example for querying. However, this kind of query is not always available, frequently requiring some kind of technical expertise to produce it. This fact clearly limits the 3D model retrieval usability.

An easy alternative for querying is simply drawing a 2D stroke-based sketch lacking of color and texture. Although such a kind of sketch is composed of few strokes, it still keeps enough structural information representing what the user is looking for. This kind of querying leads to *sketch-based 3D model retrieval*.

Our contribution in this paper is to present a new approach for sketch-based 3D model retrieval. Our proposal, named HKO-KASD (*Histogram of Keyshape Orientations - Keyshape Angular Spatial Descriptor*), uses a filter-refine strategy for the retrieving task. The filtering step is carried out by means of the proposed global descriptor named HKO, and the refining step is carried out using a local approach using the proposed KASD descriptor.

The results of our proposal show an increase in precision for many classes of 3D models with respect to current strategies applied for sketch-based 3D model retrieval. Particularly, our method achieves significant improvement over 3D models with a well defined structure as shown later.

The remaining part of this document is organized as follows. Section 2 briefly presents related work in the area of 3D model retrieval. Section 3 describes our approach in detail, discussing the proposed global descriptor and the local matching. Section 4 discusses the conducted experiments and analyses the achieved results. Finally, Section 5 presents some conclusions.

DOI: 10.2312/3DOR/3DOR12/047-050

2. Related Work

Sketch-based retrieval is a young and challenging area not only in the case of 3D model retrieval but also in image retrieval systems [EHBA11]. The general idea for comparing a sketch (a 2D object) against 3D models is representing each 3D model by a set of 2D projection images. An interesting work in this context is the proposal of Yoon et al. [YSSK10]. They used suggestive contour images as projections from 14 different viewpoints. Then, they applied the HOG [DT05] descriptor, and finally, they compare HOG descriptors using a metric function.

Saavedra et al. [SBSS11] also use suggestive contour images to get 2D projections. For comparing sketches against suggestive contour images they propose to use a structural-based local approach called STELA, which uses the HELO descriptor [SB10] in a filtering step.

HOG and HELO descriptors compute a set of local orientations to form an orientation histogram. These techniques compute orientations using gradient information which are very sensitive to noise and outliers. In addition, although STELA is a keyshape-based local descriptor, this does not discriminate between different types of *keyshapes*.

Following the proposal of Saavedra et al., we propose a more robust approach for retrieving 3D models having a rough sketch as query. We use a filter-refine strategy by means of global and local descriptors. We use a global descriptor to determine the appropriate suggestive contour image for each 3D model. To this end, we propose the *Histogram of Keyshape Orientations* (HKO) as a global descriptor, avoiding using gradient information that may be sensitive to noise. For the local approach used in the refining step, we propose to use *Keyshape Angular Spatial Descriptor*, a new descriptor (KASD) taking into account the spatial distribution of *keyshapes* around a referent *keyshape* where an angular-partitioned local region is defined.

3. HKO-KASD Approach

A 3D model is transformed into a set of projections from different viewpoints. We use 14 suggestive contour images for each 3D model, each corresponding to a specific viewpoint as specified by Yoon et al. [YSSK10]. Our approach comprises two stages. The first stage is a filtering step which determines the most appropriate projection for each 3D model having a query sketch. This is carried out by means of a global descriptor (HKO), that we propose in this work.

The next stage is a refining step, which will determine the final ranking. We want to exploit the structural information provided by sketches or contour images. In this way, we propose to use *keyshapes* as structural components, as in the work of Saavedra et al. [SBSS11]. In addition, to take into account locality information, we compute a local descriptor for each *keyshape*. We propose a new local descriptor named KASD representing the *keyshape* spatial distribution around a referent *keyshape*.

3.1. Keyshapes

Let I be a sketch or suggestive contour described as an edge map. We represent I by a set of strokes $I = \{S_1, S_2, \ldots, S_{N_s}\}$. Since we do not have information about the start and end point of real strokes, we approach them by means of *edgelinks*.

An *edgelink* is a sequence of edge points starting or ending on a terminal point or branch point. In a formal way, let S_k be a *endgelink* representing a stroke, $S_k = x_1, x_2, \ldots, x_{T_k}$, where x_1 and x_{T_k} belong to the set of terminal and branch points of the underlying image. In Figure 1 we show an example of a sequence of *edgelinks*. Henceforth, an *edgelink* will be called simply **stroke**.

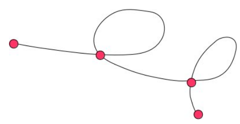

Figure 1: *A synthetic example of a sequence of* edgelinks*.*

After representing an image I as a set of strokes, we proceed to detect *keyshapes*. We only consider straight lines as *keyshapes*. Although lines are simple shapes, these still keep enough structural information which will be exploited by our descriptors. Then, a stroke S of I is decomposed into a set of straight lines. Therefore, image I also is represented by the set of lines. Thus, we define $I = l_1, l_2, \ldots l_n$, where n is the total number of detected lines or *keyshapes* and l_i is a detected straight line.

