

# 3D Sketch-Based 3D Model Retrieval

Bo Li, Yijuan Lu, Azeem Ghumman, Bradley Strylowski, Mario Gutierrez,  
Safiyah Sadiq, Scott Forster, Natacha Feola, Travis Bugerin

Department of Computer Science, Texas State University

San Marcos, Texas, USA

{b\_l158, lu, a\_z65, b\_s211, mag262, s\_s451, smf85, n\_f27, tlb165}@txstate.edu

## ABSTRACT

How to draw 3D sketches in a three-dimensional space and how to use a hand-drawn 3D sketch to search similar 3D models are brand new and challenging research topics. In this paper, we make an initial study on 3D sketching and propose a novel 3D sketch-based 3D model retrieval system. Our system allows users to freely draw 3D sketches in the air as well as to find similar 3D models given human-drawn 3D sketches. Promising retrieval performance has been achieved in experiments based on 300 collected 3D sketches and a recent large scale sketch-based 3D shape retrieval benchmark.

## Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval

## Keywords

Sketch-based 3D model retrieval; 3D sketch; Kinect

## 1. MOTIVATION

Since prehistoric times, sketching has been a universal form of communication used by humans to depict the visual world. Today, sketching has become one of the most natural ways to provide a visual search query for search, e.g., one can search images, videos, and 3D models on the Internet by sketching an object or scene on a touch phone/tablet.

Currently, all existing sketch-based 3D model retrieval systems rely on 2D sketching technology, which requires users to draw a sketch on a two-dimensional plane (paper, touch screen). Constraining a sketch to two dimensions limits the 3D information that the shape can communicate, creating a huge semantic gap between the iconic representation of a 2D sketch and the accurate 3D coordinate representation of a 3D model. This gap makes 2D sketch-based 3D model retrieval very challenging [3].

Bothered by this gap, an interesting question is raised: “Can we do 3D sketching such that we can perform 3D

sketch-based 3D model retrieval?” If it were possible for users to sketch an object in all three dimensions in a 3D space (for example, in the air) by utilizing hand gestures, the 3D sketch should provide a better description of the object than a 2D sketch. The 3D sketch not only encodes 3D information, such as depth and features of more facets of the object, but also includes the salient 3D feature lines of its counterpart of 3D models. However, there is a scarcity of comprehensive research supporting methods that allow users to sketch 3D objects in a 3D space. In addition, how to understand (translate) 3D sketches drawn by human hands and how to match 3D sketches with 3D models become new research problems.

In this paper, we perform an initial study of 3D sketching technology and develop a 3D sketching virtual drawing “board”. We investigate 3D sketch and model matching problems and build a novel 3D sketch-based 3D model retrieval system, whose performance is tested on a large scale sketch-based 3D shape retrieval benchmark dataset and a collected human 3D sketch dataset.

To our best knowledge, this work is the first attempt to explore 3D sketching in a 3D space and to develop an innovative retrieval system that enables users to search 3D models based on hand-drawn 3D sketches. The implications of this work could be tremendous, as 3D sketching allows for more direct communication in a user’s drawing, which could not only enhance 3D model retrieval accuracy, but also provide a possibility for other human sketch related applications, such as virtual try-on systems for clothes, glasses and watches. The main contributions introduced in this paper are highlighted as follows:

- A novel 3D sketching virtual drawing “board” (software) is proposed and implemented, which allows users to freely draw 3D sketches in the air (a real 3D space). Based on this developed 3D sketching virtual drawing board, the first human 3D sketch dataset is collected.
- A 3D sketch-based 3D model retrieval system is introduced for the first time to solve the matching problem between 3D sketches and models.
- Our proposed approach shows promising application potential for 3D sketch understanding or recognition, large scale 3D model search, and on-line 3D model shopping, etc.
- Our work will explicitly guide the research in 3D sketching and provide a path for large scale sketch-based image, video, and object retrieval.

