# FourieRF: Few-Shot NeRFs via Progressive Fourier Frequency Control

## Supplementary Material

## A. Supplementary Results on NeRF Synthetic and LLFF dataset

First, we note that as part of the supplementary materials we have **included a interactive website**, which provides qualitative comparisons with two closest baselines to our method (methods that have comparable running times): TensoRF [3] and ZeroRF [18]. As can be seen in the provided website, our method significantly outperforms those baselines in terms of quality of reconstructions in the few-shot setting, and has fewer artefacts, especially when considering very few input views. We provide qualitative comparisons both in terms of the novel view synthesis (RGB) as well as depth estimation compared to these baseline approaches. We encourage the reader to consider the videos provided in the interactive website (please allow a few seconds to load the videos) to see the improvement provided by our method.

In addition, below we include details on the statistics of our evaluations on the LLFF dataset in tables 4,5,6 and on the NeRF synthetic dataset in tables 7 and 8. For the LLFF dataset we reproduced the ZeroRF experiments to obtain the per scene score.

## B. Details on Implementation and settings

Finally, we provide details surrounding the settings of our implementation and experiments. All of our experiments were run in an nvidia RTX 4090 graphics card. We build our code base in top of the TensoRF [3] repository. Our repository can be found in the following link: https://github.com/diego1401/FourieRF.

Our method uses the AdamW optimizer [9, 11] with $\beta_1 = 0.9, \beta_2 = 0.98$, and a weight decay of $0.2$ for synthetic scenes and $0$ for real scenes. When performing frequency-control on real scenes we have to deal with matter appearing in front of the camera as a form of overfitting, as highlighted by FreeNeRF [22]. We find that applying their occlusion regularization works without any modification in our pipeline, thus we use their hyper parameters to compute our metrics. Moreover, we note that it is also efficient to use the gradient scaling introduced in the Floaters No More paper [16], this approach does not require to set a hyper parameter. We train for 10k iterations to match with our baselines, mainly ZeroRF [18].

The key hyper parameters of our method differ in the synthetic and real datasets. This can be attributed to the fact that the synthetic dataset has a solid color, white background, which alters the behavior of our method.

For the synthetic dataset, the clipping threshold is initialized as $f_0 = 0.3$, and it is linearly increased with $\delta = \frac{1}{2000} = 2 \times 10^{-3}$. We use the same configuration parameters as TensoRF [3] with the following differences. We apply a TV loss (with weight $w_{TV} = 1.0$) on the appearance and density features. We find that setting the weight decay to $0.2$ in the optimizer is the key to removing floaters (in our method and in ZeroRF [18]).

For the real dataset, the clipping threshold is initialized as $f_0 = 0.01$, and it is linearly increased until the end of training, i.e. $\delta = \frac{1}{10000} = 10^{-4}$. We use the same configuration parameters as TensoRF [3] with the following differences. We apply a TV loss (with weight $w_{TV} = 1.0$) on the appearance and density features, and an L1 loss (with weight $w_{L1} = 10^{-4}$) on the density features. We find that applying the L1 loss in this type of scenes is more efficient than setting a weight decay for the optimizer.

Table 4. Details quantitative comparison on the LLFF real dataset 3 views.

