

SEMANTIC SKETCH-BASED 3D MODEL RETRIEVAL

Bo Li, Yijuan Lu, Ribel Fares

Department of Computer Science, Texas State University
{b_l58, lu, rf1190}@txstate.edu

ABSTRACT

Query-by-Sketch is an intuitive way to retrieve relevant 3D models given a user sketch. It has many promising application potentials in human computer interaction, 3D animation, game design, *etc.* However, it is a very challenging task due to the big semantic gap between human-drawn sketches and 3D models. A human-drawn sketch is usually composed of several simplified and exaggerated curves as the iconic representation of an object, while a 3D model of the object is generally an accurate representation of its geometry information. Such big semantic gap makes the search based on a direct 2D-3D comparison suffer low accuracy and high computational cost. In this paper, we propose a novel semantic sketch-based 3D model search to bridge the semantic gap by first recognizing the potential semantic meanings of the user sketch and then performing 2D-3D matching for the 3D models within the predicted categories. The experimental results demonstrate that our approach achieves significant improvements in both search accuracy and efficiency, which further validate that our approach can effectively bridge the semantic gap between user sketches and 3D models.

Index Terms— Sketch-based 3D model retrieval, semantic information, sketch recognition, machine learning

1. INTRODUCTION

Recently, sketch-based 3D model retrieval attracts a lot of attention due to its promising application potentials in 3D animation, human computer interaction, and game design. However, it is a very challenging task since there exists a big semantic gap between user sketches and 3D models in the database. Human sketches always have arbitrary styles, iconic representations in 2D space, high-level abstraction, and drastic simplification, which bring a lot of difficulties in sketch description and representation. 3D model of an object is generally an accurate representation of its geometry information. Such big semantic gap makes direct comparison between 2D sketches and 3D models almost impossible.

In the last 10 years, more than a dozen of sketch-based 3D model retrieval algorithms have been proposed [4, 1, 5, 2] to bridge such semantic gap. These methods can be categorized into two groups. The first group of methods tries to bridge

the gap by rendering a large number of 2D images from 3D models and matching 2D sketch images with 2D rendering images. However, such kinds of methods usually suffer from high computational cost and low retrieval accuracy due to two major reasons: 1) A large number of rendering images need to be compared, which may involve many noisy (irrelevant) models; 2) 2D rendering images have comprehensive geometric representation, which is still different from sketches' iconic representations in 2D space. The second group of methods is trying to reduce the gap by first generating a 3D model based on a user sketch and matching the generated 3D model with all the 3D models in the database directly. Since it is difficult to generate an accurate 3D model based on a simple user sketch drawn from one view, the search performance of such kind of methods is still unsatisfactory. In summary, all of these methods fail to bridge the semantic gap without considering the semantic information of either the sketch queries or the target 3D models.

Motivated by the above findings, in this paper, we propose a novel semantic sketch-based 3D model search to bridge the semantic gap. Our approach tries to understand the semantic meaning users are expressing through their sketches before searching the corresponding 3D models. By building an intelligent sketch recognizer, our approach can first predict the potential semantics of the user sketch. Then by searching the 3D models in the predicted semantic categories, the best matchings can be found. The experimental results demonstrate that our approach achieves significant improvements in both search accuracy and efficiency, which further validates that our approach can effectively bridge the semantic gap between user sketches and 3D models. The main contributions introduced in this paper are summarized as follows:

- 1) We conduct a comprehensive study of the semantic gap between user sketches and 3D models and propose a novel semantic sketch-based 3D model search algorithm to bridge such semantic gap.
- 2) We develop an intelligent sketch recognizer through supervised learning to correctly capture the semantic meanings of users' sketches.
- 3) Our work will explicitly guide the research in human sketch recognition and opens broad research possibility for sketch-based 3D model retrieval and annotation.

2. RELATED WORK

In this section, we review existing sketch-based 3D model retrieval techniques according to the categorization in Section 1.

Most existing sketch-based 3D model retrieval algorithms fall in the first group and they can be further classified into two subcategories: (1) local feature-based approach, such as Eitz et al.’s [4] local feature and Bag-of-Words based method BOF-SBR; (2) global feature-based approach, such as Li et al.’s [4, 5] 2D-3D alignment based method SBR-2D-3D. To speed up the retrieval process, the first subcategory extracts local features and then usually clusters and quantizes them based on the Bag-of-Words model, which is fast but also non-trivially reduces the discriminability of the local features because of losing/decreasing the spatial and structural relationships among the local features during the quantization process. On the contrary, the global feature-based approach performs a global feature matching, which can be more accurate but often obviously slower. For example, in the SHREC’12 Sketch-Based 3D Shape Retrieval Contest [4], SBR-2D-3D has achieved the best accuracy while it is the slowest as well. BOF-SBR is the fastest while its performance and robustness are apparently inferior if compared with SBR-2D-3D.