## 2. BACKGROUND

Sketch-based 3D model retrieval targets on retrieving 3D models given a hand-drawn query sketch. Recently, sketch-based 3D model retrieval has attracted much attention since it can be widely used in sketch-based rapid prototyping, recognition, mobile 3D search, 3D printing, 3D animation production and etc. Many related algorithms have been proposed [3] which use a 2D hand-drawn sketch as a query. A series of Shape Retrieval Contest (SHREC) tracks on this topic have been held in conjunction with the Eurographics Workshops on 3D Object Retrieval (3DOR) over the past three years. Some new benchmark datasets, such as the large scale SHREC'13 Sketch Track Benchmark (SHREC13STB) [2] which contains 7,200 2D sketches and 1,258 3D models of 90 classes, have been built and released to the public. Due to the semantic gap between the two different representations of rough sketches and accurate 3D models, sketch-based 3D model retrieval remains one of most challenging research topics in the field of 3D model retrieval. In order to bridge the gap, we propose a 3D sketching solution and develop a sketch-based 3D model retrieval system that uses human 3D sketches, which is described in detail in the following sections.

## 3. OUR APPROACH

### 3.1 3D Sketching

How to design a smart, user friendly, and inexpensive 3D sketching virtual drawing “board”, which allows users to easily create 3D sketches, is one of crucial problems of this project. We propose to implement this virtual “board” with the Microsoft Kinect, considering it is a popular and low cost motion sensing input device and offers a built-in color VGA video camera, depth sensor, and multi-array microphone.

We utilize the RGB video camera on the Kinect to allow the user to see exactly what they are drawing in the surrounding 3D space (in the air), and make use of the depth sensor to capture the depth information communicated in the 3D sketches. The video and depth sensors have a pixel resolution of 640x480 and run at 30 frames per second. With this capability, users are able to engage the tracking function of Kinect to monitor the 3D locations of their hands in a 3D space while drawing. To facilitate sketching and retrieval, we also develop a voice-activated Kinect-based 3D sketching Graphical User Interface (GUI) (shown in Fig. 1) that supports the tracking function. The proposed interface can not only track the movement of a user’s hand, but also support voice commands (e.g. start, left, right, pause, resume, search, and reset) allowing a user to pause in the middle of a drawing or completely restart a sketch as well as to switch between left and right hands while drawing. Users can also select front or side view display and point or line mode for sketch viewing. All the above operations can be performed in real-time. In experiments, we have observed that the interactive feedback greatly helps the user in drawing a 3D sketch and improves users’ drawing experience.

Although the Microsoft Kinect is able to detect and track 48 individual points on the user’s body, the captured hand motion data contains a lot of noise, mostly due to the fact that when users perform sketching in the air they could shake and rarely have perfectly steady hands while drawing. The introduced noise can significantly affect the 3D model retrieval performance. To combat this, a Kalman filter is

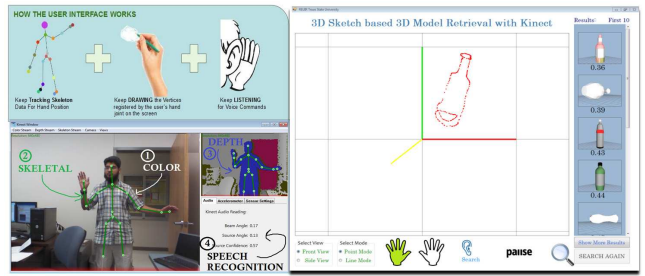


Figure 1: System Graphical User Interface (GUI).

applied, which is an optimal estimator that infers parameters of interest from inaccurate and uncertain observations. In our experiments, we find the Kalman filter successfully filters partial noise and predicts smoother sketch curves.

### 3.2 3D Sketch-Based 3D Model Retrieval System

Based on the above 3D sketching platform, we build an efficient 3D sketch-based 3D model retrieval system. Figure 2 shows the framework of our retrieval system, which contains both online and offline processes and consists of three major components: data processing, feature extraction, and matching.

