Rethinking Deblurring Strategies for 3D Reconstruction: Joint Optimization vs. Modular Approaches

Akash Malhotra[1,2], Nacéra Seghouani[2], Ahmad Abu Saiid[3], Alaa Almatuwa[3] and Koumudi Ganepola[3]

[1]*Amadeus, Sophia Antipolis, France*

[2]*Université Paris-Saclay, LISN, France*

[3]*Université Paris-Saclay, CentraleSupélec, France*

{*akash.malhotra, nacera.seghouani*}*@lisn.fr,* {*ahmad.ahmad, alaa-jawad-abdulla-ali.almutawa,*

Keywords: Multiview Synthesis, 3D Reconstruction, Deblurring, Neural Radiance Fields (NeRF), Image Restoration.

Abstract: In this paper, we present a comparison between joint optimization and modular frameworks for addressing deblurring in multiview 3D reconstruction. Casual captures, especially with handheld devices, often contain blurry images that degrade the quality of 3D reconstruction. Joint optimization frameworks tackle this issue by integrating deblurring and 3D reconstruction into a unified learning process, leveraging information from overlapping blurry images. While effective, these methods increase the complexity and training time. Conversely, modular approaches decouple deblurring from 3D reconstruction, enabling the use of stand-alone deblurring algorithms such as Richardson-Lucy, DeepRFT, and Restormer. In this study, we evaluate the trade-offs between these strategies in terms of reconstruction quality, computational complexity, and suitability for varying levels of blur. Our findings reveal that modular approaches are more effective for low to medium blur scenarios, while Deblur-NeRF, a joint optimization framework, excels at handling extreme blur when computational costs are not a constraint.

1 INTRODUCTION

Multiview photorealistic 3D reconstruction, used in virtual reality, autonomous navigation, and visual effects, enables the creation of realistic 3D representations. These underlying 3D representations are often created by techniques such as Neural Radiance Fields (NeRF) (Mildenhall et al., 2020) and 3D Gaussian Splatting (Kerbl et al., 2023), which have introduced learning-based algorithms. However, these methods heavily rely on clean, high-quality input images that leads to poor performance for handheld captures, which often have out-of-focus blur and motion blur. These degradations are particularly problematic as they impair feature matching between views and introduce uncertainties in geometry estimation, leading to inconsistent or incomplete 3D reconstruction.

Several methods address blur through integration of deblurring directly into the reconstruction pipeline, such as Deblur-NeRF (Ma et al., 2021), BAD-NeRF (Wang et al., 2023), and PDRF (Peng and Chellappa, 2023). Although these approaches achieve higher reconstruction fidelity through joint optimization of image deblurring and scene representation, they come with significant drawbacks: increased number of parameters, longer training times, higher computational complexity and more complex model architectures. Alternatively, standalone deblurring algorithms can be used to preprocess the images before 3D Reconstruction. Recent deblurring methods have made substantial progress in enhancing the visual quality of noisy images (Zhang et al., 2022). However, their impact on downstream 3D reconstruction tasks remains unexplored.

This work evaluates two approaches for addressing blur in 3D reconstruction: the modular framework, where images are preprocessed with standalone deblurring models before training NeRF; and the joint-optimization framework, where deblurring and reconstruction are performed simultaneously in Deblur-NeRF (Ma et al., 2021) framework. Through experiments and complexity analysis on synthetic and real-world scenes, we quantify the trade-offs between reconstruction quality and computational complexity of the two approaches. For the modular pipeline, we evaluated both traditional algorithms such as Richardson-Lucy (Fish et al., 1995) and modern learning-based methods such as DeepRFT (Mao

Malhotra, A., Seghouani, N., Saiid, A. A., Almatuwa, A. and Ganepola, K.
Rethinking Deblurring Strategies for 3D Reconstruction: Joint Optimization vs. Modular Approaches.
DOI: 10.5220/0013378800003912

In *Proceedings of the 20th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2025) - Volume 3: VISAPP*, pages 816-823
ISBN: 978-989-758-728-3; ISSN: 2184-4321

et al., 2023) and Restormer (Zamir et al., 2022).

