3D Sketch-based 3D Model Retrieval with Convolutional Neural Network

Yuxiang Ye[1], Bo Li[2,1], Yijuan Lu[1]

[1]Texas State University, [2]University of Southern Mississippi

Overview

- **Research topic:** 3D Sketch-Based 3D model retrieval (SBR)
  - ❖ Retrieve 3D models from a dataset given a user's hand-drawn 3D sketch
  - ❖ Promising in: game design, 3D animation and human computer interaction, etc
- **3D Sketching:** we collect 3D sketches using Microsoft Kinect
  - ❖ Encodes 3D information, depth and features of more facets of the object
  - ❖ Includes the salient 3D feature lines of its counterpart of 3D models

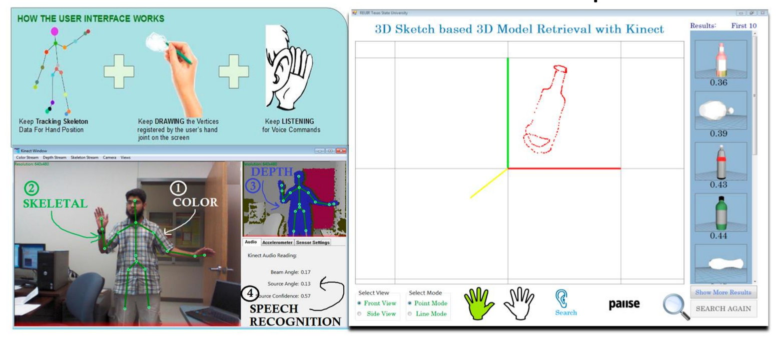

Fig. 1. 3D sketching collection [1]

- **Challenges**
  - ❖ Complexity: 3D sketching is more complex than 2D sketching
    - ➢ Drawing a 3D horse is more difficult than drawing a 2D horse on a paper
  - ❖ Variation: one thousand people may draw the same object in one thousand different ways
    - ➢ Two people are not even able to draw exact the same 3D dogs
  - ❖ Uncertainty: 3D sketches only record 3D coordinates of all the individual points captured from human's hand movement during sketching
    - ➢ A lot of noisy and inaccurate points are captured due to hand shaking, object occlusion, and camera delay
- **Research results:** The experimental results reveal our approach outperforms other competing state-of-the-arts and demonstrate promising potentials of our approach on 3D sketch based application.
- **Contributions**
  - ❖ A novel 3D sketch-based 3D model retrieval system is introduced to solve the matching problem between 3D sketches and 3D models
  - ❖ Our CNN-SBR system combines multiple machine learning and 3D vision processing techniques, which will explicitly guide the research in 3D sketch understanding
  - ❖ Comprehensive experiments have been conducted to evaluate the state-of-the-art sketch based retrieval approaches on 3D sketch-based 3D model retrieval
  - ❖ The experimental results not only show our approach outperforms other state-of-the-arts, but also demonstrates promising application potential of our approach on 3D sketch understanding, on-line 3D model shopping, and large scale 3D model search, etc

Algorithm

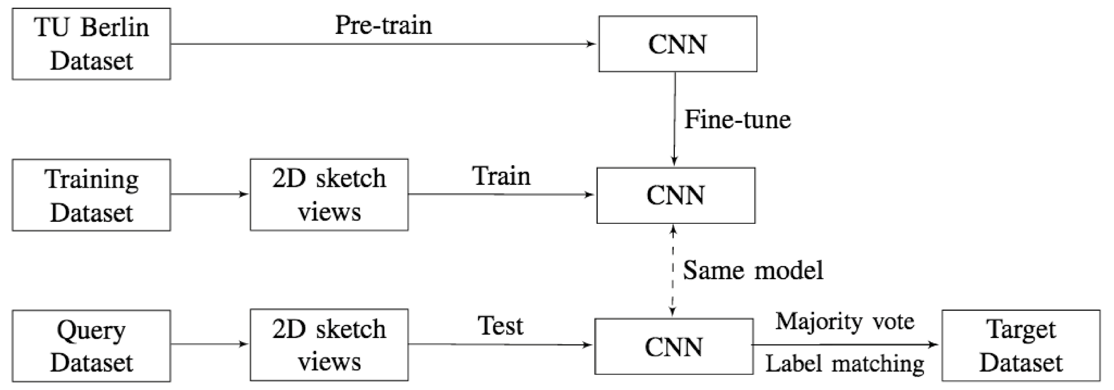

Fig. 2. Illustration of CNN-SBR architecture

- **We employ advanced deep learning method and propose a novel 3D sketch based 3D model retrieval system**
  - ❖ Pre-train our deep CNN model on TU Berlin dataset and obtain well-learned weights for our CNN model
  - ❖ Convert all the 3D sketches to multiple 2D sketch views for both training and query datasets, and perform data augmentation for these 2D sketch views
  - ❖ Fine-tune the CNN model using previously well-learned weights: instead of employing mature learned pre-trained model (training epoch over 500), we choose a semi-mature model at epoch 50
  - ❖ Use majority vote and simple label matching to generate the output result

Algorithm (Cont.)