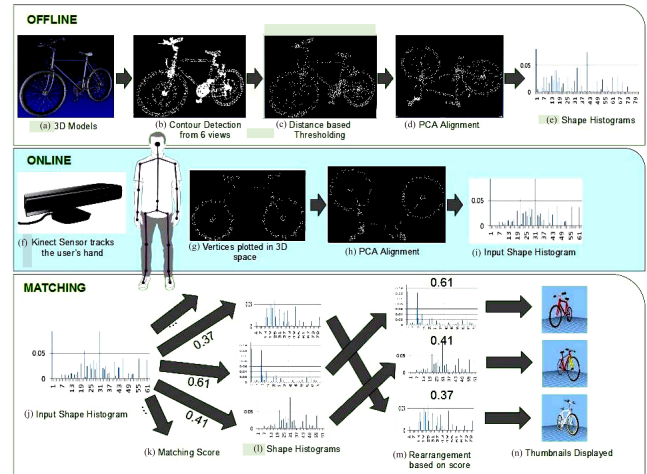


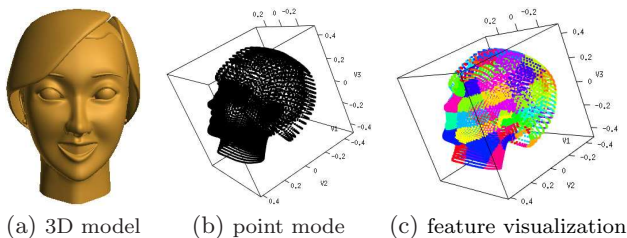
Figure 2: System framework.

(1) **Data processing.** The major task of the data processing component is to generate 3D outlines of 3D models. First, 3D normalization is performed for each 3D model: 3D alignment based on Principle Component Analysis (PCA) is first performed, and then translate the origin to the center of the model’s bounding sphere, and finally scale the model such that the radius of the bounding sphere is 1. Second, the 3D outline for the model is generated by integrating the 3D contour points of its six principle views (front, back, left, right, top, and bottom views). One sample result is demonstrated in Fig. 2 (b). Third, an approximately uniform point sampling is performed by setting a 3D distance threshold between the 3D points. This step is to improve the robustness

of our algorithm w.r.t the different resolutions of 3D outlines and sketches. Fig. 2 (c) shows the final 3D outline of the bicycle model after performing the sampling. Finally, a PCA alignment is executed (Fig. 2 (d)) on the 3D outline. The same normalization process will be applied on an online hand-drawn 3D sketch.

**(2) Feature extraction.** Descriptive visual features need to be extracted from both 3D models and 3D sketches in order to perform effective and efficient sketch-model matching and retrieval. Extensive research has been conducted to generate a number of 3D shape descriptors in characterizing complete and perfect 3D models. However, many of them require a 3D mesh model as input and only a few of them are applicable to represent a 3D sketch which is essentially a sparse point cloud (a very abstract and inaccurate representation of a 3D object). Therefore, a large research space remains unexplored in terms of comparing an inaccurate point cloud of a 3D sketch with an accurate 3D model.

In this paper, we investigate using the 3D shape histogram [1] as a representative feature for both 3D models and 3D sketches considering its descriptiveness, high efficiency, and simplicity. The 3D shape histogram divides the 3D space occupied by a 3D model/sketch into a set of shell, sector or spiderweb bins and counts the percentage of the vertices falling in each bin to form a histogram as the 3D shape histogram descriptor. One visualization example of the process is shown in Fig. 3. Considering the inherent nature of the representations for a 3D sketch and a 3D outline, as well as the efficiency issue, for each 3D sketch or outline, its 3D shape histogram [1] descriptor is extracted based on the spiderweb model (20 shells, 6 sectors, 120 bins in total) (Fig. 2 (e)).



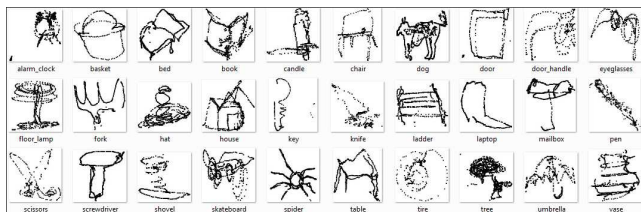
**Figure 3: An example to visualize the 3D shape histogram feature (different bins) of the m349 model in the Princeton Shape Benchmark (PSB) [4] dataset.**

**(3) 3D sketch-3D model matching.** The histogram of the 3D sketch is compared with the histograms of all the 3D outlines generated from the 3D models based on the Euclidean distance (Fig. 2 (k)). Then, the distances are sorted in ascending order (Fig. 2 (m)), and finally the thumbnails of the top ten 3D models are listed accordingly in real-time on the right side of the GUI, as shown in Fig. 2 (n). Users can also browse the next 10 results by saying the voice command “Show more results”.