| Method | Statistic | fortress | room | horns | orchids | leaves | fern | flower | trex | mean |
|---|---|---|---|---|---|---|---|---|---|---|
| FreeNeRF [22] | PSNR ↑ | 23.437 | 22.020 | 18.506 | 15.286 | 16.250 | 21.187 | 20.413 | 19.941 | 19.630 |
| | SSIM ↑ | 0.583 | 0.834 | 0.585 | 0.407 | 0.521 | 0.662 | 0.617 | 0.687 | 0.612 |
| | LPIPS ↓ | 0.319 | 0.190 | 0.355 | 0.377 | 0.350 | 0.286 | 0.291 | 0.297 | 0.308 |
| ZeroRF [18] | PSNR ↑ | 20.633 | 18.833 | 13.688 | 13.900 | 16.275 | 18.700 | 17.880 | 16.786 | 17.087 |
| | SSIM ↑ | 0.435 | 0.663 | 0.233 | 0.275 | 0.533 | 0.523 | 0.490 | 0.517 | 0.459 |
| | LPIPS ↓ | 0.386 | 0.392 | 0.612 | 0.527 | 0.398 | 0.422 | 0.423 | 0.451 | 0.451 |
| Ours | PSNR ↑ | 22.109 | 20.271 | 18.290 | 15.103 | 16.524 | 20.965 | 21.062 | 20.103 | 19.303 |
| | SSIM ↑ | 0.573 | 0.792 | 0.627 | 0.422 | 0.587 | 0.667 | 0.674 | 0.745 | 0.636 |
| | LPIPS ↓ | 0.305 | 0.294 | 0.336 | 0.359 | 0.290 | 0.271 | 0.266 | 0.271 | 0.299 |

Table 5. Details quantitative comparison on the LLFF real dataset 6 views.

| Method | Statistic | fortress | room | horns | orchids | leaves | fern | flower | trex | mean |
|---|---|---|---|---|---|---|---|---|---|---|
| FreeNeRF [22] | PSNR ↑ | 28.728 | 27.302 | 23.592 | 17.263 | 19.047 | 24.647 | 24.665 | 24.596 | 23.730 |
| | SSIM ↑ | 0.832 | 0.910 | 0.792 | 0.555 | 0.685 | 0.796 | 0.797 | 0.864 | 0.779 |
| | LPIPS ↓ | 0.162 | 0.117 | 0.218 | 0.291 | 0.260 | 0.196 | 0.162 | 0.154 | 0.195 |
| ZeroRF [18] | PSNR ↑ | 23.767 | 27.083 | 19.188 | 14.425 | 18.475 | 23.533 | 21.780 | 21.957 | 21.276 |
| | SSIM ↑ | 0.802 | 0.880 | 0.606 | 0.318 | 0.670 | 0.753 | 0.712 | 0.796 | 0.692 |
| | LPIPS ↓ | 0.195 | 0.211 | 0.387 | 0.519 | 0.319 | 0.280 | 0.277 | 0.279 | 0.308 |
| Ours | PSNR ↑ | 29.031 | 28.792 | 23.273 | 17.484 | 19.187 | 24.466 | 24.510 | 22.019 | 23.595 |
| | SSIM ↑ | 0.878 | 0.920 | 0.815 | 0.558 | 0.727 | 0.792 | 0.822 | 0.810 | 0.790 |
| | LPIPS ↓ | 0.144 | 0.165 | 0.217 | 0.313 | 0.214 | 0.210 | 0.174 | 0.243 | 0.210 |

Table 6. Details quantitative comparison on the LLFF real dataset 9 views.

| Method | Statistic | fortress | room | horns | orchids | leaves | fern | flower | trex | mean |
|---|---|---|---|---|---|---|---|---|---|---|
| FreeNeRF [22] | PSNR ↑ | 29.421 | 29.927 | 25.154 | 19.083 | 20.678 | 26.073 | 26.182 | 24.522 | 25.130 |
| | SSIM ↑ | 0.865 | 0.938 | 0.846 | 0.662 | 0.756 | 0.831 | 0.843 | 0.875 | 0.827 |
| | LPIPS ↓ | 0.124 | 0.091 | 0.174 | 0.237 | 0.222 | 0.159 | 0.133 | 0.139 | 0.16 |
| ZeroRF [18] | PSNR ↑ | 24.350 | 26.883 | 21.675 | 16.125 | 19.200 | 24.400 | 23.240 | 24.629 | 22.563 |
| | SSIM ↑ | 0.797 | 0.903 | 0.733 | 0.465 | 0.700 | 0.787 | 0.762 | 0.850 | 0.750 |
| | LPIPS ↓ | 0.195 | 0.189 | 0.314 | 0.424 | 0.300 | 0.242 | 0.250 | 0.229 | 0.268 |
| Ours | PSNR ↑ | 29.567 | 29.011 | 24.799 | 19.046 | 20.839 | 25.774 | 26.488 | 24.562 | 25.011 |
| | SSIM ↑ | 0.881 | 0.931 | 0.860 | 0.636 | 0.775 | 0.825 | 0.854 | 0.876 | 0.830 |
| | LPIPS ↓ | 0.153 | 0.171 | 0.194 | 0.283 | 0.200 | 0.187 | 0.158 | 0.198 | 0.193 |

