Synthetic Data for Non-rigid 3D Reconstruction using a Moving RGB-D Camera

Shafeeq Elanattil[1,2], and Peyman Moghadam[1,2]

[1]Robotics and Autonomous Systems, CSIRO Data61, Brisbane, Australia
[2]Queensland University of Technology, Brisbane, Australia
{shafeeq.elanattil, peyman.moghadam}@data61.csiro.au

Abstract

We introduce a synthetic dataset for evaluating no-rigid 3D reconstruction using a moving RGB-D camera. The dataset consist of two subjects captured with four different camera trajectories. For each case we provide frame-by-frame ground truth geometry of the scene, the camera trajectory and foreground mask. More information about the dataset can be found at: https://research.csiro.au/robotics/databases

1 Introduction

There are only a few public data sets available for evaluating RGB-D based non-rigid 3D reconstruction approaches. The dataset published with VolumeDeform [1] provides canonical and live reconstruction of their approach at every 100^{th} frame. This can only be used for comparing with their approach output. The dataset published with KillingFusion [2] has the canonical ground truth. This can be used for comparing canonical reconstruction. Both of these approaches do not provide any ground truth for comparing live reconstruction.

For evaluating live 3D reconstruction, 3D ground truth has to be in the world coordinate frame. Otherwise camera pose is needed for transforming live reconstruction to the world coordinate frame. All previous approaches [3, 1, 2] do not estimate camera pose, therefore these approaches cannot be compared with live ground truth. In order to address this limitation we releasing a synthetic dataset for evaluating RGB-D based non-rigid 3D reconstruction approaches.

2 The Dataset

Our dataset consists of

1. Complete scene geometry at first frame for evaluating canonical reconstruction.
2. Live scene geometry at each frame in world coordinates.
3. Ground truth of camera trajectory.
4. Ground truth of foreground mask.

Elanattil [4] outlines the detail of the design and production of this synthetic dataset.

2.1 Data Description

Our dataset's folder structure is as follow.

```
Subject1
├── Geometry_Ground_Truth
│   ├── Background_Geometry
│   ├── Canonical_Geometry
│   └── Live_Geometry
└── Input_Data
    ├── trajectory_0
    ├── trajectory_40
    └── ...
Subject2
└── ...
```

Inside the dataset folder two subject's data are given. Each Subject folder consists of Geometry_Ground_Truth and Input_Data sub-folders. Geometry_Ground_Truth folder consists of three sub-folders named Background_Geometry, Canonical_Geometry, and Live_Geometry respectively. Background_Geometry folder contains a mesh file of the background of the scene and Canonical_Geometry folder contains a mesh file of the canonical geometry of the foreground. Live_Geometry folder contains mesh file correspond to the each frame. The file name has the form:

- **mesh_XXXXX.ply**:live mesh file of foreground;

where **XXXXX** is an integer number representing the frame number within the data sequence.

The Input_Data folder contains the input data collected from the non-rigid synthetic scene camera trajectories named of the form trajectory_D. We used saw-tooth camera trajectory for evaluating the tracking capability of our approach. The saw-tooth camera trajectory is created by moving to and fro a distance D millimeters between the adjacent frames. Each trajectory folder have a folder tree structure as below:

```
trajectory_0
└── trajectory.txt
```

```
│
│  │
│  ├── image
│  │   ├── frame_00000.png
│  │   └── ...
│  ├── depth
│  │   ├── depth_00000.png
│  │   └── ...
│  └── foreground_mask
│      ├── labelled_00000.png
│      └── ...
├── trajectory_40
│   └── ...
└── ...
```

Here image, depth and foreground_mask folders contains RGB and depth and foreground mask images respectively. The file names have the follwoing form:

- **frame_XXXXX.png**: the RGB image of the scene;
- **depth_XXXXX.png**: the depth image;
- **labelled_XXXXX.png**: the corresponding fore-ground mask of the scene;

where **XXXXX** is an integer number representing the frame number within the data sequence. The trajectory folder also contains a trajectory.txt file in which the ground truth camera trajectory is given. Each line in trajectory.txt contains single camera pose information. Line number corresponding to the frame index of the sequence.

- The format of each line is: **'qw qx qy qz tx ty tz'**.
- **qw qx qy qz** (4 floats) give the orientation of the optical center of the camera in form of a unit quaternion with respect to the world co-ordinate system.
- **tx ty tz** (3 floats) give the position of the optical center of the color camera with respect to the world co-ordinate system.

The camera intrinsic parameters for both RGB and depth images in our data is

$$\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1050 & 0 & 480 \\ 0 & 1050 & 480 \\ 0 & 0 & 1 \end{bmatrix}.$$

2.2 Obtaining The Data

This dataset can be downloaded from [5]. When making use of this data we ask that [4, 5] are cited.

2.3 Example Data Visualizations

Figure 1 shows some sample data from our dataset.

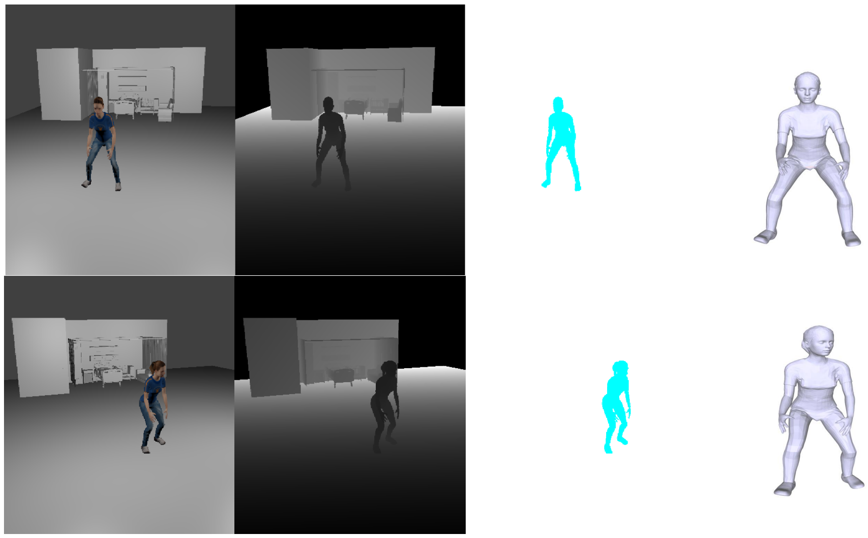

Figure 1: Some sample data from Subject2 is shown. For each row RGB image, depth image, foreground mask, and corresponding live geometry mesh is shown from left to right.

References

[1] M. Innmann, M. Zollhöfer, M. Nießner, C. Theobalt, and M. Stamminger, "VolumeDeform: Real-time volumetric non-rigid reconstruction," in *European Conference on Computer Vision (ECCV)*, 2016, pp. 362–379.

[2] M. Slavcheva, M. Baust, D. Cremers, and S. Ilic, "Killingfusion: Non-rigid 3 D reconstruction without correspondences," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 3, no. 4, 2017, p. 7.

[3] R. A. Newcombe, D. Fox, and S. M. Seitz, "Dynamicfusion: Reconstruction and tracking of non-rigid scenes in real-time," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 343–352.

[4] S. Elanattil, P. Moghadam, S. Sridharan, C. Fookes, and M. Cox, "Non-rigid reconstruction with a single moving RGB-D camera," in *International Conference on Pattern Recognition (ICPR)*, 2018.

[5] S. Elanattil and P. Moghadam, “Synthetic data for non-rigid 3D reconstruction using a moving RGB-D camera. v2.” CSIRO. Data Collection, 2018.