## 4. EXPERIMENTS AND DISCUSSIONS

### 4.1 Kinect300 Dataset Collection

Based on the developed 3D sketching virtual drawing “board”, we have collected a 3D sketch dataset named Kinect300 which comprises 300 sketches in 30 object categories (see Fig. 4), each with 10 sketches. Seventeen users (4 females



**Figure 4: Example 3D sketches (one example per class, shown in one view) of our Kinect300 dataset.**

and 13 males) with an average age of 21 years participated in the 3D sketch data collection, where each user drew sketches of several categories. Among 17 users, only two of the males have a background in art.

### 4.2 Retrieval Experiments

To comprehensively evaluate the performance of our retrieval system presented in Section 3.2, we performed the following three types of experiments based on six commonly used evaluation metrics [4]: Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-Measure (E), Discounted Cumulative Gain (DCG) and Precision-Recall (PR).

- **outline-model retrieval:** query is a 3D outline generated from a 3D model; targets are 3D models from the *same* dataset. This experiment tests the retrieval performance of searching a similar 3D model given a perfect 3D outline (not a rough human 3D sketch). Here, the SHREC13STB benchmark (target dataset only) is chosen as the target dataset, which contains 1,258 target 3D models of 90 classes.
- **sketch-sketch retrieval:** query is a 3D sketch; targets are 3D sketches from the same dataset. This experiment tests the performance of similar 3D model retrieval given a hand-drawn 3D sketch. Since users sketch objects differently, the variety of human sketches even within categories can be extremely high. We want to test whether sketches within the same category drawn by various users can be retrieved given a single user sketch. Kinect300 is used as the target dataset in this experiment.
- **sketch-model retrieval:** query is a hand-drawn 3D sketch; targets are 3D models of real objects. This experiment tests the performance of similar 3D model retrieval given a human 3D sketch. The Kinect300 3D sketch dataset is used as the query set and the SHREC13STB benchmark (target dataset only) is used as the target dataset. Please note that only 21 of the 30 classes in Kinect300 have relevant 3D models in the SHREC13STB target dataset, thus when computing the retrieval performance we only consider the results of these 21 classes. The 9 classes that have no relevant 3D models are: alarm clock, basket, candle, door handle, eyeglasses, fork, key, pen, and scissors.

Thanks to the high efficiency of the 3D shape histogram and the Kalman Filter, all the above retrieval experiments are performed in real-time. For example, it takes only 1.22 sec to perform a 3D model retrieval given a hand-drawn 3D

sketch (type 3) on a modern computer (CPU: Intel Xeon X5675 @3.07 GHz). The experimental results are listed in Fig. 5 and Table 1. As expected, the performance of outline-model retrieval is better than the performance of sketch-sketch retrieval, which outperforms the sketch-model retrieval. We can observe that the 3D shape histogram has good descriptive ability in capturing the vertex distribution pattern of a 3D model and a perfect 3D outline. But it doesn't have strong discrimination power in differentiating all 3D models coming from different categories. Thus, given a 3D outline, 3D models from irrelevant categories may be returned. We also find that the 3D shape histogram is sensitive to noise. All the 3D sketches we currently collect are quite abstract and noisy since most users have little drawing experience. When users are sketching, they often pause their hands or move their hands back and forth in the air. The sketch lines they draw are rarely smooth and continuous. The paused areas contain a large amount of dense points. Although the Kalman filter can help remove outliers and smooth the tracked sketching 3D points, the noise can still propagate into the query sketch. This partially explains why the performance of sketch-sketch retrieval is even worse.