Table 7. Details quantitative comparison on the NeRF synthetic dataset 4 views.

| Method | Statistic | chair | drums | ficus | hotdog | lego | materials | mic | ship | mean |
|---|---|---|---|---|---|---|---|---|---|---|
| FreeNeRF [22] | PSNR ↑ | 20.22 | 14.99 | 17.35 | 23.58 | 20.43 | 21.36 | 15.05 | 17.52 | 18.81 |
| | SSIM ↑ | 0.843 | 0.746 | 0.809 | 0.899 | 0.818 | 0.857 | 0.802 | 0.687 | 0.808 |
| | LPIPS ↓ | 0.109 | 0.280 | 0.144 | 0.108 | 0.156 | 0.174 | 0.218 | 0.318 | 0.188 |
| ZeroRF [18] | PSNR ↑ | 23.04 | 16.91 | 20.12 | 29.11 | 22.11 | 20.50 | 24.76 | 19.01 | 21.94 |
| | SSIM ↑ | 0.880 | 0.791 | 0.866 | 0.944 | 0.868 | 0.848 | 0.944 | 0.707 | 0.856 |
| | LPIPS ↓ | 0.074 | 0.131 | 0.100 | 0.075 | 0.085 | 0.132 | 0.050 | 0.256 | 0.113 |
| Ours | PSNR ↑ | 24.13 | 17.33 | 18.56 | 27.26 | 22.41 | 21.15 | 23.35 | 19.64 | 21.73 |
| | SSIM ↑ | 0.895 | 0.804 | 0.848 | 0.933 | 0.871 | 0.858 | 0.929 | 0.724 | 0.858 |
| | LPIPS ↓ | 0.107 | 0.206 | 0.120 | 0.088 | 0.122 | 0.129 | 0.056 | 0.283 | 0.139 |

Table 8. Details quantitative comparison on the NeRF synthetic dataset 6 views.

| Method | Statistic | chair | drums | ficus | hotdog | lego | materials | mic | ship | mean |
|---|---|---|---|---|---|---|---|---|---|---|
| FreeNeRF [22] | PSNR ↑ | 26.72 | 18.16 | 18.46 | 27.18 | 24.32 | 21.63 | 25.64 | 20.23 | 22.77 |
| | SSIM ↑ | 0.916 | 0.827 | 0.840 | 0.929 | 0.887 | 0.853 | 0.942 | 0.729 | 0.865 |
| | LPIPS ↓ | 0.071 | 0.176 | 0.161 | 0.096 | 0.132 | 0.202 | 0.066 | 0.290 | 0.149 |
| ZeroRF [18] | PSNR ↑ | 27.62 | 20.88 | 22.21 | 29.93 | 26.26 | 21.41 | 27.40 | 22.13 | 24.73 |
| | SSIM ↑ | 0.926 | 0.869 | 0.898 | 0.949 | 0.913 | 0.849 | 0.954 | 0.756 | 0.889 |
| | LPIPS | 0.074 | 0.131 | 0.100 | 0.075 | 0.085 | 0.132 | 0.050 | 0.256 | 0.113 |
| Ours | PSNR ↑ | 26.62 | 19.30 | 19.43 | 28.84 | 27.09 | 21.46 | 25.78 | 22.89 | 23.93 |
| | SSIM ↑ | 0.918 | 0.838 | 0.860 | 0.939 | 0.915 | 0.856 | 0.942 | 0.767 | 0.879 |
| | LPIPS ↓ | 0.095 | 0.182 | 0.124 | 0.108 | 0.103 | 0.141 | 0.072 | 0.261 | 0.136 |