Supplementary Material: RigidFlow: Self-Supervised Scene Flow Learning on Point Clouds by Local Rigidity Prior

A. More results of pseudo scene flow labels

In Sec. 5.3, we compare the pseudo labels generated by our self-supervised learning method and the predictions of the models trained by our method in **EPE**. Here, we further compare them under the other three evaluation metrics, **AS**, **AR**, and **Out**. As the result shows in Fig. 1, our observations in Sec. 5.3 still hold. Specifically, during the training, the quality of our generated pseudo labels is gradually improved and consistently better than that of the predicted flow. Therefore, we can apply the generated pseudo labels as supervision. More qualitative results from our self-supervised learning method are provided in Sec. D.1.

We also compare pseudo labels and the model predictions in very early training stages on FT3D_s. As shown in Table 1, at the beginning of training, although the model outputs are inaccurate, the generated pseudo labels are almost better than the outputs, which enables training.

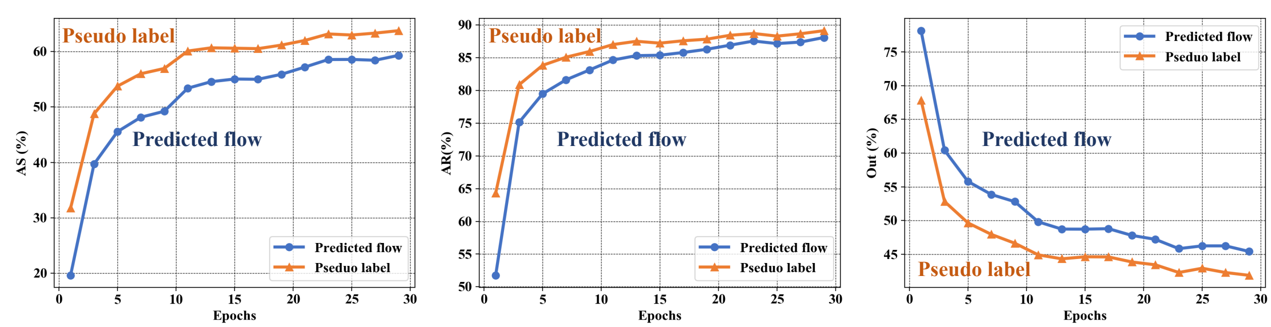
B. More experimental details

All experiments are conducted on the large-scale synthetic FlyingThings3D [6] dataset and the real-world KITTI 2015 [7, 8] dataset. Since 3D data are not directly provided by the two original datasets, we use the processed datasets prepared by HPLFlowNet [2] and FlowNet3D [5].

Our proposed pseudo label generation method is implemented based on PyTorch [9], FLOT [10], and S3DPC [4]. And we adopt FLOT [10] and FlowNet3D [5] as the scene flow estimation model in our work.

Table 1. Comparison of model outputs and pseudo labels in very early training stages on FT3D_s. Flows are evaluated on **EPE** ↓.

Iteration number of training	0 iter	10 iter	20 iter	40 iter	80 iter
Model outputs	2.9858	1.4627	1.2006	1.0095	0.7671
Pseudo labels	**2.7576**	**1.1758**	**1.0547**	**0.8794**	**0.5963**



(a) **AS** ↑ (b) **AR** ↑ (c) **Out** ↓

Figure 1. The comparison between model predictions (**Blue curve**) and generated pseudo labels (**Orange curve**) on training samples. During the training, the quality of the pseudo labels is consistently better than that of the predicted flow. This allows us to apply the pseudo labels as supervision.

Table 2. Licenses of the assets used in our paper.

Assets	License websites
PyTorch [9]	https://github.com/pytorch/pytorch/blob/master/LICENSE
FLOT [10]	https://github.com/valeoai/FLOT/blob/master/LICENSE
S3DPC [4]	https://github.com/yblin/Supervoxel-for-3D-point-clouds
FlowNet3D [5]	https://github.com/xingyul/flownet3d/blob/master/LICENSE
FlyingThings3D [6]	http://lmb.informatik.uni-freiburg.de/Publications/2016/MIFDB16
KITTI 2015 [7,8]	http://www.cvlibs.net/datasets/kitti/

Table 3. Comparison with the self-supervised loss proposed in FlowStep3D [3] on KITTI_o.