Cao et al. [2] is one of representative examples for the second group. It first reconstructs a planar or curved 3D query surface based on a 2D line drawing drawn by a user. However, no database level retrieval performance has been provided.

In a word, without considering the semantic information of either the sketch queries or the target 3D models, all the methods above fail to bridge the semantic gap between the sketches and 3D models. Our proposed semantic sketch-based 3D model retrieval makes the first attempt to use the semantic classification information to reduce the semantic gap, as well as to adequately utilize the better-performing global feature matching to improve the retrieval efficiency.

3. SEMANTIC SKETCH-BASED 3D RETRIEVAL

The framework of the proposed semantic sketch-based retrieval approach is shown in Fig. 1. It consists of two stages: sketch recognition training stage and sketch-based retrieval stage. In the sketch recognition training stage, a large sketch training dataset is selected first, which contains sketches from a variety of categories. Then, sketch features are explored and extracted to well describe input sketches’ attributes. And an intelligent sketch classifier is built up to recognize a user sketch into potential sketch categories. In the sketch-based retrieval stage, a query sketch is first fed into the developed sketch classifier and the possibilities of the input sketch belonging to all the categories are predicted. The higher the possibility is, the larger chance the user’s sketch describes the same object. Therefore, the 3D models in the top candidate categories are much closer to the input sketch. So, a general sketch-based 3D model retrieval algorithm is employed

to rank the models in the top L candidate categories. And the models in the remaining categories are ranked based on their corresponding categories’ prediction values. By a complementary utilization of the recognition and retrieval techniques, the proposed approach can successfully bridge the semantic gap between sketches and 3D models.

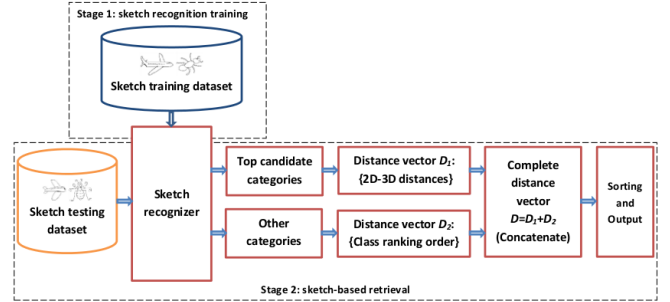


Fig. 1. Our semantic sketch-based retrieval approach.

3.1. Stage 1: Sketch recognition training

(1) **Sketch feature extraction.** Eitz et al. [3] extracted local features from a sketch based on the ideas of Scale-Invariant Feature Transform (SIFT) and Histogram of Gradient (HoG) features and embedded them into a Bag-of-Words framework as the feature representation. This feature representation has two big limitations: 1) it only captures the local information of sketches and totally ignores the global attributes; 2) it cannot handle the rotation of sketches. In this work, a hybrid feature is developed by further integrating a set of rotation-invariant global features for a sketch. The hybrid feature vector is generated by combining the 500-dimensional Bag-of-Words local feature vector in [3] and a 119-dimensional global feature vector devised by us. First, each sketch is resized into a 300×300 image, then the thickness of the sketch lines is shrunk to a single pixel. Then, a global feature vector is extracted from the sketch as shown in Fig. 2. It is comprised of 9 distance histograms: 5 radial distance histograms of the sketch pixels with respect to 5 selected reference points/lines, 2 radial distance histograms of the first intersection points and 2 radial angle histograms of the sketch pixels with respect to the two centers **C** and **FPC**. All the histograms are divided into 11 bins and the mean and standard deviation values of each histogram are also included, thus generating a 117-dimensional (13×9) global feature vector. In addition, the distance between **C** and **FPC**, and the sum of the distances between sketch pixels and **FPC** are also considered. Our experimental results show that a moderate improvement (around 3%~5%) in sketch recognition has been achieved after incorporating these global features.

(2) **Sketch classifier training.** Similar as [3], Support Vector Machine is chosen to build a sketch recognition model. The same parameter settings in [3] are utilized, such as local

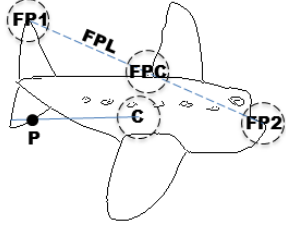


Fig. 2. Illustration of our 5 reference points or lines for the global features: **C**, **FP1**, **FP2**, **FPL** and **FPC**, where **C** is the centroid of a sketch, **FP1** and **FP2** are the two farthest points with respect to the centroid **C**, **FPL** is the line between the two farthest points; and **FPC** is the center of the **FPL** line. **P** is an example of first intersection point.

feature definitions, “soft” kernel-codebook coding choice, vocabulary size, and 3-fold cross-validation selection except for RBF kernel ($\gamma=0.1$ and $C=20$).

