

SHREC'19 Track: Extended 2D Scene Image-Based 3D Scene Retrieval

Hameed Abdul-Rashid^{†1}, Juefei Yuan*^{†1}, Bo Li*^{†1}, Yijuan Lu^{†2}, Tobias Schreck^{†3}, Perez Rey[‡], Mike Holenderski^{‡4}, Dmitri Jarnikov^{‡4}, Vlado Menkovski^{‡4}, Masaki Anono^{‡5}, Haisheng Li^{‡6}, Yujuan Wu^{‡6}, Yanping Zheng^{‡6}, Yujia Du^{‡6}, Haoze Yu^{‡6}, Guang Zeng^{‡6}, Yang Wei^{‡6}, Yujia Zhai^{‡6}, Xiaoqun Wu^{‡6}, Zhaoqun Li^{‡7}, Cheng Xu^{‡7}, Biao Leng^{‡7}, [INSERT TRAN'S GROUP]

¹ University of Southern Mississippi, USA

² Texas State University, USA

³ Graz University of Technology, Austria

⁴ Eindhoven University of Technology, Netherlands

⁵ Toyohashi University of Technology, Japan

⁶ Beijing Technology and Business University, China

⁷ Beihang University, China

Abstract

[Insert abstract here]

1. Introduction

2D scene image-based 3D scene model retrieval is to retrieve 3D scene models given an input 2D scene image. It has vast related applications, including highly capable autonomous vehicles like the Renault SYMBIOZ [Ren] [Tip], multi-view 3D scene reconstruction, VR/AR scene content generation, and consumer electronics apps, among others. However, this task is far from trivial and lacks substantial research due to the challenges involved as well as a lack of related retrieval benchmarks. Consequently, existing 3D model retrieval algorithms have been limited to focus on single object retrieval. Seeing the benefits of advances in retrieving 3D scene models based on a scene image query makes this research direction useful, promising, and interesting as well.

To promote this interesting yet challenging research, we organized a 2018 Eurographics Shape Retrieval Contest (SHREC) track [ARYLL18] titled “2D Scene Image-Based 3D Scene Retrieval”, by building the first 2D scene image-based 3D scene retrieval benchmark **SceneIBR2018**, comprising 10,000 2D scene images and 1,000 3D scene models. All the images and models are equally classified into 10 indoor as well as outdoor classes.

However, as can be seen, **SceneIBR2018** contains only 10 distinct scene classes, and this is one of the reasons that all the three deep learning-based participating methods have achieved excellent

performance on it. Considering this, after the track we have tripled the size of **SceneIBR2018**, resulting in an extended benchmark **SceneIBR2019**, which has 30,000 2D scene images and 3,000 3D scene models. Similarly, all the 2D images and 3D scene models are equally classified into 30 classes. We have kept the same set of 2D scene images and 3D scene models belonging to the initial 10 classes of **SceneIBR2018**.

Hence, this track seeks participants who will provide new contributions to further advance 2D scene images-based 3D scene retrieval for evaluation and comparison, especially in terms of scalability to a larger number of scene categories, based on the new benchmark **SceneIBR2019**. Similarly, we will also provide corresponding evaluation code for computing a set of performance metrics similar to those used in the Query-by-Model retrieval technique.

2. Benchmark

2.1. Overview

Building process. The first thing for the benchmark design is category selection, for which we have referred to several of the most popular 2D/3D scene datasets, such as Places [ZLK*17] and SUN [Xea10]. The criteria for the category selection is popularity. Finally, we selected the most popular 30 scene classes (including the initial 10 classes in **SceneIBR2018**) from the 88 available category labels in the Places88 dataset [5], via a voting mechanism among three people (two graduate students as voters and a faculty member as the moderator) based on their judgments. We want to

[†] Track organizers. *Corresponding author. For any question related to the track, please contact Bo Li. E-mail: bo.li@usm.edu.

[‡] Track participants.

mention that the 88 common scenes are already shared by ImageNet [Dea09], SUN [Xea10], and Places [ZLK*17]. Then, to collect data (images and models) for the additional 20 classes, we gathered from Flickr and Google Image for images, and downloaded SketchUp 3D scene models (originally, in “.SKP” format, but we provide “.OBJ” format as well after transformation) from 3D Warehouse [8].