Self-supervised method	**Training data**	**Prediction Model**	**EPE** ↓	**AS** ↑
Self-supervised loss in FlowStep3D [3]	KITTI_r	FLOT	0.126	47.4
Ours	KITTI_r	FLOT	**0.102**	**48.4**

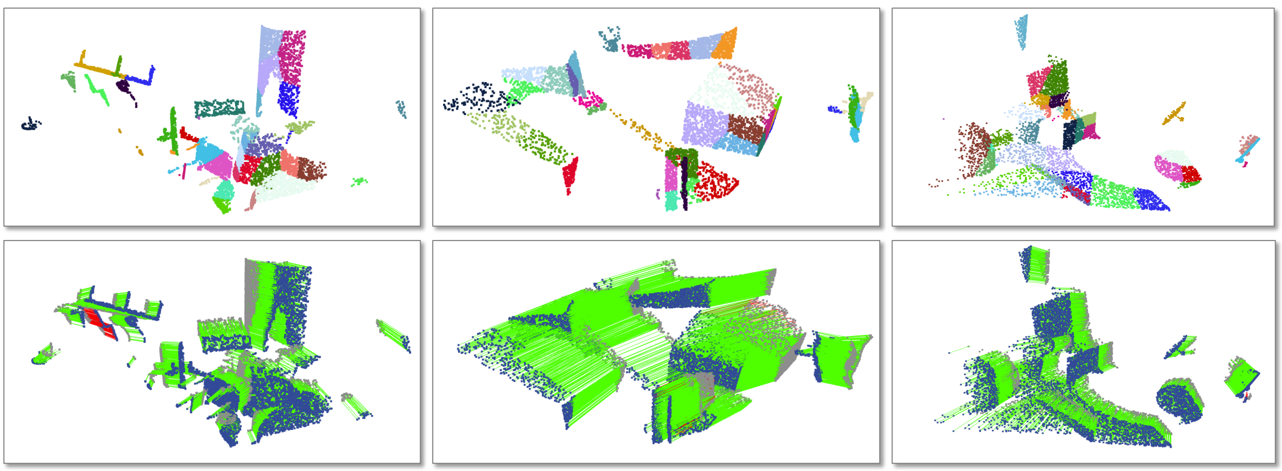


Figure 2. Pseudo scene flow labels generated by our method on $\text{FT3D}_\mathbf{s}$. Top: supervoxel results of the source point cloud. Different colors indicate different supervoxels. Bottom: generated pseudo scene flow labels. **Green line** indicates the **correct** pseudo label measured by **AR**. **Red line** indicates the **incorrect** pseudo label.

The licenses of the assets used in our paper are shown in Table 2.

C. Comparison with the self-supervised loss function in FlowStep3D

FlowStep3D [3] proposes a self-supervised loss function by combining Chamfer and Laplacian regularization losses. In Table 3, we compare our self-supervised learning method with the loss function proposed in FlowStep3D on KITTI_o. Under the same experimental setting, ours performs better.

D. More Visualization

D.1. Visualization of pseudo scene flow labels

In our paper, we train models with our self-supervised learning method on $\text{FT3D}_\mathbf{s}$ training set and $\text{KITTI}_\mathbf{r}$, respectively. Pseudo labels generated by our self-supervised learning method for some scenes in $\text{FT3D}_\mathbf{s}$ training set are shown in Fig 2. And some pseudo labels for $\text{KITTI}_\mathbf{r}$ are shown in Fig. 3.

D.2. Visualization of model predictions

We evaluate the models trained by our method on FT3D_s, KITTI_s, and KITTI_o. Qualitative results are shown in Fig 4.