3.2. Stage 2: Sketch-based retrieval

Given a query sketch q and a target 3D dataset $M = \{m_i\}$, a distance vector D needs to be generated to measure the dissimilarities between q and all the models in M , as follows.

(3) **Sketch classification.** A query sketch is fed into the developed sketch classifier. The possibilities of the input sketch belonging to all the categories are predicted and the top L candidate categories are found.

(4) **2D-3D matching.** A sketch-based retrieval algorithm is applied on all the models in the top L candidate categories and the distances between the models and the input sketch, named D_1 , are calculated. In this paper, the state-of-the-art sketch-based retrieval approach SBR-2D-3D [4, 5] is selected. The proposed semantic sketch-based retrieval algorithm that incorporates SVM-based sketch recognition into the SBR-2D-3D algorithm is named as **SBR_SVM-2D-3D**.

(5) **Distance vector generation.** It is to assign distances between the input sketch and the models in the left categories as the second part of D , named D_2 . In this paper, D_2 of these models is set as the ranking orders of their respective categories. After we normalize the values in D_1 into $[0,1]$, D_1 and D_2 are concatenated sequentially into one vector D .

(6) **Ranking and output.** All the distances in D are sorted and the relevant models are listed accordingly.

4. EXPERIMENTS AND DISCUSSION

In this section, we test our approach on a latest sketch-based 3D model retrieval benchmark, named SHREC’13 Sketch Track Benchmark [1]. To measure the sketch recognition performance, eight popular metrics are utilized including True Positive rate (TP), False Positive rate (FP), Precision (P), Recall (R), F-Measure (F), Matthews Correlation Coefficient

(MCC), area under the Receiver Operating Characteristic Curve (ROC), and area under Precision-Recall Curve (PRC). To comprehensively evaluate the retrieval performance, we select the following performance metrics: Precision-Recall (PR) diagram, Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-Measure (E), Discounted Cumulative Gain (DCG) [6] and Average Precision (AP).

SHREC’13 Sketch Track Benchmark. Recently, a SHREC’13 Sketch Track Benchmark [1], for the Shape Retrieval Contest (SHREC) 2013 Track on the topic of large scale sketch-based retrieval, was developed based on the shared categories of the sketch recognition dataset built by Eitz et al. [3] and the Princeton Shape Benchmark (PSB) benchmark [6]. This benchmark contains 7200 hand-drawn sketches, which are uniformly distributed on 90 classes, and 1258 relevant 3D models selected from the PSB benchmark to form the target 3D dataset. To evaluate retrieval algorithms based on a learning approach, the “Training” and “Testing” datasets are also built by randomly selecting 50 sketches per class for training and the rest 30 sketches for testing, while the complete target model dataset is remained as a whole for both training and testing purpose.

Sketch recognition. The developed sketch recognizer in our approach is tested on the above sketch track benchmark dataset and Eitz et al.’s complete sketch recognition benchmark [3]. Its recognition performance is compared with the newest sketch recognition algorithm LSR (local feature based approach) proposed in [3]. As shown in Table 1, our approach achieves better results in all the metrics on both benchmark datasets. It validates that our sketch recognizer is more robust to the sketch rotation and can well describe user sketches by incorporating global attributes of sketches. For our algorithm, the time taken to recognize the 2700 sketches is 269.95 seconds, thus averagely 0.1 second is needed to classify a sketch.

Table 1. Average classification performance comparison in terms of eight metrics. The first two rows are for the SHREC’13 Sketch Track Benchmark; the last two rows (*) are for the Eitz et al.’s complete sketch benchmark.

	TP	FP	P	R	F	MCC	ROC	PRC
Our	0.613	0.004	0.623	0.613	0.614	0.612	0.982	0.664
LSR	0.594	0.005	0.597	0.594	0.593	0.590	0.974	0.637
Our*	0.545	0.002	0.549	0.545	0.544	0.544	0.772	0.326
LSR*	0.520	0.002	0.523	0.520	0.519	0.518	0.759	0.298

Sketch-based retrieval. The proposed **SBR_SVM-2D-3D** method is tested on the SHREC’13 Sketch Track Benchmark and compared with the participating approaches in the contest [1], including SBR-2D-3D, FDC, and EFSD. In our approach, the number of sample points (NUM) for each sketch is set as 50 (the same setting used in SBR-2D-3D) and a variety of length values (L) for the candidate category list is tested. Fig. 3~4 and Table 2 show the comparison results.