Our experiments and analysis reveal several key insights for practical applications:

1. For low to medium blur, also called decent blur, it is better to use the modular pipeline with DeepRFT(Mao et al., 2023) as the deblurring algorithm, especially in relation to the compute budget.
2. For images with extreme blur, DeblurNeRF is preferable, especially when the computing budget is not constrained.
3. Larger models for preprocessing in the modular framework are not always better, as evidenced by DeepRFT outperforming Restormer and other Transformer based methods.

The remainder of this paper is organized as follows. Section 2 reviews related work on various deblurring techniques. Section 3 provides background information on the fundamentals of NeRF and blur modeling. Section 4 details our methodology, including datasets, experimental setup, and evaluation metrics. Section 5 presents our experimental results and analysis of both joint optimization and modular approaches. Finally, Section 6 concludes on our main results for practical applications.

2 RELATED WORK

Research in addressing blur for 3D reconstruction has evolved from classical image restoration to modern neural approaches, with recent work focusing on joint optimization techniques.

Evolution of Image Deblurring

Early deblurring approaches relied on analytical methods such as Fourier-based techniques (Richardson, 1972) and Bayesian deconvolution (Fergus et al., 2006). While these methods established important theoretical foundations, they struggled with spatially varying blur and complex degradation patterns. Later studies explored blur detection (Koik and Ibrahim, 2013) and kernel estimation (Smith, 2012), while work on camera response functions (Grossberg and Nayar, 2004) provided foundational understanding of imaging systems. The field progressed to multi-image techniques that leveraged information from multiple views or frames (Li et al., 2023), showing improved results but requiring careful image alignment and registration.

Modern Deblurring Approaches

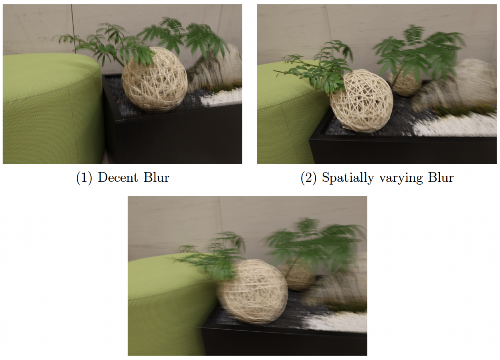
Deep learning has revolutionized image deblurring through two main approaches: single-image and multi-image methods. Single-image techniques have seen rapid advancement through architectures like Restormer (Zamir et al., 2022), which employs transformers for modeling long-range dependencies, multi-stage progressive restoration frameworks like MPRNet (Zamir et al., 2021), and DeepRFT (Mao et al., 2023), which integrates frequency-domain processing. These methods are often optimized for perceptual quality metrics and standard image quality assessments (Zhang et al., 2022) rather than downstream tasks. Multi-image approaches like BiT (Zhong et al., 2023) and GShift-Net (Li et al., 2023) leverage temporal consistency to handle complex motion blur patterns and maintain consistency across multiple views.

Joint Optimization with Neural Radiance Fields

The emergence of Neural Radiance Fields (NeRF) (Mildenhall et al., 2020) has spurred new approaches that jointly handle deblurring and 3D reconstruction. Research has shown that input image quality significantly impacts NeRF performance (Liang et al., 2023) (Rubloff, 2023). Deblur-NeRF (Ma et al., 2021) pioneered this direction by incorporating deformable sparse kernels into the NeRF framework. BAD-NeRF (Wang et al., 2023) extended this approach by integrating bundle adjustment, while PDRF (Peng and Chellappa, 2023) introduced progressive refinement. These methods achieve high-quality results but at the cost of increased computational complexity and training time.

Scope of this Work

While previous studies have advanced deblurring techniques or joint optimization frameworks independently, there is no systematic comparison of these approaches in the context of 3D reconstruction. Our work bridges this gap by evaluating when the added complexity of joint optimization frameworks is justified versus when simpler, modular solutions using state-of-the-art deblurring methods suffice. We analyze these trade-offs across different blur conditions and computational constraints, providing practical insights for choosing appropriate techniques in real-world applications.



(1) Decent Blur (2) Spatially varying Blur (3) Extreme Blur

Figure 1: Different types and levels of blur for the Blurball scene.

