A Diverse Real-World Dataset for Training and Evaluation on 3D Pointclouds

Florian Gawrilowicz (flgw@dtu.dk), roboimagedata.compute.dtu.dk

Technical University of Denmark, Department of Applied Mathematics and Computer Science

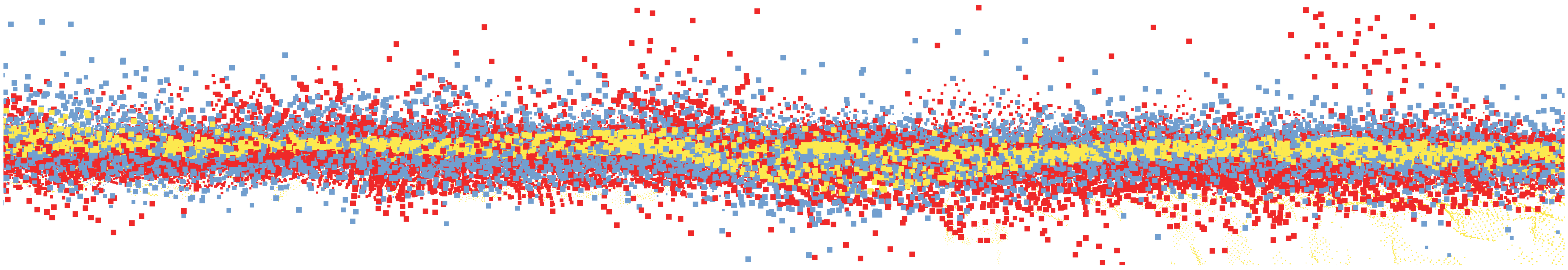

Figure 1: Super-imposed point clouds: Stuctured light scan (SLS), Multi-view stereo (MVS) by Furukawa et al. [2] and aligned using robust implicit moving least squares [4] (RIMLS) processing. While the structured light results exhibits less outliers than the MVS approach, it still is noisy, whereas the combined processing produces results of much higher fidelity.

1 Idea

A **unique** and **diverse** dataset of **real-world** point clouds for post-processing of multi-view stereo results, e.g. denoising, normal estimation, and consolidation.

2 Unique

There are many datasets of CAD models (Shapenet, ModelNet40, SHREC15), or for semantic segmentation (ScanNet, Matterport). In orders to learn point properties, like normals, these are re-sampled and noise is added.
Our dataset provides aligned results of different scanning modalities, i.e., results of three different MVS algorithms (Campbell, Furukawa [2], Tola) in conjunction with structured light scans for each camera position.

3 Diverse

The Robo-Image Dataset [1], we build upon, includes 124 different scenes of every-day objects.

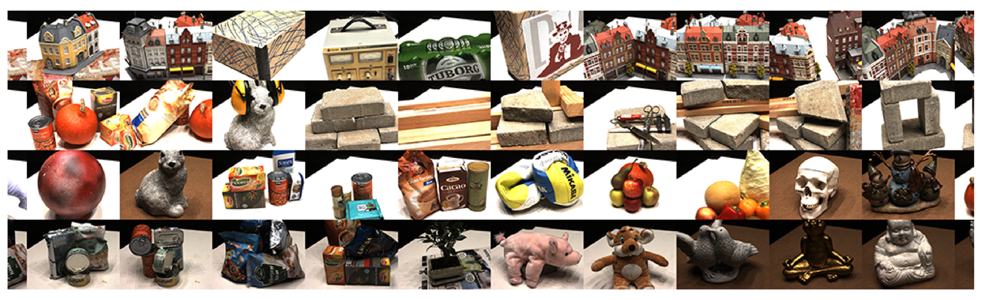

Figure 2: 40 of the 124 scenes in our data set

4 Real-world

For each scene, images from 49 or 64 calibrated camera positions under seven different lighting conditions are taken. This allows to generate new, or use the provided MVS results in a realistic setting, i.e., the scans exhibit realistic noise characteristics.