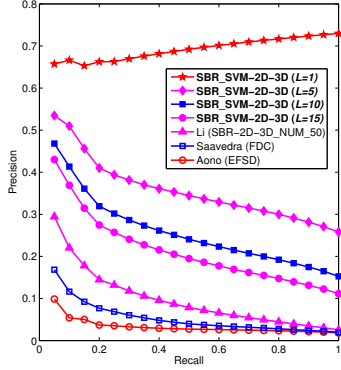


Fig. 3. Precision-Recall diagram performance comparison between our **SBR_SVM-2D-3D** (different L values) method and the participating approaches in the SHREC’13 Sketch Track Contest on the “Testing” dataset.

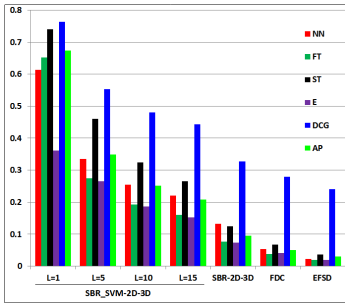


Fig. 4. Other performance metrics comparison between our semantic algorithm **SBR_SVM-2D-3D** (different L values) and the participating approaches in the SHREC’13 Sketch Track Contest on the “Testing” dataset.

(1) Our semantic retrieval approach significantly improves the retrieval performances (100%~700% better accuracy) than other state-of-the-art sketch-based retrieval algorithms. It further validates that our approach can bridge the semantic gap between the diverse query sketches and 3D models effectively. What’s more, the average retrieval time t for each query is also substantially decreased. The retrieval time of our approach includes two parts: sketch recognition and 2D-3D matching. In fact, the time for recognizing a sketch is only about 0.1 second, which adds little online computational load. Thus, our sketch recognizer can help find a much smaller number of 3D models for direct 2D-3D comparison, which not only significantly improves the retrieval

Table 2. Timing information comparison: t is the average response time per query based on a modern computer.

Method	SBR_SVM-2D-3D				SBR-2D-3D	FDC	EFSD
	$L=1$	$L=5$	$L=10$	$L=15$			
t (s)	1.43	4.88	9.32	13.26	43.93	0.02	20.24

accuracy, but also manifestly reduces the computational time.

(2) In this experiment, keeping only the top first candidate category is the best choice in terms of both accuracy and efficiency. However, for other datasets, the situation may be varied due to different sketch recognition performances.

(3) Sketch-based 3D model retrieval based on semantic information will have broad application potentials on the applications which require real-time response time.

5. CONCLUSIONS AND FUTURE WORK

A semantic sketch-based 3D model retrieval algorithm is proposed in this paper by first performing sketch recognition to find a set of candidate categories for the sketch and then applying direct 2D-3D comparison on the models within the candidate classes. It is an important improvement to encompass the semantic gap between the sketches and models. The experimental results demonstrate that the proposed method achieves the significant improvements in both retrieval accuracy and computationally efficiency. The developed sketch recognition algorithm also further improves sketch recognition by integrating global sketch features. As the main future work, we plan to study the integration of unsupervised or semi-supervised sketch recognition module when the label information of the target 3D model dataset is not available.

Acknowledgments

This work is supported by Army Research Office grant W911NF-12-1-0057, NSF CRI 1058724, and Texas State University Research Enhancement Program (REP), to Dr. Yijuan Lu.

6. REFERENCES

- [1] <http://www.itl.nist.gov/iad/vug/sharp/contest/2013/SBR/>, 2013.
- [2] L. Cao, J. Liu, and X. Tang. 3D object retrieval using 2D line drawing and graph based relevance reedback. In *ACM Multimedia*, pages 105–108, 2006.
- [3] M. Eitz, J. Hays, and M. Alexa. How do humans sketch objects? *ACM Trans. Graph.*, 31(4):44, 2012.
- [4] B. Li and et al. SHREC’12 track: Sketch-based 3D shape retrieval. In *3DOR*, pages 109–118, 2012.
- [5] B. Li and H. Johan. Sketch-based 3D model retrieval by incorporating 2D-3D alignment. *Multimedia Tools and Applications*, pages 1–23 (online first version), 2012.
- [6] P. Shilane, P. Min, M. M. Kazhdan, and T. A. Funkhouser. The Princeton shape benchmark. In *SMI*, pages 167–178, 2004.