3 BACKGROUND

Neural Radiance Fields

Neural Radiance Fields (NeRF) (Mildenhall et al., 2020) provide a powerful paradigm for 3D reconstruction by representing a scene as a continuous 5D function that maps spatial coordinates and viewing directions to color and density. Formally, let (x,y,z) and (θ,ϕ) denote, respectively, a 3D location and a viewing direction. NeRF learns a function:

$$(c,\sigma) = F_\Theta(\gamma(x,y,z),\phi(\theta,\phi)) \tag{1}$$

where c is the emitted color, σ is the volume density, $\gamma(\cdot)$ is a positional encoding to map coordinates into a higher-dimensional space, and Θ are the learnable parameters of the neural network. Rendering a pixel color $C(r)$ of a ray r cast into the scene involves integrating the contributions of sampled points along that ray:

$$C(r) = \sum_{i=1}^{N} T_i\big(1-\exp(-\sigma_i\delta_i)\big)c_i \tag{2}$$

where

$$T_i = \exp\left(-\sum_{j=1}^{i-1}\sigma_j\delta_j\right) \tag{3}$$

and δ_i is the distance between consecutive sample points along the ray. By optimizing NeRF parameters to minimize the discrepancy between rendered and captured images, one can achieve high-fidelity novel view synthesis, provided the input images are sharp and noise-free.

However, real-world captures often contain motion blur due to camera or subject movement during exposure. Such blur distorts the observed pixel colors, hindering NeRF's ability to infer accurate scene geometry and radiance distributions. Standard NeRF, lacking mechanisms to account for blur, typically yields suboptimal reconstruction quality under these conditions.

Joint Optimization Framework

Addressing blur within the reconstruction pipeline can follow two main strategies: joint optimization and modular frameworks. Joint optimization frameworks incorporate the modeling of blur directly into the NeRF training process. Deblur-NeRF (Ma et al., 2022) exemplifies this philosophy by introducing a Deformable Sparse Kernel (DSK) that models spatially varying blur kernels. Instead of a uniform convolution, Deblur-NeRF approximates the blur of a pixel p as a sparse weighted combination of neighboring colors:

$$b_p = c_p * h \tag{4}$$

where b_p and c_p are blurred and sharp pixel colors respectively, and h is a blur kernel. To improve computational efficiency, it leverages a sparse set of kernel points:

$$b_p = \sum_{q\in N(p)} w_q c_q \tag{5}$$

with $N(p)$ denoting the neighborhood of pixel p and w_q the learned weights. Additionally, Deblur-NeRF refines the ray origins for each pixel by introducing offsets Δo_q:

$$r_q = (o_q + \Delta o_q) + t\,d_q \tag{6}$$

allowing the model to compensate for spatially varying blur patterns. This joint optimization of NeRF parameters and kernel properties enables the network to restore sharpness while simultaneously improving reconstruction fidelity, effectively learning

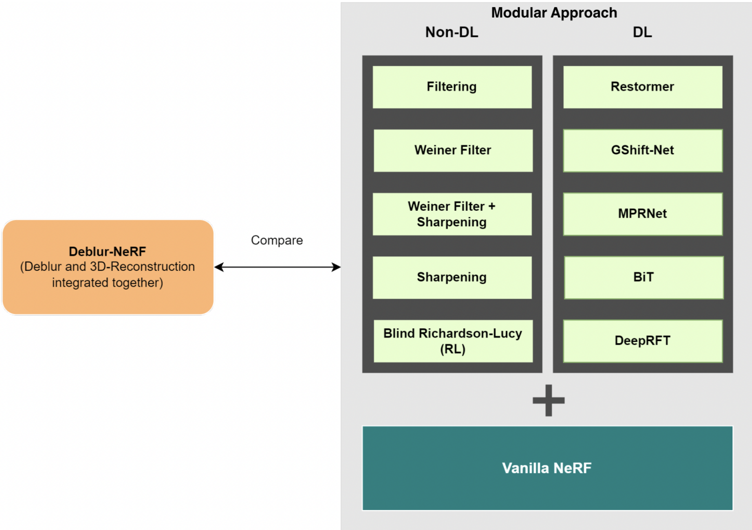


Figure 2: Experimental workflow. Comparison of DeblurNeRF with the modular approach. Deblurring and 3D reconstruction (using NeRF) are decoupled. Many non deep learning and deep learning based algorithms for deblurring are compared in the modular approach.

a scene representation that is robust to blur. The trade-off, however, is significantly increased complexity, computational overhead, and resource usage, potentially limiting scalability.