Among these three experiments, the most challenging is sketch-model retrieval, which attempts to retrieve similar 3D models from the SHREC13STB benchmark given a 3D sketch in the Kinect300 dataset. The query dataset and target dataset are different. The query sketch and the target 3D models are not even one to one correspondent since we don't ask users to sketch existing 3D models in the 3D benchmark dataset. Additionally, our collected Kinect300 dataset is a diverse dataset, which contains many challenging categories such as dog, human, face, house.... For these categories, drawing a simple and compact sketch can be a challenging task. Therefore, it is not surprising to see that the performance of the third experiment is far below the performance of other two. It also raises a challenging but interesting problem specific to this type of 3D model retrieval: how to effectively compare an inaccurate hand-drawn 3D sketch with an accurate 3D model? However, many simple categories, including wineglass, sword, airplanes, and balloons, display good retrieval results that suggest further research could significantly improve the retrieval performance. Figure 6 shows two examples using sketches with different complexity.

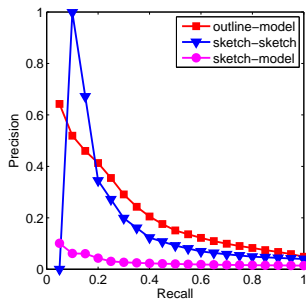


Figure 5: Precision-Recall performance.

There is still much room left for further improvement in this challenging task of 3D sketch-based 3D model retrieval. More descriptive shape descriptors are deserved for our further exploration for this specific type of retrieval. In addition, the 3D sketches we currently collect are quite abstract and noisy. Better denoising and smoothing algorithms will

Table 1: Other performance metrics.

Benchmark	NN	FT	ST	E	DCG
outline-model	0.391	0.156	0.238	0.121	0.486
sketch-sketch	0.167	0.087	0.139	0.092	0.360
sketch-model	0.029	0.021	0.038	0.021	0.254

be investigated to further remove noisy points (e.g. the dot in the alarm clock example in Fig. 4) created during 3D sketching. To further improve the performance of our retrieval system, supervised learning could be applied to help better understand 3D sketch structure before performing matching and retrieval.

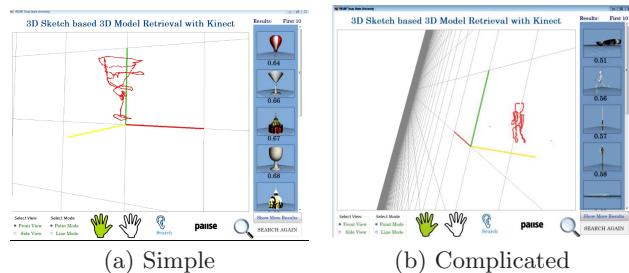


Figure 6: Two example retrieval results using sketches with different complexity.

## 5. CONCLUSIONS AND FUTURE WORK

3D sketching in 3D space and 3D sketch-based 3D model retrieval are brand new research topics. Currently, software does not exist that allows users to include depth information in their sketches. Very little preliminary work exists in this field, allowing for exciting and interesting research results. In this paper, a novel 3D sketching virtual drawing “board” is proposed and a 3D sketch-based 3D model retrieval system is developed. Future goals include developing descriptive and discriminant 3D shape descriptors to represent 3D sketches and 3D models, collecting a larger number of 3D sketches from more diverse users, and training our system for better 3D sketch and 3D model matching.

## Acknowledgments

This work is supported by Army Research Office grant W911NF-12-1-0057, NSF CRI-1305302, NSF CNS-1358939 and NSF OCI-1062439.

## 6. REFERENCES

- [1] M. Ankerst, G. Kastenmüller, and et al. 3D shape histograms for similarity search and classification in spatial databases. In *SSD*, pages 207–226, 1999.
- [2] B. Li, Y. Lu, and et al. SHREC’13 track: Large scale sketch-based 3D shape retrieval. In *3DOR*, pages 89–96, 2013.
- [3] B. Li, Y. Lu, and et al. A comparison of 3D shape retrieval methods based on a large-scale benchmark supporting multimodal queries. *CVIU*, 131:1–27, 2015.
- [4] P. Shilane, P. Min, and et al. The Princeton shape benchmark. In *SMI*, pages 167–178, 2004.