Input: Original 3D sketch dataset S
Output: Enlarged 3D sketch dataset T with random shifts, rotations, and flips
initialization;
$w = width_{original} - width_{target}$;
foreach $I \in S$ **do**
  for $i \leftarrow 1$ **to** 500 **do**
    $C \leftarrow copy(I)$;
    $x_{shift} \leftarrow random(0, w)$;
    $y_{shift} \leftarrow random(0, w)$;
    $C \leftarrow shift(C, x_{shift}, y_{shift})$;
    $roll \leftarrow random(0, 1)$;
    if $roll < 0.5$ **then**
      $rd \leftarrow random(-5, 5)$;
      $C \leftarrow rotate(C, rd)$;
    end
    $roll \leftarrow random(0, 1)$;
    if $roll < 0.5$ **then**
      $C \leftarrow flip(C)$
    end
    $append(T, C)$;
  end
end

Algorithm 1: Data augmentation algorithms

- **Data processing**
  - ❖ To adapt the framework for 2D sketch-based CNN model, we need to convert the 3D sketches to 2D sketch views
  - ❖ Project all the coordinates in each 3D sketch to its six square faces if we regard a 3D sketch as regular hexahedron
  - ❖ Map the 3D coordinates to 2D depth image
  - ❖ Replicate both TU Berlin dataset and 2D sketch views by 500 times using random rotation, shift and flip (see **Algorithm 1**)
- **Majority vote and label matching**
  - ❖ Rescale the similarities between a 3D sketch and target 3D model categories to range [0, 1]. A higher value means bigger similarity
  - ❖ Count the number of top-1 labels among six similarity vectors for each target 3D model category
  - ❖ Compute the average similarity between this sketch and target 3D model categories based on six similarity vectors
  - ❖ Rank all the target 3D model categories using the summation of the top-1 label count and the average similarity
  - ❖ Rank all the related models accordingly

Experiments

- **To comprehensively evaluate the performance of our CNN-SBR system, we participated in 2016 Shape Retrieval Contest (SHREC'16) track which targets on 3D sketch-based 3D model retrieval**
  - ❖ **SHREC'16 3D Sketch Track Benchmark consists of two parts:**
    - ➢ **Kinect300 dataset**: a 3D sketch dataset which consists of 300 3D sketches (30 categories, 10 sketches per category) and 21 categories have the corresponding models in the target SHREC13STB dataset. Fig. 3. shows some example 3D sketches
    - ➢ **SHREC13STB dataset**: a 3D model dataset which consists of 1258 models unevenly distributed in 90 categories
      - ○ Fig. 4. shows some example 3D models

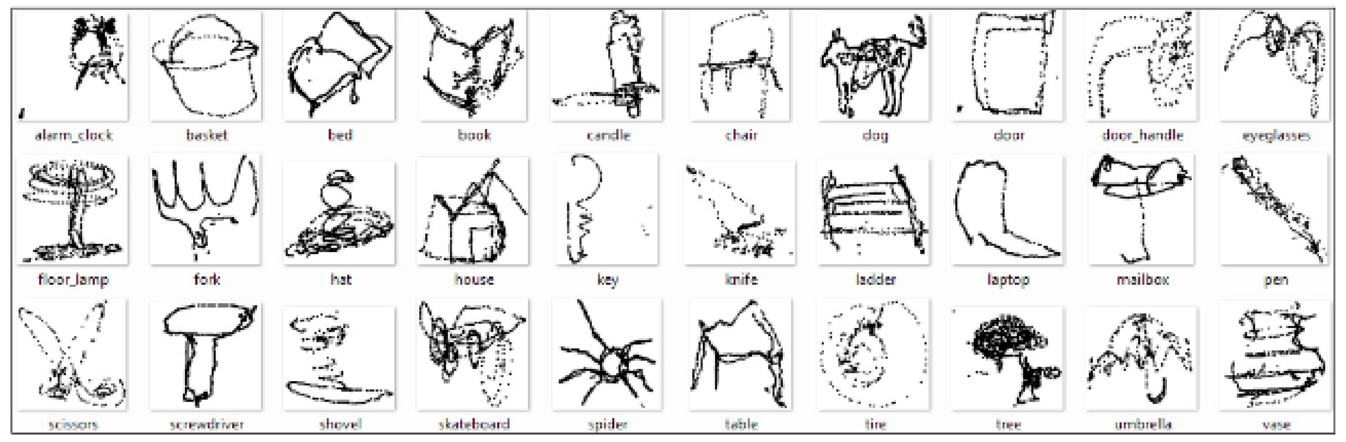

Fig. 3. Example 3D sketches of Kinect300 dataset [1,2]