Modular Framework

In the modular framework, various deblurring techniques are employed as pre-processing steps to enhance image quality before 3D reconstruction. In this work, we evaluated and compared many approaches based on deep learning and non-deep learning, which are briefly explained in Section 2. They can take a single image or multiple images as inputs. Multiple images can be given as inputs if the images are in a form of a video or are positionally close to each other in a multiview setting.

Ultimately, the choice between joint optimization and modular solutions involves balancing computational complexity, model capacity, and reconstruction fidelity. Joint optimization methods such as Deblur-NeRF closely align the blur compensation process with scene representation but demand substantial computational resources. Modular pipelines, by contrast, allow users to exploit off-the-shelf deblurring models to preprocess images before training a standard NeRF, improving scalability and ease of use. Understanding these complementary strategies sets the stage for informed pipeline design, particularly as the field moves toward more practical and resource-efficient solutions for robust 3D reconstruction under real-world conditions of motion and other kind of blurs.

Metrics

When all or some of the reference images (sharp or noiseless) are available, the standard 3D reconstruction metrics PSNR, SSIM (Wang et al., 2004), and LPIPS (Zhang et al., 2018) are utilized to evaluate the reconstructions.

To compare the standalone deblurring algorithms used in the modular framework, we utilize FFT Blur Score (Rosebrock, 2020). It provides a frequency-domain perspective by quantifying the residual blur in images based on their high-frequency content and is calculated as:

$$\text{Blur Score} = \frac{\text{Max\_FFT}_d - \text{FFT\_Value}_i}{\text{Max\_Blur\_Dist}} \tag{7}$$

where Max_FFT_d is the maximum FFT score in the dataset, FFT_Value_i is the FFT score for the current image, and Max_Blur_Dist is an empirically determined constant that captures the maximum FFT distance between clear and blurry frames. This metric normalizes the blur score between 0 (sharpest) and 1 (most blurred). Unlike PSNR, SSIM, and LPIPS, which rely on reference images, the FFT Blur Score can evaluate blur independently, making it particularly useful in scenarios lacking sharp ground truth.

4 METHODOLOGY

Our methodology systematically evaluates the joint optimization framework using DeblurNeRF, and the

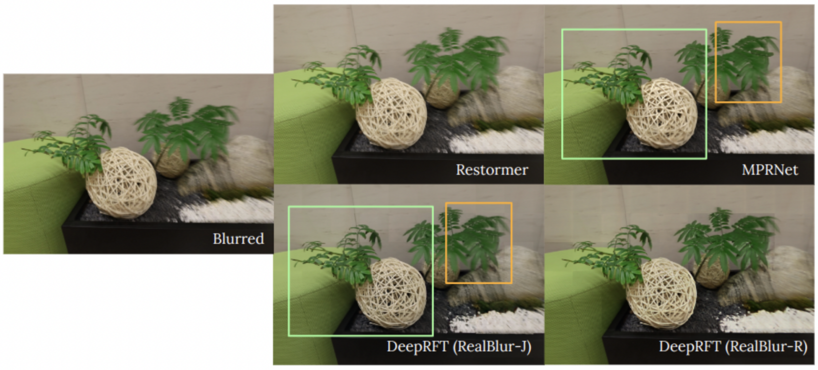


Figure 3: Qualitative comparison of different deblurring models on Blurball scene for spatially varying blur.

modular approach that combines standalone deblurring techniques with NeRF, as shown in Figure 2. We employed a two-stage evaluation process. In the first stage, using the synthetic Blurwine scene, we comprehensively tested a broad range of deblurring techniques in the modular approach. Among traditional deblurring algorithms, the sharpening filter, the Wiener filter, a combination of both, and the blind Richardson-Lucy were tested. From the deep learning approaches that act on a single image at a time, MPRNet, Restormer and DeepRFT were tested. We also considered multi-image or temporal models such as GShift-Net and the Blur Interpolation Transformer (BiT), which exploit additional frames or viewpoints to improve deblurring quality (cf. Section 3). In addition to applying GShift-Net to the video of the blurwine scene, we also tested GShift-Net with the neighboring images instead with respect to the camera positions, and call it ”GShift-Net Adjusted”.

