EigenGS Representation: From Eigenspace to Gaussian Image Space

Supplementary Material

Additional Results & Visualizations

The supplementary material includes additional visualizations and detailed experimental results that highlight key aspects of our method, where the visualization is primarily using the FFHQ dataset. All experimental settings are aligned with those described in the main paper. Section I presents comprehensive quantitative and qualitative results across different datasets. Our method is able to achieve consistent performance across diverse image categories and resolutions. Section II provides an analysis of our frequency-aware learning mechanism. We show the advantage of color space selection in Section III.

I. Quantitative and Qualitative Evaluations

We present comprehensive quantitative and qualitative results across our four benchmark datasets to show the effectiveness of the ImageNet-trained, universal EigenGS model. Figures I to IV show the visualization results using our ImageNet-trained EigenGS, which is noteworthy as it was trained on an unaligned collection of images with arbitrary sizes. The successful adaptation of this universal model to different datasets highlights the robustness of our approach.

Tables I to IV provide detailed performance metrics comparing three scenarios: the baseline GaussianImage method, dataset-specific trained EigenGS, and our ImageNet-trained universal EigenGS. In that, the dataset-specific trained EigenGS can represent the performance upper bound of our method, while the scores of ImageNet version corresponds to visualization. Overall, though the ImageNet-trained model shows a slight performance decrease compared to dataset-specific training, it consistently outperforms the baseline method across all datasets.

For CelebA, the reconstruction shows excellent preservation of facial features even at early iterations. Meanwhile, the FFHQ results demonstrate robust handling of high-resolution facial details. The Cats dataset shows our method's ability to capture broader structural elements. Lastly, despite the challenging variety of viewpoints in the Stanford Cars dataset, our method can still achieve high-quality reconstruction. These results particularly confirm that our EigenGS representation maintains consistent performance across different datasets even when using the ImageNet-trained model, which suggests that our approach successfully captures universal image statistics and generalizes well across domains.

II. Spatial Frequency

Our frequency-aware learning mechanism reveals interesting characteristics in how different spatial frequencies are handled in image reconstruction. Through visualization of the Gaussian components shown in Figure V, we can observe the larger spatial extent of the low-frequency Gaussians, $\mathcal{N}_l$, indicating their role in capturing larger structural elements. In contrast, the high-frequency set $\mathcal{N}_h$ consists of smaller Gaussians that provide fine detail refinement, as evidenced by their more compact elliptical boundaries.

The low-frequency set $\mathcal{N}_l$ comprises approximately 10% of the overall number of Gaussians, which establishes a foundational structure that remains largely stable during optimization. This foundation serves as a ground layer upon which high-frequency details are subsequently refined. This hierarchical approach proves particularly effective in preventing "penny-round-tile" artifacts, which typically emerge when all Gaussians converge to uniformly small sizes and locate side-by-side on the image to create a visually distracting pattern, as shown in the main paper.

III. Color Space

Our experimental results show that the choice of color space may impact both the visual quality and convergence characteristics of EigenGS. By visualizing the reconstruction process Figures VI to VII, we observe distinct differences between RGB and YCbCr color space implementations, particularly in handling PCA-based reconstruction artifacts.

The YCbCr color space offers natural advantages through its separation of luminance (Y: 16-235) and chrominance (Cb/Cr: 16-240) components. This design provides inherent margins for handling reconstruction values that may fall outside the typical 0-255 range during PCA-based processing. In contrast, when using RGB color space, all three channels are processed with equal weighting, and thus, the reconstruction is more susceptible to outliers, particularly in regions with extreme values.

A particularly notable advantage of YCbCr implementation is its superior convergence speed. As evident in the visualizations, YCbCr-based reconstruction shows significant improvements within just 10 iterations, achieving sharper results compared to RGB. This acceleration in convergence can be attributed to the reduced dimensionality of color information: YCbCr isolates intensity variations to the Y channel, with only two channels handling color information. This characteristic appears to enhance optimization stability by reducing inter-channel interference during the reconstruction process.

Figure I. Qualitative results on the CelebA dataset using ImageNet-trained EigenGS.

