SimCol Evaluation Details

Anita Rau

August 24, 2022

Abstract

This document aims to give a more detailed description of the submission format for the SimCol-to-3D 2022 challenge. The document describes the expected submission *zip*-folder and the evaluation metrics used for the leaderboard. Further information can be found on the challenge website https://www.synapse.org/#!Synapse:syn28548633/wiki/617126 and the challenge GitHub project https://github.com/anitarau/simcol.

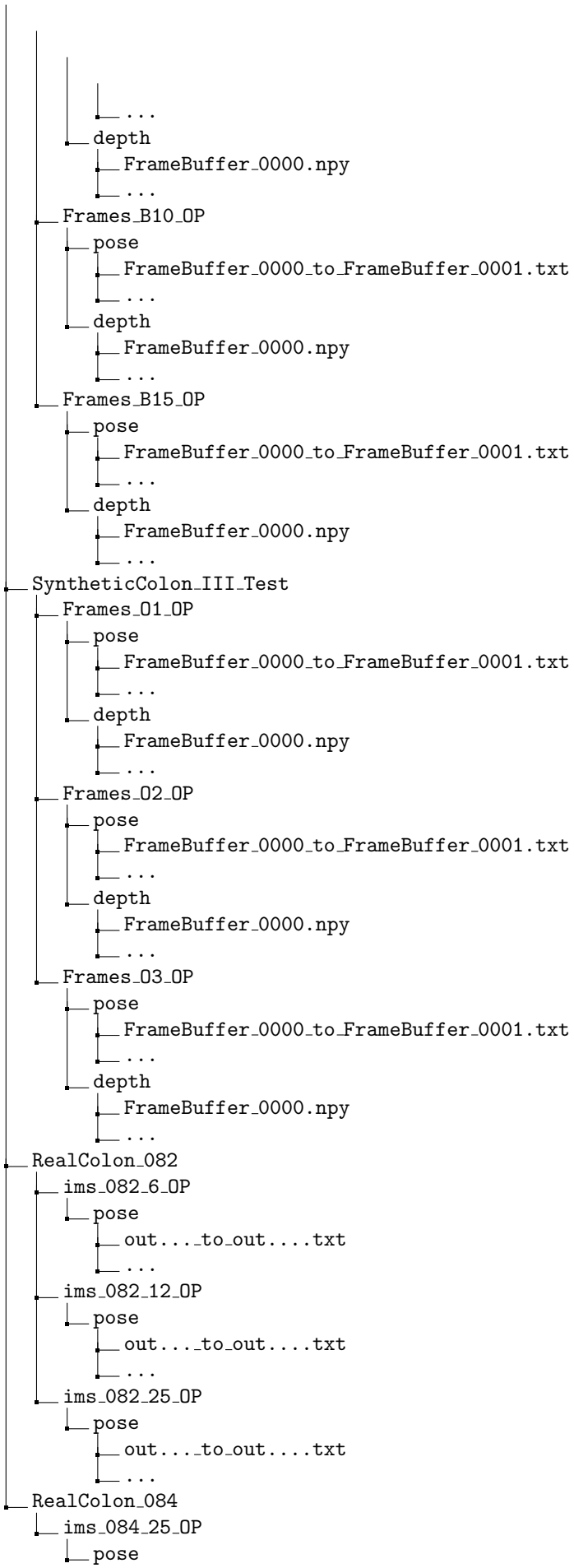
1 Submission Format

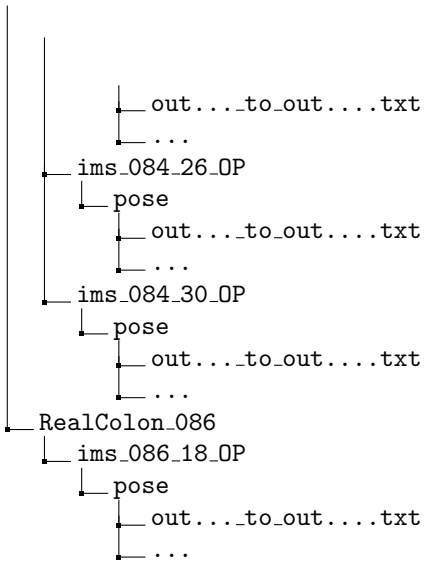
The challenge consists of three sub-tasks. Participants can choose to compete in one, two, or all three tasks.

1.1 Submission .zip File

Independently of the number of tasks that a group submits to, only one .zip file should be submitted. If all three tasks are tackled, the structure of this zip files should be as follows:

```
SimCol_submission
├── SyntheticColon_I_Test
│   ├── Frames_S5_OP
│   │   ├── pose
│   │   │   ├── FrameBuffer_0000_to_FrameBuffer_0001.txt
│   │   │   └── ...
│   │   └── depth
│   │       ├── FrameBuffer_0000.npy
│   │       └── ...
│   ├── Frames_S10_OP
│   │   ├── pose
│   │   │   ├── FrameBuffer_0000_to_FrameBuffer_0001.txt
│   │   │   └── ...
│   │   └── depth
│   │       ├── FrameBuffer_0000.npy
│   │       └── ...
│   └── Frames_S15_OP
│       ├── pose
│       │   ├── FrameBuffer_0000_to_FrameBuffer_0001.txt
│       │   └── ...
│       └── depth
│           ├── FrameBuffer_0000.npy
│           └── ...
└── SyntheticColon_II_Test
    └── Frames_B5_OP
        └── pose
            └── FrameBuffer_0000_to_FrameBuffer_0001.txt
```