Based on the performance results from the Blurwine scene, we selected the most promising algorithms - DeepRFT, Restormer, and MPRNet - for evaluation on the more challenging real-world scenes (Blurball and Blurobject). This selective testing approach allowed us to focus computational resources on the most effective methods while maintaining experimental rigor. After applying these methods on the input images, we use NeRF for 3D reconstruction and record the reconstruction quality. In parallel, we use DeblurNeRF as a joint optimization framework and record its reconstruction quality. The comparison between them is discussed in Section 5.

Our experiments utilize the following three scenes, each designed to evaluate the performance of joint optimization and modular frameworks under varying blur conditions. Examples of different types and levels of blur are shown in Figure 1. The scenes are as follows:

1. **Blurwine:** Introduced by (Ma et al., 2021), this synthetic motion blur scene consists of 34 images, split into 29 for training and 5 for testing. Images were generated with controlled motion blur to facilitate quantitative evaluation against ground truth. Each scene contains both blurred and sharp (reference) images.
2. **Blurball:** Also introduced by (Ma et al., 2021), it is a real-world blurry scene with 27 images, split into 23 for training and 4 for testing. These images were captured under extreme motion blur conditions using deliberate camera shake. Ground truth reference images were captured using a tripod setup to ensure stability.
3. **Blurobject:** For the current study, we created a novel real-world motion blur scene, containing 33 images, divided into 28 for training and 5 for testing. These images were captured using a Canon 2000D camera under manual exposure, introducing mild blur levels reflective of everyday scenarios. Unlike Blurball, these images were not derived from videos but from individual image captures for a generalized scenario.

Each scene includes a combination of sharp and blurred reference images. The blurred images reflect varying levels of motion blur, ranging from controlled synthetic settings in Blurwine to moderate and extreme real-world conditions in Blurobject and Blurball, respectively. Scenes were processed using COLMAP (Schonberger and Frahm, 2016) scripts to compute camera poses, which are also given as input to NeRF.

We employ multiple complementary metrics to comprehensively evaluate reconstruction quality. For the scene with ground truth images (Blurwine), we use PSNR, SSIM, and LPIPS to evaluate reconstructions. For real-world scenes (Blurball and Blurobject), we assess reconstruction quality using the subset of sharp images as reference. In the absence of sharp reference images, as in the case of Blurobjects scene, we employ FFT Blur Score. These metrics are explained in Section 3. All experiments maintain consistent NeRF training protocols. We also analyze computational efficiency through floating point operations per second (FLOPs) counts, memory usage, and total training time to understand the practical implications of each approach.

Figure 4: Qualitative comparison of different deblurring models on Blurobject scene.

Table 1: Blur FFT Scores on Blurobject scene (Lower scores are better).

Deblurring Technique	**Blur FFT Score**
Original	0.71
Restormer	0.54
MPRNet	0.57
DeepRFT	**0.42**

5 RESULTS AND DISCUSSION

Our experiments evaluated the effectiveness of joint optimization and modular frameworks in handling motion blur for 3D reconstruction across synthetic and real-world scenes.

As shown in Table 2, in the Blurwine scene, with controlled blur, Deblur-NeRF's joint optimization strategy achieved superior reconstruction quality with a PSNR of 27.47, SSIM of 0.86, and LPIPS of 0.14. Among modular approaches, DeepRFT performed best with a PSNR of 22.89, SSIM of 0.73, and LPIPS of 0.23. Traditional methods like Wiener filter and Blind Richardson-Lucy performed poorly, often worse than the original blurred inputs used with NeRF, possibly due to their inability to model complex, spatially varying blur patterns. While GShift-Net showed some improvement over traditional methods (PSNR: 20.49), it still lagged significantly behind other deep learning approaches.

Table 2: Quantitative comparison of various deblurring techniques on the Blurwine scene.