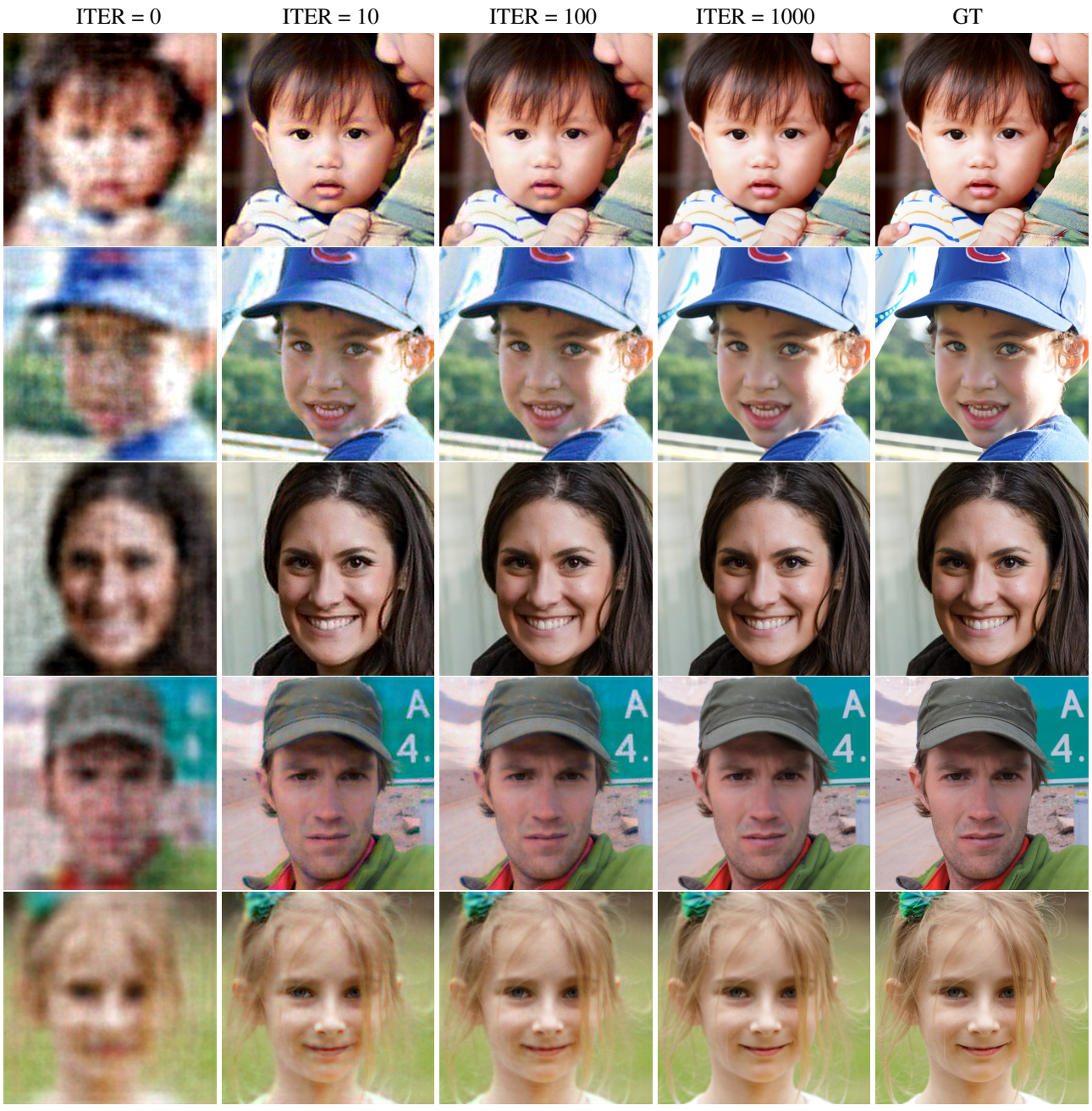

Figure II. Qualitative results on the FFHQ dataset using ImageNet-trained EigenGS.