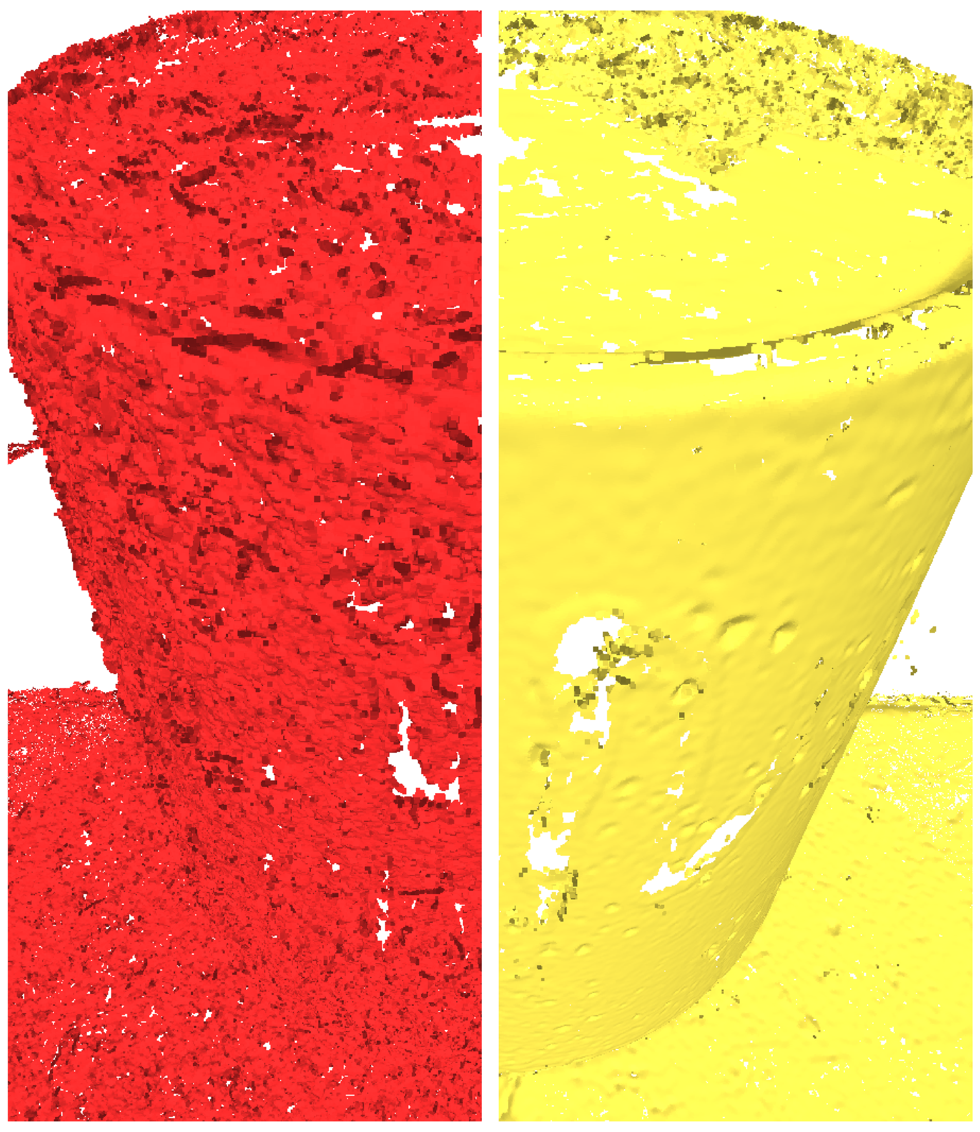

Figure 3: MVS vs. estimated ground truth

5 Contribution

SLS are generally of higher quality than MVS points, i.e. less noise. But there is no 1-to-1 correspondence between the two pointclouds. So it can't be used for training and evaluation of consolidation methods. Furthermore, noise in the positions results in noisy normals, because they usually have to be computed based on the noisy point positions.

To overcome these issues, we align and denoise the points in conjunction using the robust implicit moving least squares [4] (RIMLS) framework.

The result is a novel and challenging dataset for normal estimation (PCPNet [3], NormalNet, MC-Net, PCNN, DGCNN) and consolidation (EC-Net) of pointclouds.