Model	**PSNR**	**SSIM**	**LPIPS**
NeRF	21.11	0.63	0.36
Deblur-NeRF	**27.47**	**0.86**	**0.14**
Filtering	10.33	0.10	0.63
Sharpening	21.53	0.67	0.28
Wiener Filter	19.75	0.54	0.41
Wiener + Sharpening	19.67	0.54	0.36
Richardson-Lucy	20.84	0.61	0.34
Restormer	22.88	0.73	0.23
GShift-Net	20.49	0.61	0.34
GShiftNet Adjusted	21.46	0.66	0.31
MPRNet	22.36	0.70	0.26
BiT	16.85	0.47	0.35
DeepRFT	**22.89**	**0.73**	**0.23**

As shown in Table 3, for the Blurball scene with extreme motion blur, Deblur-NeRF maintained strong performance (PSNR: 27.39, SSIM: 0.77) through its joint optimization of NeRF parameters and spatially varying blur kernels. DeepRFT demonstrated robustness to severe blur (PSNR: 24.90, SSIM: 0.66). However, in the Blurobject scene with moderate blur, DeepRFT slightly outperformed Deblur-NeRF with a PSNR of 22.11 and SSIM of 0.47, compared to Deblur-NeRF's PSNR of 21.14 and SSIM of 0.44.

Our experiments revealed several unexpected findings. DeepRFT, despite having a smaller model size than Restormer, consistently achieved better reconstruction quality across all scenes. This is also seen in Table 1. This suggests that incorporating Fourier domain processing into neural architectures can be more effective than simply increasing model capacity. Another surprising result was that single-image methods (DeepRFT and Restormer) outperformed multi-image approaches like BiT and GShift-Net. While BiT achieved a PSNR of 16.85 and GShift-Net 20.49 on the synthetic scene, DeepRFT and Restormer achieved 22.89 and 22.88 respectively, indicating that additional temporal information did not necessarily translate to better deblurring performance for 3D reconstruction.

The computational analysis reveals significant differences between the approaches, as can be seen in Table 4. Deblur-NeRF requires substantial resources due to its deformable sparse kernel optimization and increased ray rendering, resulting in approximately five times the FLOPs per pixel compared to VanillaNeRF. This complexity demands at least 32 GB memory and training times spanning multiple days for large scenes. In contrast, the modular approach

Table 3: Reconstruction results on Blurball and Blurobject scene without any Deblurring (NeRF), with DeblurNeRF and with modular approach using vaious deblurring models.

Method	**Blurball**			**Blurobject**		
	PSNR	**SSIM**	**LPIPS**	**PSNR**	**SSIM**	**LPIPS**
NeRF	24.03	0.62	0.40	17.32	0.32	0.44
Deblur-NeRF	**27.38**	**0.77**	**0.24**	21.14	0.44	0.39
Restormer	23.17	0.61	0.41	21.51	0.47	0.32
MPRNet	23.24	0.62	0.39	21.57	0.47	0.33
DeepRFT	24.89	0.66	0.35	**22.11**	**0.47**	**0.29**

Table 4: Approximate computational complexity comparison of DeblurNeRF and modular approach with DeepRFT.

Metric	**Deblur-NeRF**	**DeepRFT + VanillaNeRF**	**DeepRFT**	VanillaNeRF
Training Time	4x	1x	Pretrained	Baseline (1x)
Memory Requirement	32 GB	16 GB	8 GB	8 GB
FLOPs per Pixel	5x	1.1x	0.1x	Baseline (1x)

using DeepRFT with VanillaNeRF is more efficient. DeepRFT's FFT-based operations and single forward pass during inference keep computational and memory requirements low while maintaining competitive reconstruction quality.

These results highlight the complementary strengths of each approach. Deblur-NeRF excels with extreme blur, but requires significant computational resources, limiting its scalability. The modular approach with DeepRFT offers a practical alternative, particularly effective for moderate blur scenarios and resource-constrained environments. For applications with extreme blur and abundant computational resources, Deblur-NeRF is optimal. However, when dealing with moderate blur or limited resources, the modular approach with DeepRFT provides an efficient and effective solution.

6 CONCLUSION

This study investigated the trade-offs between joint optimization and modular frameworks for mitigating blur in 3D reconstruction. Our experiments revealed that Deblur-NeRF excels at handling extreme blur through joint optimization, while the modular approach with DeepRFT offers an efficient alternative for moderate blur scenarios. Traditional methods proved inadequate for complex blur patterns found in the real world scenes, while modern deep learning methods showed better performance. Surprisingly, DeepRFT outperformed both the larger Restormer model and multi-image approaches like BiT and GShift-Net, suggesting that incorporating Fourier domain processing into neural networks is a promising yet underexplored direction.