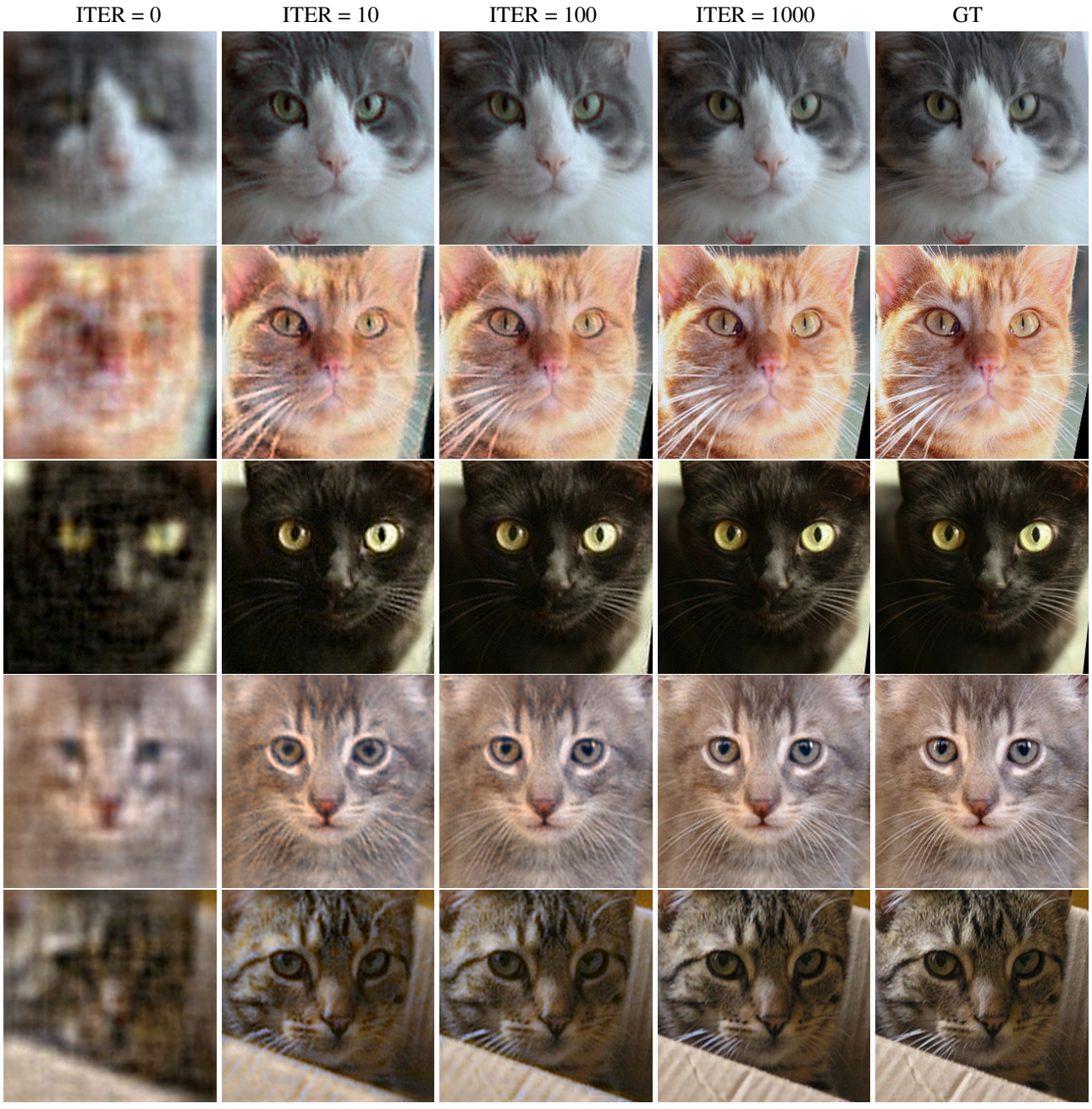

Figure III. Qualitative results on the Cats dataset using ImageNet-trained EigenGS.

Figure IV. Qualitative results on the Cars dataset using ImageNet-trained EigenGS.