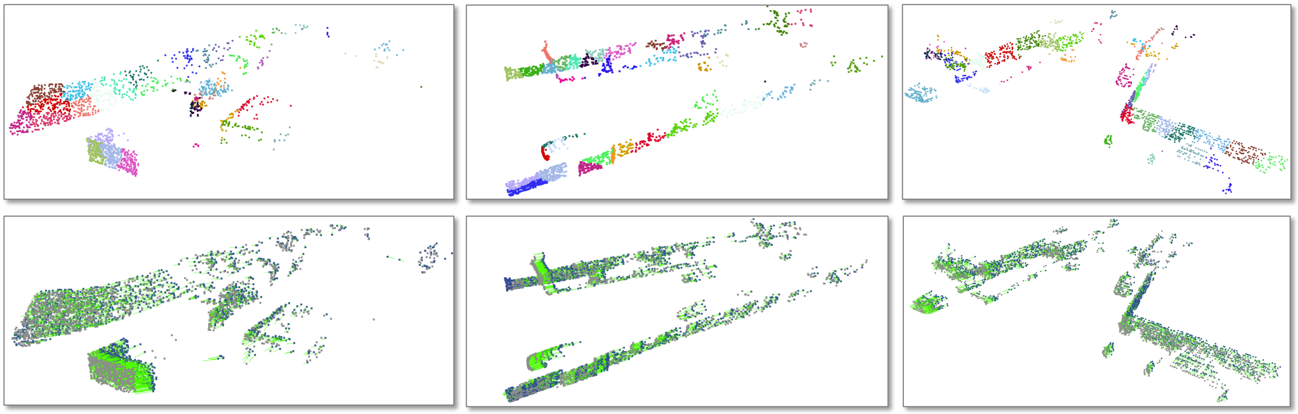


Figure 3. Pseudo scene flow labels generated by our method on $\text{KITTI}_{\mathbf{r}}$. Top: supervoxel results of the source point cloud. Different colors indicate different supervoxels. Bottom: generated pseudo scene flow labels. **Green line** indicates the generated pseudo label.

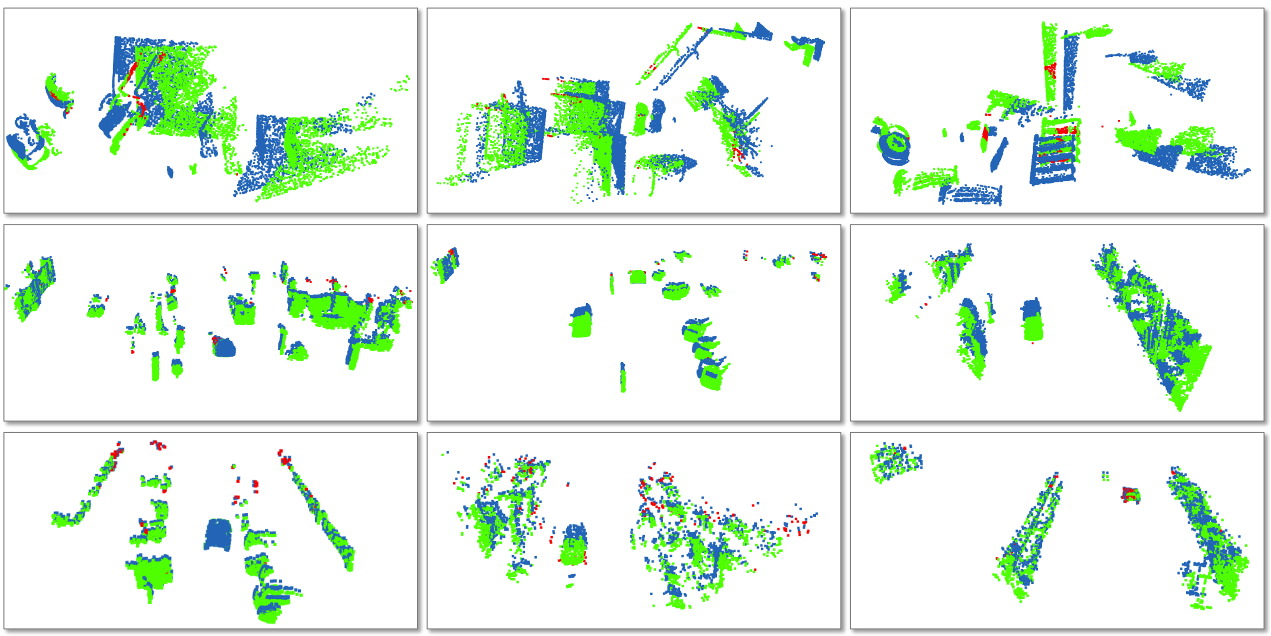


Figure 4. Qualitative results on FT3D_{s} (top), KITTI_{s} (middle), and KITTI_{o} (bottom). Bule points represent the source point cloud. Green points represent the points warped by the correct scene flow predictions. Red points represent the points warped by the incorrect predictions. The scene flow predictions are measured by **AR**.

D.3. Visualization of pseudo scene flow labels for non-rigid objects

In our method, we assume that the motion of each part in a real-world scene is rigid. Therefore, by adopting a small supervoxel size, an object will be over-segmented into several parts with different rigid parameters. When the object is non-rigid, our method would approximate the non-rigid motion as the rigid motions of parts and generate scene flow labels. Fig. 5 shows pseudo scene flow labels for some non-rigid objects in MPI Sintel dataset [1]. Here, we use the network trained on FT3D_{s} by our self-supervised method to predict flow and feed the flow to our piecewise pseudo label generation module to produce pseudo labels for these non-rigid objects.