Our contribution of the Blurobject scene dataset provides a compelling benchmark based on real world scenarios. This dataset fills a critical gap between synthetic and extreme blur datasets, providing a valuable resource for evaluating deblurring techniques. The findings emphasize the importance of matching deblurring strategies to application requirements.

Future work could extend this analysis to a broader range of scenes and newer 3D reconstruction techniques like Gaussian Splatting. The strong performance of DeepRFT could motivate further research into efficient architectures that can maintain high quality of reconstruction with reduced computational demands.

Supplementary. Please refer to the following github repository for more code and more details. https://github.com/AlaaAlmutawa/BDRP.

REFERENCES

Fergus, R., Singh, B., Hertzmann, A., Roweis, S. T., and Freeman, W. T. (2006). Removing camera shake from a single photograph. *ACM Transactions on Graphics (TOG)*, 25(3):787–794.

Fish, D. A., Brinicombe, A. M., Pike, E. R., and Walker, J. G. (1995). Blind deconvolution by means of the richardson-lucy algorithm. *JOSA A*, 12(1):58–65.

Grossberg, M. D. and Nayar, S. K. (2004). Modeling the space of camera response functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(10):1272–1282.

Kerbl, B., Kopanas, G., Leimkühler, T., and Drettakis, G. (2023). 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):139–1.

Koik, B. T. and Ibrahim, H. (2013). A literature survey on blur detection algorithms for digital imaging. In *2013 1st International Conference on Artificial Intelligence, Modelling and Simulation*, pages 272–277.

Li, D., Shi, X., Zhang, Y., Cheung, K.-T., See, S., Wang, X., and Li, H. (2023). A simple baseline for video restoration with grouped spatial-temporal shift. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9822–9832.

Liang, H., Wu, T., Hanji, P., Banterle, F., Gao, H., Mantiuk, R., and Oztireli, C. (2023). Perceptual quality assessment of nerf and neural view synthesis methods for front-facing views. *arXiv preprint arXiv:2303.15206*.

Ma, L., Li, X., Liao, Z., Zhang, Q., Wang, X., Wang, J., and Sander, P. V. (2021). Deblur-nerf: Neural radiance fields from blurry images. *arXiv preprint arXiv:2111.14292*.

Mao, X., Liu, Y., Liu, F., Li, Q., Shen, W., and Wang, Y. (2023). Intriguing findings of frequency selection for image deblurring. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 1905–1913.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. (2020). Nerf: Representing scenes as neural radiance fields for view synthesis. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 405–421. Springer.

Peng, C. and Chellappa, R. (2023). Pdrf: Progressively deblurring radiance field for fast scene reconstruction from blurry images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 2029–2037.

Richardson, W. H. (1972). Bayesian-based iterative method of image restoration. *JOSA*, 62(1):55–59.

Rosebrock, A. (2020). Opencv fast fourier transform (fft) for blur detection in images and video streams. Accessed: 2024-12-16.

Rubloff, M. (2023). What are the nerf metrics? Accessed: 2023-12-19.

Schonberger, J. L. and Frahm, J.-M. (2016). Structure-from-motion revisited. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4104–4113.

Smith, L. (2012). Estimating an image's blur kernel from edge intensity profiles. Technical report, Naval Research Laboratory.

Wang, P., Zhao, L., Ma, R., and Liu, P. (2023). Bad-nerf: Bundle adjusted deblur neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4170–4179.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612. Accessed: 2023-12-19.

Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., and Yang, M.-H. (2022). Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5728–5739.

Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., Yang, M.-H., and Shao, L. (2021). Multi-stage progressive image restoration. *CoRR*, abs/2102.02808.

Zhang, K., Ren, W., Luo, W., Lai, W.-S., Stenger, B., Yang, M.-H., and Li, H. (2022). Deep image deblurring: A survey. *arXiv preprint arXiv:2202.10881*.

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., and Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595.

Zhong, Z., Cao, M., Ji, X., Zheng, Y., and Sato, I. (2023). Blur interpolation transformer for real-world motion from blur. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5713–5723.