CelebA		ITER=0	100	500	1000	5000	10000
GaussianImage	PSNR	-	10.3 ± 2.4	21.9 ± 1.2	30.0 ± 2.1	43.8 ± 2.9	45.2 ± 2.8
	SSIM	-	0.47 ± 0.05	0.82 ± 0.04	0.95 ± 0.02	0.99 ± 0.001	0.99 ± 0.001
	%	-	0	0	0	91	97
Ours (CelebA-trained EigenGS)	PSNR	29.9	37.8 ± 3.8	41.6 ± 3.9	43.6 ± 3.6	47.1 ± 3.1	48.1 ± 3.1
	SSIM	0.93	0.98 ± 0.02	0.99 ± 0.01	0.99 ± 0.003	0.99 ± 0.001	0.99 ± 0.001
	%	-	28	66	89	98	99
Ours (ImageNet-trained EigenGS)	PSNR	28.7	35.4 ± 3.4	39.6 ± 3.7	42.2 ± 3.6	46.3 ± 3.2	47.3 ± 3.1
	SSIM	0.91	0.96 ± 0.03	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.002	0.99 ± 0.001
	%	-	10	47	80	98	98

Table I. Quantitative comparison on the CelebA dataset with PSNR, SSIM, and percentage of images achieving PSNR larger than 40 dB.

FFHQ		ITER=0	100	500	1000	5000	10000
GaussianImage	PSNR	-	10.4 ± 1.7	21.8 ± 0.9	29.4 ± 1.6	39.2 ± 1.9	40.1 ± 1.9
	SSIM	-	0.41 ± 0.05	0.77 ± 0.05	0.94 ± 0.03	0.99 ± 0.001	0.99 ± 0.001
	%	-	0	0	0	98	99
Ours (FFHQ-trained EigenGS)	PSNR	28.0	34.4 ± 2.4	36.4 ± 2.6	37.5 ± 2.6	40.7 ± 2.5	41.8 ± 2.4
	SSIM	0.87	0.95 ± 0.02	0.98 ± 0.01	0.99 ± 0.01	0.99 ± 0.003	0.99 ± 0.002
	%	-	41	76	83	98	99
Ours (ImageNet-trained EigenGS)	PSNR	27.2	34.1 ± 2.5	36.1 ± 2.6	37.2 ± 2.7	40.5 ± 2.6	41.6 ± 2.4
	SSIM	0.84	0.95 ± 0.02	0.97 ± 0.01	0.98 ± 0.01	0.99 ± 0.003	0.99 ± 0.002
	%	-	39	71	79	97	99

Table II. Quantitative comparison on the FFHQ dataset with PSNR, SSIM, and percentage of images achieving PSNR larger than 35 dB.

Cats		ITER=0	100	500	1000	5000	10000
GaussianImage	PSNR	-	11.2 ± 2.2	22.2 ± 1.3	30.4 ± 2.2	42.4 ± 4.9	43.2 ± 4.9
	SSIM	-	0.47 ± 0.10	0.82 ± 0.07	0.96 ± 0.02	0.99 ± 0.02	0.99 ± 0.02
	%	-	0	0	0	69	74
Ours (Cats-trained EigenGS)	PSNR	30.6	38.1 ± 4.8	41.3 ± 5.2	42.8 ± 5.1	45.3 ± 4.5	46.1 ± 4.5
	SSIM	0.92	0.97 ± 0.02	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.001	0.99 ± 0.001
	%	-	35	57	70	89	90
Ours (ImageNet-trained EigenGS)	PSNR	29.6	37.6 ± 5.2	41.1 ± 5.6	42.9 ± 5.4	45.9 ± 4.7	46.6 ± 4.6
	SSIM	0.90	0.96 ± 0.04	0.99 ± 0.02	0.99 ± 0.01	0.99 ± 0.002	0.99 ± 0.002
	%	-	35	59	67	88	90

Table III. Quantitative comparison on the Cats dataset with PSNR, SSIM, and percentage of images achieving PSNR larger than 40 dB.

