## SHREC 2019 - Extended 2D Scene Image-Based 3D Scene Retrieval Track Proposal

**Title:** SHREC 2019 Extended 2D Scene Image-Based 3D Scene Retrieval

**Organizers:**

Juefei Yuan,University of Southern Mississippi, USA

Bo Li, University of Southern Mississippi, USA

Yijuan Lu, Texas State University, USA

Tobias Schreck, Graz University of Technology, Austria

**Objective:** In the months following our SHREC 2018 - 2D Scene Image-Based 3D Scene Retrieval (SceneIBR2018) track [1], we have extended the number of the scene categories from the initial 10 classes to 30 classes [2], resulting a new benchmark **Scene IBR**, which has 30,000 scene images and 3,000 3D scene models. For that reason, we seek to further evaluate the performance of existing and new 2D scene image-based 3D scene model retrieval algorithms using this extended and more comprehensive new benchmark.

**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 [3] [4], multi-view 3D scene reconstruction, VR/AR scene content generation, and consumer electronics apps. However, this task is far from trivial and lacks substantial research due to the challenges involved and 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) tracks [21, 2, 22, 3] titled “2D Scene Image-Based 3D Scene Retrieval”, by building the first 2D scene image-based 3D scene retrieval benchmark **SceneIBR2018**, which has10,000 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. Considering this, after the track we have tripled the size of **SceneIBR2018**

to make each has 30 classes, and built an extended version for it, resulting an extended benchmark **Scene IBR**, which has 30,000 scene images, and 3,000 3D scene models. Similarly, all the 2D sketches, 2D images and 3D scenes are equally classified into 30 classes. We have kept the same scene images and models in the initial 10 classes of **SceneIBR2018.**

Hence, this track seeks participants who will provide new contributions to advance 2D scene images-based 3D scene retrieval for evaluation and comparison, especially in terms of scalability to the number of scene categories, on the new benchmark **Scene\_IBR**. 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.

**Benchmark:**

**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 [23] and SUN [15]. 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 [23], 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 mention that the 88 common scenes are already shared by ImageNet [5], SUN [15], and Places [23]. Then, to collect data (sketches, images, and models) for the additional 20 classes, we gathered from Flicker and Google Image for sketches and images, and downloaded SketchUp 3D scene models (originally in .SKP format, but we provide .OBJ format as well after transformation) from 3D Warehouse [1].

**Benchmark verview.**

Our extended 2D scene image-based 3D scene retrieval benchmark **Scene\_IBR** expands the initial 10 classes of SceneIBR2018 with 20 new classes totaling a comprehensive dataset of 30 classes. The Scene\_IBR contains a complete dataset of 30,000 2D scene images (1000 per class) and 3,000 3D scene models (100 per class). Example for each class are demonstrated in both **Fig. 1** and **Fig. 2**.

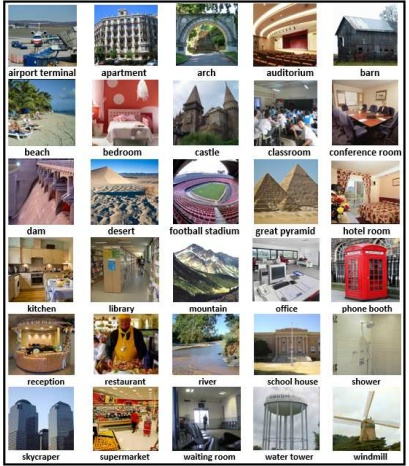
In the same manner as the SceneIBR2018 track, we randomly pull 700 images and 70 models from each class for training and the remaining 300 images and 30 models are used for testing, as shown in **Table 1.**

If a method involves a learning-based approac, results for both the training and testing datasets need to be submitted. Otherwise, the retrieval results based on the test and complete datasets are needed. To provide future users a complete reference of our benchmark, we will evaluate the participating algorithms on both the testing dataset and the complete **Scene\_IBR** benchmark.

**Table 1.** Training and testing datasets (per class) of our **Scene\_IBR** benchmark.

|  |  |  |
| --- | --- | --- |
| **Scene\_IBR** Benchmark | Image | Model |
| Train | 700 | 70 |
| Test | 300 | 30 |
| Total (per class) | 1000 | 100 |
| Total (all 30 classes) | 20,000 | 3000 |

**2D Scene Image Dataset**



**Fig. 1** Example 2D scene images (one example per class, shown in one view) in our **Scene\_IBR** benchmark.

**3D Scene Dataset**



**Fig. 2** Example 3D scene models (one example per class) in our **Scene\_IBR** benchmark.

**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) [5]. We have developed the related code to compute these metrics and will provide them to participants.

**The Procedural Aspects:**The complete dataset will be made available on the 25th of January and the results will be due in six weeks after that. Every participant is expected to perform the queries and send us their retrieval results. We will then do the performance assessment. Participants and organizers will collaborate to write a joint SHREC track competition report to detail the results and evaluations. Results of the track will be presented during the Eurographics 3DOR Workshop 2019 in Genova, Italy.

**Preliminary Timeline**

|  |  |
| --- | --- |
| January 23 | - Call for participation. |
| **January 25** | -Distribution of the database. Participants can start the retrieval or train their algorithms. |
| **February 8** | - Please register before this date. |
| **March 8** | - Submission of the results on the test (for learning-based methods) or the complete (for non-learning based approaches) datasets and one-page description of their method(s). |
| March 11 | - Distribution of relevance judgments and evaluation scores. |
| March 13 | - Track is finished and results are ready for inclusion in a track report. |
| March 15 | - Submit the track report for review. |
| March 25 | - Reviews done, feedback and notifications given. |
| April 5 | - Camera-ready track paper submitted for inclusion in the proceedings. |
| May 5-6 | - Eurographics Workshop on 3D Object Retrieval 2019, featuring SHREC’2019. |



**References:**

[1] Hameed Abdul-Rashid and et al., SHREC 2018: 2D Scene Image-Based 3D Scene Retrieval, Eurographics Workshop on 3D Object Retrieval 2018 (3DOR 2018), 2018

[2] Juefei Yuan\*, Hameed Abdul-Rashid\*, **Bo Li**, Yijuan Lu. *Sketch/Image-Based 3D Scene Retrieval: Benchmark, Algorithm, Evaluation*. The IEEE 2nd International Conference on Multimedia Information Processing and Retrieval. March 28-30, San Jose, CA, USA (Invited Paper), January 2019, Accepted.[3] Renault. Renault SYMBOIZ Concept. http://www.renault.co.uk/vehicles/concept-cars/

symbioz-concept.html.

[4] L. T. Tips. Driving a multi-million dollar autonomous car. http://www.youtube.com/watch?

v=vlIJfV1u2hM\&feature=youtu.be.

[2] Jia Deng and et al. ImageNet: A large-scale hierarchical image database. CVPR 2009: 248-255.

[3] B. Zhou and et al. Places: A 10 million image database for scene recognition. IEEE Trans. Pattern Anal. Mach. Intell., 40(6):1452–1464, 2018.

## [4] 3D Warehouse. <https://3dwarehouse.sketchup.com/?hl=en>.

[5] B. Li and et al. A comparison of 3D shape retrieval methods based on a large-scale benchmark supporting multimodal queries. *Computer Vision and Image Understanding*, 131:127, 2015.