FlipNeRF: Flipped Reflection Rays for Few-shot Novel View Synthesis  
—Supplementary Material—

Seunghyeon Seo  Yeonjin Chang  Nojun Kwak  
Seoul National University  
{zzzlssh, yjean8315, nojunk}@snu.ac.kr

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4-view</td>
<td>8-view</td>
<td>3-view</td>
<td>6-view</td>
</tr>
<tr>
<td>Learning Rate</td>
<td>$[1 \times 10^{-3}, 1 \times 10^{-5}]$</td>
<td>$[2 \times 10^{-3}, 2 \times 10^{-5}]$</td>
<td>1ere-2</td>
</tr>
<tr>
<td>Warm-up Step</td>
<td>512</td>
<td>1024</td>
<td>512</td>
</tr>
<tr>
<td>Delay Multiplier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_1$ (for $L_{\text{NLL}}$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_2$ (for $L'_{\text{NLL}}$)</td>
<td>$[4 \times 10^{-1}, 1 \times 10^{-4}]$</td>
<td>$[4 \times 10^{-2}, 1 \times 10^{-5}]$</td>
<td>$[4 \times 10^{-1}, 1 \times 10^{-4}]$</td>
</tr>
<tr>
<td>$\lambda_3$ (for $L_{\text{UE}}$)</td>
<td>$[1 \times 10^{-4}, 1 \times 10^{-1}]$</td>
<td>$[1 \times 10^{-5}, 1 \times 10^{-2}]$</td>
<td>$[1 \times 10^{-4}, 1 \times 10^{-1}]$</td>
</tr>
<tr>
<td>$\lambda_4$ (for $L'_{\text{UE}}$)</td>
<td>1ere-2</td>
<td>1ere-3</td>
<td>1ere-3</td>
</tr>
<tr>
<td>$\lambda_5$ (for $L_{\text{BFC}}$)</td>
<td>1ere-1</td>
<td>1ere-2</td>
<td>1ere-1</td>
</tr>
<tr>
<td>$\lambda_6$ (for $L_{\text{Ori.}}$)</td>
<td>1ere-1</td>
<td>1ere-2</td>
<td>1ere-1</td>
</tr>
</tbody>
</table>

Table A: Details of hyperparameters and loss balancing terms. For each dataset, the more training views are provided, the smaller $\lambda$’s we set to prevent over-regularization, except $\lambda_1$ for $L_{\text{NLL}}$. $[a, b]$ indicates the annealing from $a$ to $b$.

A. Implementation Details

The detailed hyperparameters and our loss balancing terms by the datasets and the number of training views are provided in Tab. A.

B. Explicit Normalization for the Estimated Surface Normals

As mentioned in Sec. 3.2, we use the weighted sum of blending weights and estimated normal vectors along a ray, i.e. $\hat{n} = \sum_{i=1}^{n} w_i n_i$, as the surface normals $\hat{n}$ to derive a flipped reflection direction $d'$. In this formulation, $\hat{n}$ is not guaranteed to be a unit vector without an explicit normalization process. However, we empirically found that the normalization rather destabilizes the training and leads to the performance degradation as shown in Fig. A. We conjecture that the inaccurately generated $\hat{n}'$ with an explicit normalization provides wrong supervisory signals, especially during the initial training stage when the surface normals are not accurately estimated. Furthermore, as illustrated in Fig. B, $\hat{n}$ is trained naturally to approximate the unit vector through the training without an explicit normalization process. With our proposed loss terms and regularization techniques for accurate normal estimation, we are able to use $\hat{n}$ as the surface normal vector without any additional...
initial phase
final phase

Figure B: Distribution of $\|\hat{n}\|_2$ without an explicit normalization process. At an initial training stage, the estimated surface normals $\hat{n}$ are not unit vectors as most of $\|\hat{n}\|_2$ are far from 1. However, our proposed $\hat{n}$ is trained to be a unit vector naturally through the training without an explicit normalization process, and most of $\|\hat{n}\|_2$ are concentrated close to 1 after the training, which indicates that $\hat{n}$ is successfully approximated to a unit vector.

C. Additional Qualitative Results

The additional qualitative comparisons are provided in Fig. C, Fig. D, and Fig. E. Furthermore, we provide more qualitative results of our FlipNeRF in Fig. F, Fig. G, and Fig. H. Our FlipNeRF shows superior rendering quality compared to other baselines.
Figure C: Additional qualitative comparisons on Realistic Synthetic 360°.

Figure D: Additional qualitative comparisons on DTU.

Figure E: Qualitative comparisons on LLFF 3-view.
Figure F: Additional qualitative results of our FlipNeRF on Realistic Synthetic 360°.
Figure G: Additional qualitative results of our FlipNeRF on DTU.
Figure H: Additional qualitative results of our FlipNeRF on LLFF 3-view.
References


