

Reference camera poses for the query images of the San Francisco Landmarks dataset

This package provides the 6DoF reference camera poses computed in our CVPR17 paper [1]. It also includes the localization benchmarks (figure 3a, b, c in the paper) that evaluate the positional accuracy for 2D image-based and 3D structure-based localization baselines.

Data format description

The reference poses for the query images of the San Francisco Landmarks dataset [2,3] are provided in two formats:

- `reference_poses_442.txt` (plain text)

Each line in this file contains query name, rotation in quaternion, and camera position in the UTM coordinates, e.g.

```
0 <query name> <1x4 quaternion> <1x3 camera position>
```

Note that we call C ($t = -R^*C$ for $P = [R \mid t]$) as the camera position.

- `reference_poses_442.mat` (matlab binary)

This file contains struct array `poses` which has fields `name` and `P`. For example, `poses(1).name` returns the name of query image and `poses(1).P` returns a 3x4 projection matrix $P = [R \mid t]$ of the query in the UTM coordinates.

- `sf0bundler2utm_similarity_transformation.txt` (plain text)

This file contains the similarity transformation from SF-0 model [4,5] to UTM coordinates:

```
cs (1x1 scale)
Rs (3x3 rotation matrix)
ts (3x1 translation vector)
```

3D point `x` in SF-0 can be transformed to UTM coordinates by `Xutm = cs * Rs * X + ts`.

References

[1] T. Sattler, A. Torii, J. Sivic, M. Pollefeys, H. Taira, M. Okutomi, T. Pajdla: Are Large-Scale 3D Models

Really Necessary for Accurate Visual Localization? CVPR 2017.

[2] D. Chen, G. Baatz, K. Koeser, S. Tsai, R. Vedantham, T. Pylvanainen, K. Roimela, X. Chen, J. Bach, M. Pollefeys, B. Girod, and R. Grzeszczuk: City-scale landmark identification on mobile devices. CVPR 2011.

[3] San Francisco Landmark Dataset. <https://purl.stanford.edu/vn158kj2087>

[4] Y. Li, N. Snavely, D. Huttenlocher, P. Fua: Worldwide Pose Estimation using 3D Point Clouds. ECCV 2012.

[5] <http://landmark.cs.cornell.edu/>