**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. 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. Flows are evaluated on EPE ↓.

<table>
<thead>
<tr>
<th>Iteration number of training</th>
<th>0 iter</th>
<th>10 iter</th>
<th>20 iter</th>
<th>40 iter</th>
<th>80 iter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model outputs</td>
<td>2.9858</td>
<td>1.4627</td>
<td>1.2006</td>
<td>1.0095</td>
<td>0.7671</td>
</tr>
<tr>
<td>Pseudo labels</td>
<td>2.7576</td>
<td>1.1758</td>
<td>1.0547</td>
<td>0.8794</td>
<td>0.5963</td>
</tr>
</tbody>
</table>

![Figure 1](image.png)

(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.

<table>
<thead>
<tr>
<th>Assets</th>
<th>License websites</th>
</tr>
</thead>
</table>

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

<table>
<thead>
<tr>
<th>Self-supervised method</th>
<th>Training data</th>
<th>Prediction Model</th>
<th>EPE ↓</th>
<th>AS ↑</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-supervised loss in FlowStep3D [3]</td>
<td>KITTIr</td>
<td>FLOT</td>
<td>0.126</td>
<td>47.4</td>
</tr>
<tr>
<td>Ours</td>
<td>KITTIo</td>
<td>FLOT</td>
<td>0.102</td>
<td>48.4</td>
</tr>
</tbody>
</table>

Figure 2. Pseudo scene flow labels generated by our method on FT3Dₘ. 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 KITTIo. 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 FT3Dₘ training set and KITTIr, respectively. Pseudo labels generated by our self-supervised learning method for some scenes in FT3Dₘ training set are shown in Fig 2. And some pseudo labels for KITTIr are shown in Fig. 3.

D.2. Visualization of model predictions

We evaluate the models trained by our method on FT3Dₘ, KITTIₘ, and KITTIo. Qualitative results are shown in Fig 4.
Figure 3. Pseudo scene flow labels generated by our method on KITTI. 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 FT3Ds (top), KITTI (middle), and KITTO (bottom). Blue 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 FT3Ds 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