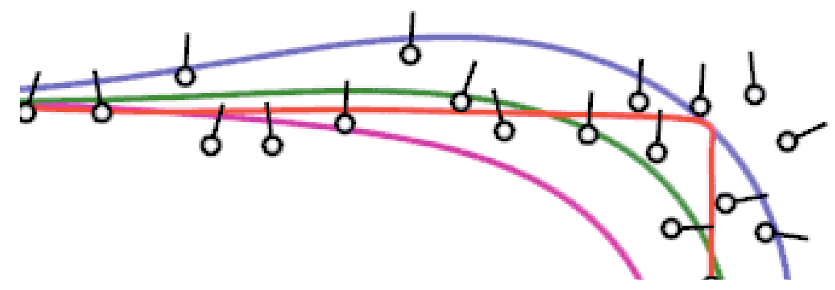

Figure 4: SPSS, APSS, IMLS, RIMLS

6 Example Training Results

To exemplify the usage, we trained PCPNet [3] on our data set. In Figure 5 the loss for training and test data is shown. It reaches a value of approximately 0.2 on the test data, which corresponds to an average difference of 16 degrees in unoriented normal direction.

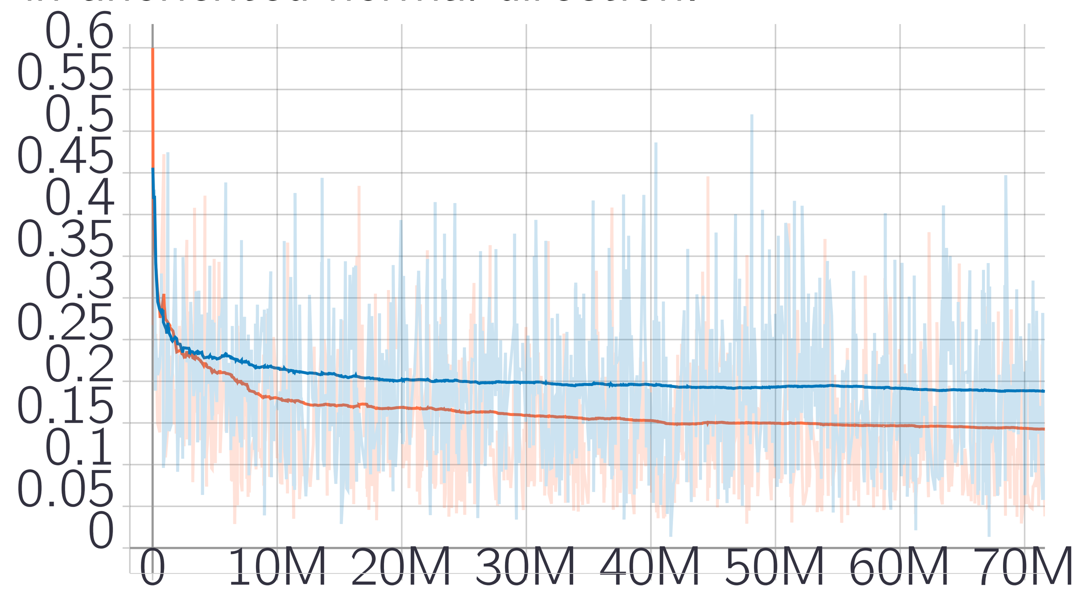

Figure 5: Loss on training (orange) and test (blue) data for PCPNet [3].

7 Future work

Evaluating more architectures, e.g. EC-Net, on our dataset. Formulating the alignment in a Bayesian framework to estimate a generative model.

References

[1] Henrik Aanæs et al. "Large-Scale Data for Multiple-View Stereopsis". In: *International Journal of Computer Vision* (2016), pp. 1–16.

[2] Yasutaka Furukawa and Jean Ponce. "Accurate, dense, and robust multiview stereopsis". In: *IEEE transactions on pattern analysis and machine intelligence* 32.8 (2010), pp. 1362–1376.

[3] Paul Guerrero et al. "PCPNet: Learning Local Shape Properties from Raw Point Clouds". In: *Computer Graphics Forum* 37.2 (2018), pp. 75–85. DOI: 10.1111/cgf.13343.

[4] A Cengiz Öztireli, Gael Guennebaud, and Markus Gross. "Feature preserving point set surfaces based on non-linear kernel regression". In: *Computer Graphics Forum*. Vol. 28. 2. Wiley Online Library. 2009, pp. 493–501.

DTU Compute
Department of Applied Mathematics and Computer Science
DTU Compute
Department of Applied Mathematics and Computer Science
DTU Compute
Department of Applied Mathematics and Computer Science