

**SHREC'18 Track: 2D Scene Image-Based 3D Scene Retrieval Track Proposal**

**Title:** SHREC'18 Track: 2D Scene Image-Based 3D Scene Retrieval

**Organizers:** Hameed Abdul-Rashid, University of Southern Mississippi, USA  
Juefei Yuan, University of Southern Mississippi, USA  
Bo Li, University of Southern Mississippi, USA  
Yijuan Lu, Texas State University, USA

**Objective:** The objective of this track is to evaluate the performance of different 2D scene image-based 3D scene retrieval algorithms using a 2D scene image query dataset and a collected 3D Warehouse model dataset.

**Introduction:**

Provided with a 2D scene image, a 2D scene image-based 3D scene retrieval algorithm is to search for relevant 3D scenes (.OBJ file) from a dataset. It is an intuitive and convenient framework which allows users to learn, search, and utilize the retrieved results for related applications. For example, automatic 3D content generation based on one or a sequence of captured images for AR/VR applications, or 3D movie, game and animation production, robotic vision (i.e. path finding), and consumer electronics apps, which facilitate users to efficiently generate a 3D scene after taking an image of a real scene. It is also very promising and has great potentials in related applications such as 3D geometry video retrieval, and highly capable autonomous vehicles such as the Renault SYMBIOZ [1][2].

However, there is little research in 2D scene image-based 3D scene shape retrieval [3][4] due to two reasons: (1) the problem itself is challenging to cope with; (2) lack of related retrieval benchmarks. Seeing the benefit of advances in retrieving 3D scene models using 2D scene image queries makes the research direction meaningful, interesting and promising.

Deng et al. [5] collected the ImageNet database initially comprising of 5,247 synsets and 3.2 million images back in 2009. Nearly ten years after its inception, there over 21,000 synsets indexed and nearly 14.2 million images. For this track, we built a smaller and more manageable dataset comprising of 10,000 scene images across 10 classes, each with 1,000 images. It avoids the bias issue since we collected the same number of images for every class, while the images variation within one class is also adequate enough.
To organize a SHREC’18 2D scene sketch-based 3D model retrieval track [6], we have collected 1,000 3D Warehouse [7] scene mesh models (transformed to .OBJ format) to correspond to the 250 scene sketches equally divided into 10 classes in the Scene250 sketch dataset [8]. For each class, we collected the same number (100) of 3D scene models as well. We reuse this 3D scene target dataset for this track and only need to collect 2D scene images to form a query data, which are not difficult to find.

This track is organized to promote this challenging research direction by soliciting state-of-the-art 2D scene image-based 3D scene retrieval methods and foreseen the future directions on this research topic. Evaluation code for computing a set of performance metrics similar to those used in the Query-by-Model retrieval technique will also be provided.

**Benchmark Overview:**

Our 2D scene image-based 3D scene shape retrieval benchmark SceneIBR utilizes 10,000 2D scene images selected from ImageNet [5] as its 2D scene image dataset and 1,000 3D Warehouse scene models (.OBJ format) as its 3D scene dataset, and both have ten classes. Each of the ten classes contains the same number of 2D scene images (1,000) and 3D scene models (100).

To facilitate learning-based retrieval, we randomly select 700 images and 70 models from each class for training and use the remaining 300 images and 300 models for testing. We conclude this in Table 1. Participants need to submit results on the training and testing datasets, respectively, if they use a learning-based approach. Otherwise, the retrieval results based on the test (300 images, 30 models) and complete (10,000 images, 1000 models) datasets are needed. To provide a complete reference for future users of our SceneIBR benchmark, we will evaluate the participating algorithms on both the testing dataset (300 images and 30 models per query) and the complete SceneIBR benchmark (10,000 images and 100 models per class).

**Table 1.** Training and testing datasets (per class) of our SceneIBR benchmark.

<table>
<thead>
<tr>
<th>SceneIBR Benchmark</th>
<th>Image</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>700</td>
<td>70</td>
</tr>
<tr>
<td>Test</td>
<td>300</td>
<td>30</td>
</tr>
<tr>
<td>Total (per class)</td>
<td>1,000</td>
<td>100</td>
</tr>
</tbody>
</table>

**2D Scene Image Dataset:**

The 2D scene image query set is composed of 10,000 scene images (10 classes, each with 1,000 images) that are all from ImageNet [5], while all the classes have relevant models in the target 3D
scene dataset which are downloaded from the 3D Warehouse [7]. One example per class is demonstrated in Fig. 1.

![Fig. 1 Example 2D scene images (one example per class) [4] in our SceneIBR benchmark.](image1)

**3D Scene Dataset:**

The 3D scene dataset is built on the selected 1,000 3D scene models downloaded from Google 3D Warehouse. Each class has 100 3D scene models. One example per class is shown in Fig. 2.

![Fig. 2 Example 3D scenes models (one example per class) in our SceneIBR benchmark.](image2)

**Evaluation Method:**

To have a comprehensive evaluation of the retrieval algorithm, we employ seven commonly adopted performance metrics in 3D model retrieval technique. They are Precision-Recall (PR) diagram, Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-Measures (E), Discounted Cumulated Gain (DCG) and Average Precision (AP). We also have developed the code to compute them.
The Procedural Aspects:
The complete dataset will be made available on the 1st of February and the results will be due in three 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 2018 in Delft, Netherlands.

Preliminary Timeline
January 22 - Call for participation.
January 25 - A few sample 2D scene images and 3D scene models will be available online.
January 29 - Please register before this date.
February 1 - Distribution of the database. Participants can start the retrieval or train their algorithms.
February 22 - Submission of the results on the test (for learning-based methods) or test & complete (for non-learning based approaches) datasets and one-page description of their method(s).
February 24 - Distribution of relevance judgments and evaluation scores.
February 26 - Track is finished and results are ready for inclusion in a track report.
February 28 - Submit the track report for review.
March 3 - Reviews done, feedback and notifications given.
March 10 - Camera-ready track paper submitted for inclusion in the proceedings.

References:
[6] SHREC’18 2D Scene-Based 3D Scene Retrieval website. 
www.3DSearchLab.net/SceneSBR.html.

■ End