Cars		ITER=0	100	500	1000	5000	10000
GaussianImage	PSNR	-	12.7 ± 0.8	23.9 ± 0.7	32.9 ± 1.6	41.8 ± 2.3	42.9 ± 2.2
	SSIM	-	0.57 ± 0.05	0.03 ± 0.01	0.98 ± 0.001	0.99 ± 0.004	0.99 ± 0.001
	%	-	0	0	10	100	100
Ours (Cars-trained EigenGS)	PSNR	24.4	31.6 ± 2.5	34.1 ± 3.1	36.6 ± 3.6	43.3 ± 4.2	44.7 ± 4.1
	SSIM	0.85	0.95 ± 0.02	0.98 ± 0.01	0.99 ± 0.01	0.99 ± 0.001	0.99 ± 0.001
	%	-	11	32	67	99	100
Ours (ImageNet-trained EigenGS)	PSNR	23.9	31.5 ± 2.5	33.9 ± 2.9	36.3 ± 3.2	43.2 ± 4.1	44.5 ± 4.1
	SSIM	0.84	0.95 ± 0.03	0.98 ± 0.01	0.99 ± 0.01	0.99 ± 0.02	0.99 ± 0.01
	%	-	8	38	62	98	99

Table IV. Quantitative comparison on the Cars dataset with PSNR, SSIM, and percentage of images achieving PSNR larger than 35 dB.