Once *keyshapes* have been detected, we classify them as horizontal line (H), vertical line (V), diagonal line with slope 1 (D_1), or diagonal line with slope -1 (D_2). In Figure 2 we displayed an example of a suggestive contour image with its corresponding *keyshape* representation.

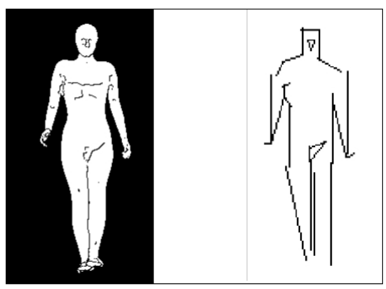

Figure 2: *A suggestive contour image with its corresponding* keyshape *representation.*

3.2. Global Approach

We compute a histogram of *keyshape* orientations (HKO) made up with the orientations of the lines detected previously. The orientation of a line varies from 0 to π using an 8-bin quantization. The final descriptor is the corresponding unitary version of the HKO descriptor.

3.3. Local Approach

Sketches are characterized by keeping structural information. This represents the components of an object. For instance, a hand-drawing of a chair is composed of vertical lines representing legs and horizontal line representing its body. Moreover, structural information also represents a spatial relationship between components. For instance, the body of a table is drawing over its corresponding legs. We are interested in our local descriptor exploits such as information.

3.3.1. Getting Local Descriptors

We named our descriptor as *Keyshape Angular Spatial Distribution* (KASD) because it takes into account the spatial distribution of the *keyshapes* around a referent *keyshape* where an angular-partitioned local circular region is defined.

Let L_R be a referent *keyshape*, we define a circular local region around L_R. The radius of the local region is three times the referent *keyshape* length. In addition, the local region is divided in angular partitions (slices). Each partition corresponds to a spatial region around L_R. An example of a local region and its partitions is depicted in Figure 3 (a).

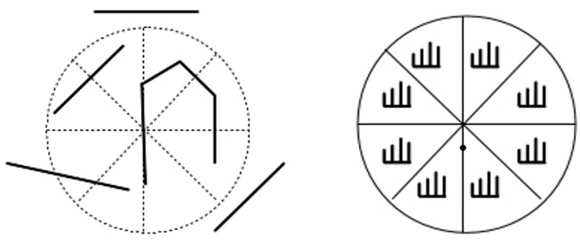

Figure 3: *(a) Local region around a referent* keyshape. *(b) Local descriptor and its 4-bin histograms for each slice.*

Having partitioned the local region of L_R, we proceed to compute a 4-bin histogram for each partition (see Figure 3) (b). This histogram represents the distribution of *keyshape* types around L_R computed for each edge pixel. We use a histogram with 4 bins because of the four classes of keyshapes. Finally, the local descriptor is the unitary version of the juxtaposition of the eight histograms.

3.3.2. Matching

Let I be an image (sketch or suggestive contour). We define $LD(I)$ as:

$$LD(I) = \bigcup_{t \in \{H,V,D_1,D_2\}} LD_t(I) \quad (1)$$

where $LD_H(I)$ is a set of local descriptor for horizontal lines in I, $LD_V(I)$ for vertical lines, $LD_{D_1}(I)$ for diagonal lines with slope 1, and $LD_{D_2}(I)$ for diagonal lines with slope -1.

We compare an input sketch against a suggestive contour for each 3D model. The matching process is carried out between local descriptors corresponding to the same *keyshape* type using the Hungarian Method [Kuh10]. The cost of matching two local descriptors is measured by the L_1 metric. The final match is the union of the partial matches, and the cost of the match is defined as the average cost of all matches.

Finally, the dissimilarity value $DV(S,C)$ between an input sketch S and a suggestive contour image C is defined as $DV(S,C) = AC(S,C)/NM(S,C)$, where, AC indicates the average matching cost and NM is the number of matches after comparing S and C.

4. Experimental Results

4.1. Dataset Description

For our experiments, we used the same benchmark as used by Yoon et al. [YSSK10] which consists of 260 3D models divided homogeneously in 13 classes (*ant*, *bear*, *bird*, *chair*, *cup*, *fish*, *glasses*, *hand*, *human*, *octopus*, *plane*, *table*, *tool*). Additionally, the benchmark provides 250 user hand-drawn sketches, which are used as input for the retrieval task evaluation. It is worth mentioning that the sketches are, in fact, rough sketches drawn by users in a free way. No constraint are imposed to the users for drawing such sketches.

4.2. Result Analysis

To measure the performance of our method we use the *first-tier precision* metric. We compare our results with recent proposals in the context of sketch-based 3D model retrieval. In particular we present results agains HELO [SB10] and STELA [SBSS11] achieving significant increasing in effectiveness for the majority of classes.