Benchmark details. Our extended 2D scene image-based 3D scene retrieval benchmark **SceneIBR2019** expands the initial 10 classes of **SceneIBR2018** with 20 new classes totaling a more comprehensive dataset of 30 classes. **SceneIBR2019** contains a complete dataset of 30,000 2D scene images (1,000 per class) and 3,000 3D scene models (100 per class). Examples for each class are demonstrated in both **Fig. 1** and **Fig. 2**.

In the same manner as the **SceneIBR2018** track, we randomly pull 700 images and 70 models out from each class for training and the remaining 300 images and 30 models are used for testing, as shown in Table 1. If a method involves a learning-based approach, results for both the training and testing datasets need to be submitted. Otherwise, retrieval results based on the complete datasets are needed.

Table 1: Training and testing datasets information of our **SceneIBR2019** benchmark.

| Datasets | Images | Models |
|----------------------|--------|--------|
| Training (per class) | 700 | 70 |
| Testing (per class) | 300 | 30 |
| Total (per class) | 1000 | 100 |
| Total (all 30 class) | 30,000 | 3,000 |

2.2. 2D Scene Image Dataset

. The 2D scene image query set is composed of 30,000 scene images (30 classes, each with 1,000 images) that are all from the Flickr and Google Image websites. One example per class is demonstrated in **Fig. 1**.

2.3. 3D Scene Dataset

. The 3D scene dataset is built on the selected 3,000 3D scene models downloaded from 3D Warehouse. Each class has 100 3D scene models. One example per class is shown in **Fig. 2**.

2.4. Evaluation Method

To have a comprehensive evaluation of the retrieval algorithm, we employ seven commonly adopted performance metrics in 3D model retrieval community: Precision-Recall (PR) diagram, Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-Measures (E), Discounted Cumulated Gain (DCG) and Average Precision (AP) [LLL*15]. We have developed the related code to compute these metrics and will provide the code to participants.

3. Participants

[Insert Participants Here]



Figure 1: Example 2D scene images (one example per class) in our **SceneIBR2019** benchmark.

4. Methods

[Insert Methods Here]

5. Results

[Insert Results Here]

6. Conclusions and Future Work

Insert Conclusion Here

References

- [ARYLL18] ABDUL-RASHID H., YUAN J., LI B., LU Y.: SHREC’18 2D Scene Image-Based 3D Scene Retrieval Track Website. <http://orca.st.usm.edu/~bli/SceneIBR2018/>, 2018. 1
- [Dea09] DENG J., ET AL: ImageNet: A large-scale hierarchical image database. In *CVPR* (2009), pp. 248–255. 2
- [LLL*15] LI B., LU Y., LI C., GODIL A., SCHRECK T., AONO M., BURTSCHER M., CHEN Q., CHOWDHURY N. K., FANG B., ET AL.: A comparison of 3d shape retrieval methods based on a large-scale benchmark supporting multimodal queries. *Computer Vision and Image Understanding* 131 (2015), 1–27. 2
- [Ren] RENAULT: Renault SYMBOIZ Concept. <http://www.renault.co.uk/vehicles/concept-cars/symbioz-concept.html>. 1
- [Tip] TIPS L. T.: Driving a multi-million dollar autonomous car. <http://www.youtube.com/watch?v=vLIJfVlu2hM&feature=youtu.be>. 1



Figure 2: Example 3D scene models (one example per class, shown in one view) in our **SceneIBR2019** benchmark.

[Xea10] XIAO J., ET AL: SUN database: Large-scale scene recognition from abbey to zoo. In *CVPR* (2010), IEEE Computer Society, pp. 3485–3492. [1](#), [2](#)

[ZLK*17] ZHOU B., LAPEDRIZA A., KHOSLA A., OLIVA A., TORRALBA A.: Places: a 10 million image database for scene recognition. *IEEE Trans. on PAMI* (2017). [1](#), [2](#)