...
depth
FrameBuffer_0000.npy
...
Frames_B10_OP
pose
FrameBuffer_0000_to_FrameBuffer_0001.txt
...
depth
FrameBuffer_0000.npy
...
Frames_B15_OP
pose
FrameBuffer_0000_to_FrameBuffer_0001.txt
...
depth
FrameBuffer_0000.npy
...
SyntheticColon_III_Test
Frames_01_OP
pose
FrameBuffer_0000_to_FrameBuffer_0001.txt
...
depth
FrameBuffer_0000.npy
...
Frames_02_OP
pose
FrameBuffer_0000_to_FrameBuffer_0001.txt
...
depth
FrameBuffer_0000.npy
...
Frames_03_OP
pose
FrameBuffer_0000_to_FrameBuffer_0001.txt
...
depth
FrameBuffer_0000.npy
...
RealColon_082
ims_082_6_OP
pose
out..._to_out....txt
...
ims_082_12_OP
pose
out..._to_out....txt
...
ims_082_25_OP
pose
out..._to_out....txt
...
RealColon_084
ims_084_25_OP
pose



If only subtasks 1 and 2 are tackled, then no real data should be submitted and the respective folders for real data should not exist in the submission .zip file. If a group only competes in subtask 1, then no pose folders should exist in the submission .zip file. Accordingly, no depth folders should exist, if only tasks 2 or 3 are tackled.

`SimCol_submission` should be zipped and submitted as a single file through Synapse.

1.2 File formats

Details on the submission format should be gathered from the project GitHub. Under `docker_templates` we provide scripts that can be run to generate sample outputs. Please refer to the comments in the `add_your_code_here.py` files!

1.2.1 Subtask 1

To limit file sizes it is essential that depth maps are saved in the float16 format. The depths should be in range [0,1] and will converted to centimeters by the challenge organizers during evaluation. It is also essential that depth maps during training are read as described here: https://github.com/anitarau/simcol/blob/fc38594f80da1e29238d4d8f60968c15c60fbeb6/docker_templates/SimCol3D-task1-depth-docker/src/add_your_code_here.py#L28. Note that using `matplotlib` to read the ground truth depths can lead to slightly different depth values.

1.2.2 Subtask 2

For subtask 2 **relative** camera poses should be submitted. Camera poses should be 4×4 projection matrices as described in Section III of the dataset paper https://arxiv.org/abs/2204.04968. We strongly recommend that participants familiarize themselves with the pose evaluation script on Github https://github.com/anitarau/simcol/blob/main/evaluation/eval_synthetic_pose.py to make sure that the submitted relative poses will be correctly converted to full camera trajectories by the organizers.

1.2.3 Subtask 3

Important: If you wish to participate in subtask 3 but do not have access to the EndoMapper project https://www.synapse.org/#!Synapse:syn26707219/wiki/615178, please contact the SimCol organizers immediately! The real data is managed by a different team and requires compliance with the EndoMapper access protocol.

EndoMapper provides COLMAP labels for (at least) Sequences 001 and 002. The test sequences used for subtask 3 were collected by the EndoMapper team. They have also kindly provided the SimCol organizers with the COLMAP labels that will be used for evaluation. We recommend that

participants familiarize themselves with the COLMAP labels provided for Sequences 001 and 002 and use them for validation of their methods.

2 Evaluation Metrics

Please refer to the evaluation scripts provided on the project GitHub.

2.1 Subtask 1

Let Y be one ground truth depth map, and let Y' be its prediction. Further, let D be the number of pixels d in the depth map and let $\mu()$ denote the median of all d. Then we use the following three evaluation metrics:

$$L_1 = \frac{1}{D}\sum_d ||Y(d) - Y'(d)||_1 \tag{1}$$

$$L_{\text{rel}} = \mu_d(||\frac{Y(d)-Y'(d)}{Y(d)}||_1) \tag{2}$$

$$L_{\text{RMSE}} = \sqrt{\frac{1}{D}\sum_d (Y(d) - Y'(d))^2} \tag{3}$$

At test time the mean over all N depth maps in a sub dataset will be reported. Note that the predicted depth maps will be scaled before evaluation. For each subsequence we compute the scale factor s. Let $\bar{d}$ denote the average depth value of a depth map Y.

$$s = \frac{\sum_n \bar{d} \cdot \bar{d}'}{\sum_n \bar{d}' \cdot \bar{d}'} \tag{4}$$

2.2 Subtasks 2 and 3

There are three metrics to evaluate the accuarcy of the predicted poses: the Absolute Translation Error (ATE), the Relative Translation Error (RTE), and the ROTation error (ROT). Let Ω denote relative poses, and let P denote absolute poses. As before, $'$ denotes the predictions. Then the losses are defined as

$$RTE = \mu_\tau(||trans({\Omega_\tau}^{-1}{\Omega'}_\tau)||), \tag{5}$$

$$ATE = \mu_\tau(||trans(P_\tau) - trans(P'_\tau)||), \quad \text{and} \tag{6}$$

$$ROT = \mu_\tau(\frac{trace(Rot(\Omega_\tau^{-1}\Omega'_\tau)) - 1}{2} \cdot \frac{180}{\pi}), \tag{7}$$

where μ_τ denotes the median over all steps τ, *trans* and *Rot* denote the translation and rotation components of a projection matrix, and $||.||$ denotes the 2-norm.

As for the depth task, we scale the predicted poses before evaluation. We compute the scale s as:

$$s = \frac{\sum_\tau trans(\Omega_\tau)^T \cdot trans(\Omega'_\tau)}{\sum_\tau trans(\Omega'_\tau)^T \cdot trans(\Omega'_\tau)} \tag{8}$$