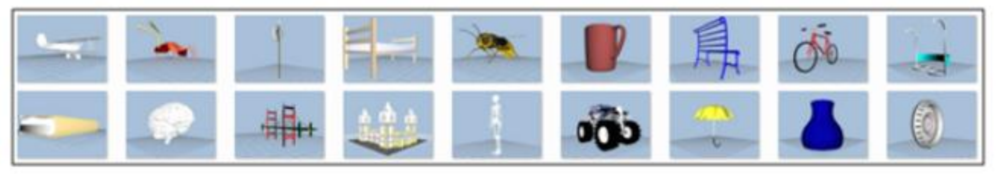

Fig. 4. Example 3D models of SHREC13STB benchmark [1, 2]

- ❖ **Evaluation Results:**
  - ➢ We perform the the comparison based on six widely-used evaluation metrics: Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-Measure (E), Discounted Cumulative Gain (DCG) and Average Precision (AP) on the test dataset of SHREC'16 3D Sketch Track Benchmark for learning based participating algorithms and on the complete dataset for non-learning based algorithms. The results are compared in Table 1
  - ➢ We also perform the Precision-Recall comparison in Fig. 5

Experiments (Cont.)

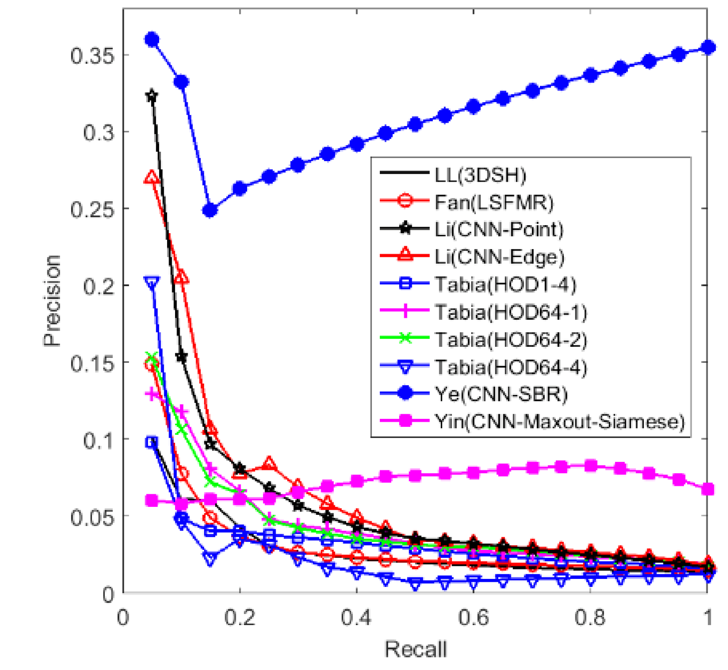

Fig.5. Precision-Recall comparison on the complete dataset of SHREC13STB for non-learning based algorithms and on the testing dataset for learning based algorithms

Participant	Method	NN	FT	ST	E	DCG	AP
Complete benchmark (Non-learning based methods)							
LL	3DSH	0.029	0.021	0.038	0.021	0.254	0.029
Fan	LSFMR	0.033	0.020	0.033	0.018	0.248	0.032
Li	CNN-Point	0.124	0.044	0.075	0.046	0.294	0.060
	CNN-Edge	0.114	0.056	0.084	0.051	0.302	0.063
Tabia	HOD1-4	0.029	0.015	0.035	0.026	0.259	0.032
	HOD64-1	0.052	0.031	0.053	0.034	0.274	0.044
	HOD64-2	0.067	0.031	0.057	0.032	0.272	0.044
	HOD64-4	0.124	0.019	0.022	0.013	0.230	0.026
Testing dataset (Learning-based methods)							
Ye	CNN-SBR	0.222	0.251	0.320	0.186	0.471	0.314
Yin	CNN-Maxout-Siamese	0.000	0.031	0.108	0.048	0.293	0.072

Table 1. Performance metrics comparison on the SHREC'16 3D Sketch Track Benchmark

References

[1] B. Li, Y. Lu, and et al., "SHREC'16 track: 3D sketch-based 3D shape retrieval," in Eurographics Workshop on 3D Object Retrieval (3DOR 2016), Lisbon, Portugal, 2016.

[2] B. Li, Y. Lu, A. Ghumman, B. Strylowski, M. Gutierrez, S. Sadiq, S. Forster, N. Feola, and T. bugerin, "3D sketch-based 3D model retrieval," in Proceedings of the 5th ACM on International Conference on Multimedia Retrieval , New York, NY, USA, 2015, pp. 555–558.

[3] M. Eitz, J. Hays, and M. Alexa, "How do humans sketch objects?" ACM Trans. Graph. , vol. 31, no. 4, pp. 44:1–44:10, Jul. 2012.

[4] Q. Yu, Y. Yang, Y.-Z. Song, T. Xiang, and T. Hospedales, "Sketch-a-net that beats humans," in arXiv:1501.07873, 2015.

[5] B. Li and Y. Lu. KinectSBR: A Kinect-Assisted 3D Sketch-Based 3D Model Retrieval System. In *ACM ICMR'15*, pages 655–656, 2015.

Acknowledgement

This work is supported by Army Research Office grant W911NF-12-1-0057, NSF CNS-1305302 and NSF CNS- 1358939 to Dr. Lu.