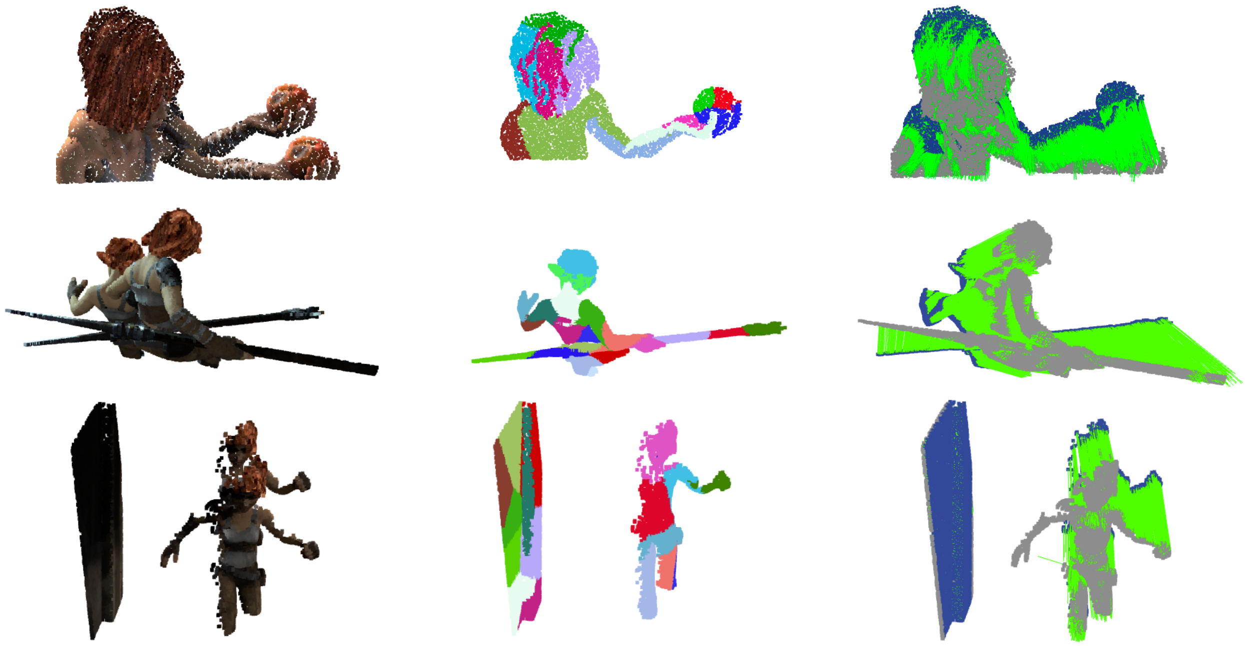


Figure 5. Pseudo scene flow labels generated by our method for some non-rigid objects in MPI Sintel dataset. Left: input point clouds. Middle: supervoxel results of the source point cloud. Different colors indicate different supervoxels. Right: generated pseudo scene flow labels. **Green line** indicates the generated pseudo label.

References

[1] Daniel J Butler, Jonas Wulff, Garrett B Stanley, and Michael J Black. A naturalistic open source movie for optical flow evaluation. In *European conference on computer vision*, pages 611–625. Springer, 2012. 3

[2] Xiuye Gu, Yijie Wang, Chongruo Wu, Yong Jae Lee, and Panqu Wang. Hplflownet: Hierarchical permutohedral lattice flownet for scene flow estimation on large-scale point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3254–3263, 2019. 1

[3] Yair Kittenplon, Yonina C Eldar, and Dan Raviv. Flowstep3d: Model unrolling for self-supervised scene flow estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4114–4123, 2021. 2

[4] Yangbin Lin, Cheng Wang, Dawei Zhai, Wei Li, and Jonathan Li. Toward better boundary preserved supervoxel segmentation for 3d point clouds. *ISPRS journal of photogrammetry and remote sensing*, 143:39–47, 2018. 1, 2

[5] Xingyu Liu, Charles R Qi, and Leonidas J Guibas. Flownet3d: Learning scene flow in 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 529–537, 2019. 1, 2

[6] Nikolaus Mayer, Eddy Ilg, Philip Hausser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4040–4048, 2016. 1, 2

[7] Moritz Menze and Andreas Geiger. Object scene flow for autonomous vehicles. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3061–3070, 2015. 1, 2

[8] Moritz Menze, Christian Heipke, and Andreas Geiger. Joint 3d estimation of vehicles and scene flow. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2, 2015. 1, 2

[9] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017. 1, 2

[10] Gilles Puy, Alexandre Boulch, and Renaud Marlet. Flot: Scene flow on point clouds guided by optimal transport. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVIII 16*, pages 527–544. Springer, 2020. 1, 2