Table 1 shows the results of our method (HKO-KASD). The result are presented independently for each class. We increase the precision for 7 of 13 classes, achieving significant improvement for *ant*, *bird*, *chair*, *cup*, *glasses*, *octopus* and *tool*. For instance, the achieved precision for *cup* is 72% better than that gained by STELA method, and the achieved precision for *tool* is 248% better. A graphical comparison in terms of first-tier precision is depicted in Figure 4.

Two examples of the retrieval task are shown in Figure

Table 1: *First-tier precision for each class.*

Class	HELO	STELA	HKO-KASD
Ant	0.147	0.126	**0.153**
Bear	0.210	0.338	0.135
Bird	0.107	0.110	**0.121**
Chair	0.088	0.121	**0.167**
Cup	0.138	0.142	**0.244**
Fish	0.162	0.152	0.124
Glasses	0.029	0.079	**0.109**
Hand	0.333	0.319	0.252
Human	0.255	0.321	0.180
Octopus	0.108	0.150	**0.195**
Plane	0.021	0.117	0.088
Table	0.135	0.120	0.110
Tool	0.079	0.045	**0.157**

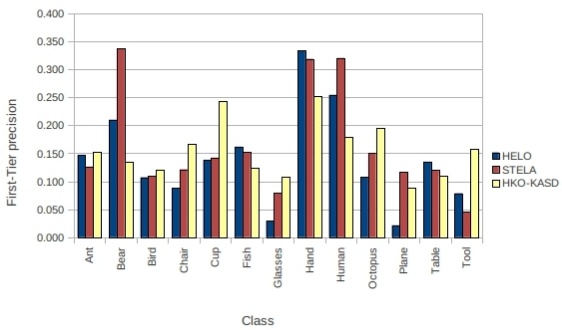

Figure 4: *The first-tier precision for each class. We compare HKO-KASD with HELO and STELA methods.*

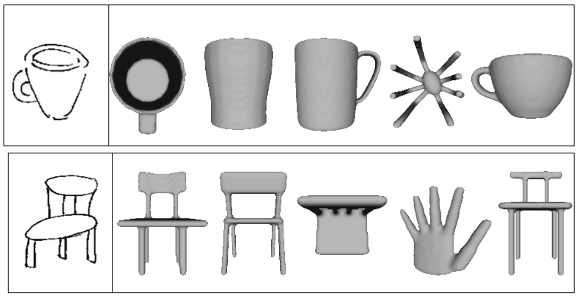

Figure 5: *Examples of using the proposed local approach.*

5. In these examples we depict only the first five retrieved models. We can note that our proposal retrieves only one false positive among the first five retrieved objects, the remaining four objects correspond to objects belonging to the same class of the query.

It is important to note two aspects in our method. First, our method retrieves relevant objects even though they have different viewpoints. Second, our method can retrieve relevant objects even though such as objects, belonging to the same class, undergo shape variations

5. Conclusions

We have presented a new approach for retrieving 3D model using sketches as queries. In this context, traditional methods do not work appropriately since sketches lack color and texture. However, sketches provide structural information that defines how an object is composed of. We use a global descriptor as the filtering stage and a local descriptor for the refining stage. Both stages rely on structural components called *keyshapes*. This proposal takes into account the structural information that sketches provide.

Our results show increasing in effectiveness for many classes of the benchmark database achieving an improvement in more than 50% for some classes.

6. Acknowledgment

This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the "IT Consilience Creative Program" support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2010-C1515-1001-0001)

References

[BKSS07] Bustos B., Keim D., Saupe D., Schreck T.: Content-based 3D object retrieval. *IEEE Computer Graphics and Applications 27*, 4 (2007), 22–27. 1

[DT05] Dalal N., Triggs B.: Histograms of oriented gradients for human detection. In *Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01* (2005), pp. 886–893. 2

[EHBA11] Eitz M., Hildebrand K., Boubekeur T., Alexa M.: Sketch-based image retrieval: Benchmark and bag-of-features descriptors. *IEEE Transactions on Visualization and Computer Graphics 17*, 11 (2011), 1624–1636. 2

[Kuh10] Kuhn H. W.: The hungarian method for the assignment problem. In *50 Years of Integer Programming 1958-2008*. 2010, pp. 29–47. 3

[SB10] Saavedra J., Bustos B.: An improved histogram of edge local orientations for sketch-based image retrieval. In *Pattern Recognition*, vol. 6376 of *Lec. Notes in Computer Science*. 2010, pp. 432–441. 2, 3

[SBSS11] Saavedra J. M., Bustos B., Scherer M., Schreck T.: Stela: sketch-based 3d model retrieval using a structure-based local approach. In *Proceedings of the 1st ACM International Conference on Multimedia Retrieval* (2011), ICMR '11, ACM, pp. 26:1–26:8. 2, 3

[YSSK10] Yoon S. M., Scherer M., Schreck T., Kuijper A.: Sketch-based 3D model retrieval using diffusion tensor fields of suggestive contours. In *Proc. of the international conference on Multimedia* (2010), pp. 193–200. 2, 3