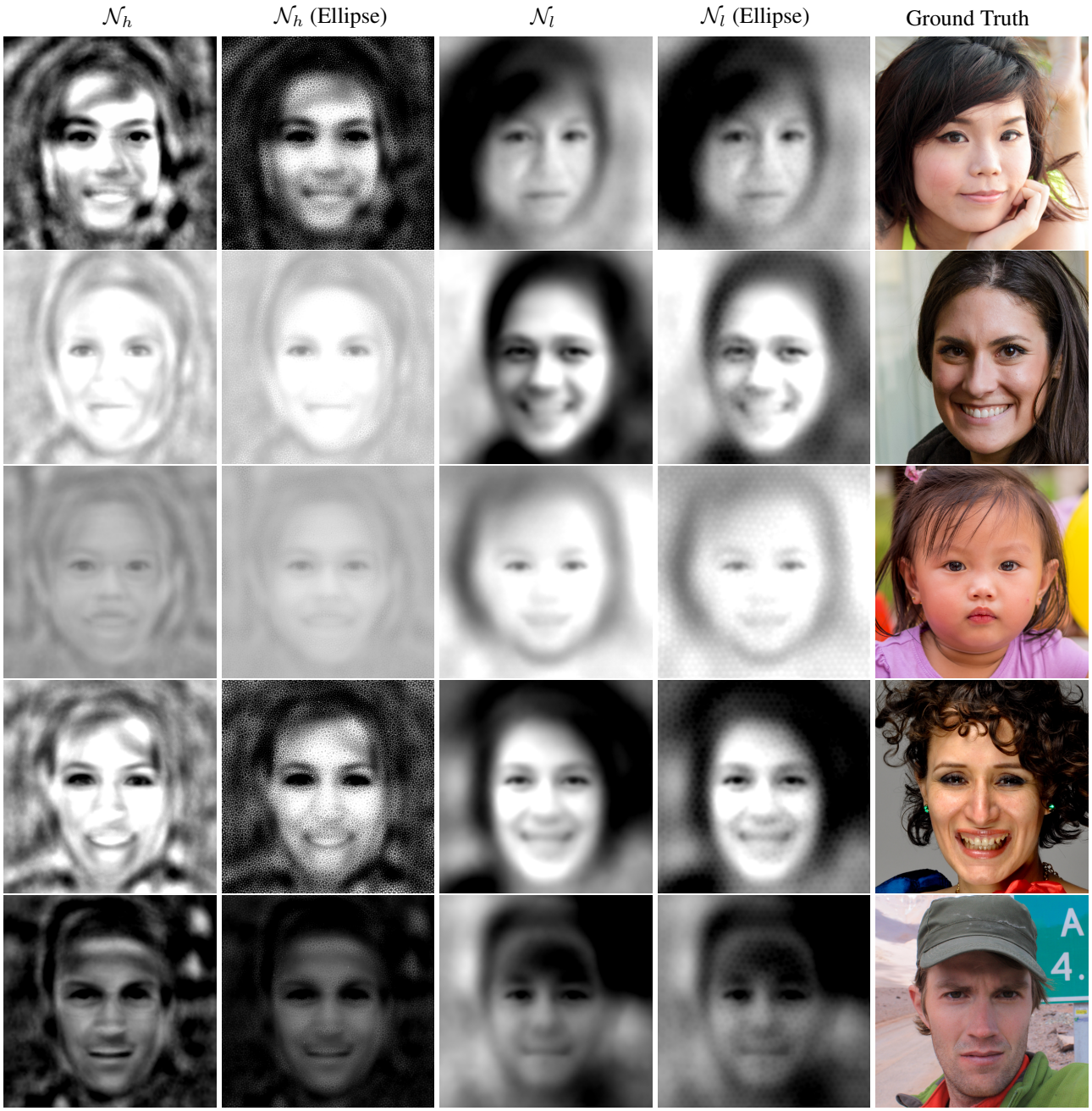

Figure V. Spatial frequency separation using normalized Y-channel rendering, where darker and brighter regions represent negative and positive Gaussian clusters respectively. From left to right: $\mathcal{N}_h$ shows reconstruction from only high-frequency Gaussians; $\mathcal{N}_h$ (Ellipse) visualizes the smaller spatial coverage of detail-oriented Gaussians; $\mathcal{N}_l$ shows reconstruction using only low-frequency Gaussians, $\mathcal{N}_l$ (Ellipse) displays the same Gaussians with reduced scale to highlight their spatial extent, and GT shows the ground truth image.

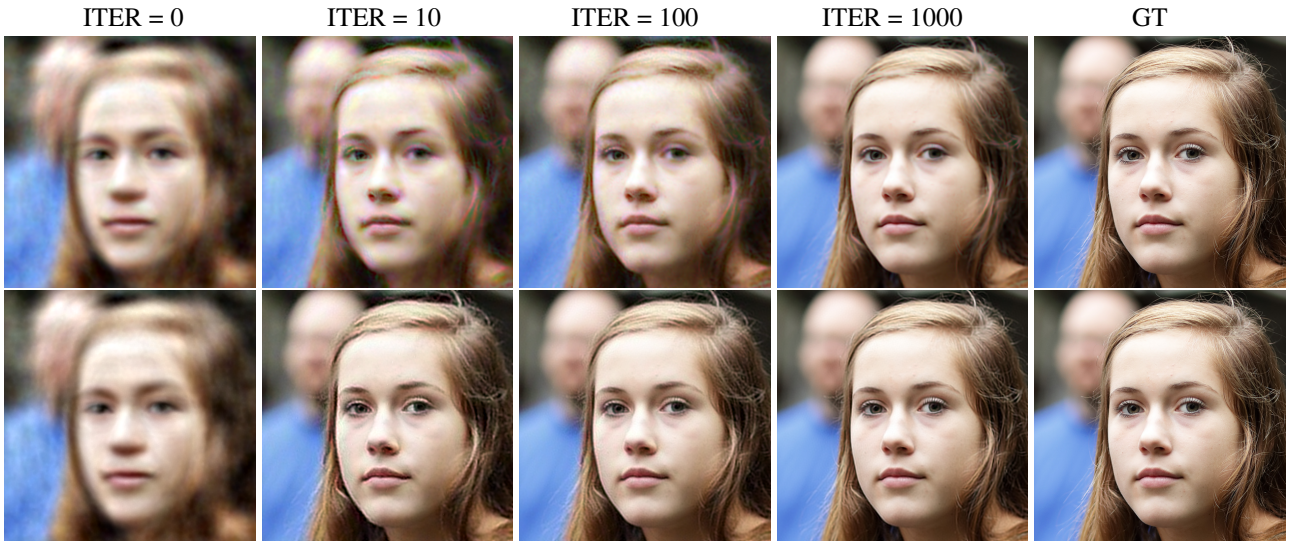

Figure VI. Color space comparison on FFHQ dataset. Top row: RGB color space reconstruction shows noticeable color shifts particularly at early iteration, *e.g.*, the red and blue fringe-like artifacts around the hair and eyes. Bottom row: YCbCr color space reconstruction shows superior early convergence with sharp detail preservation, even for the hair regions, exhibiting minimal color artifacts throughout the optimization process.

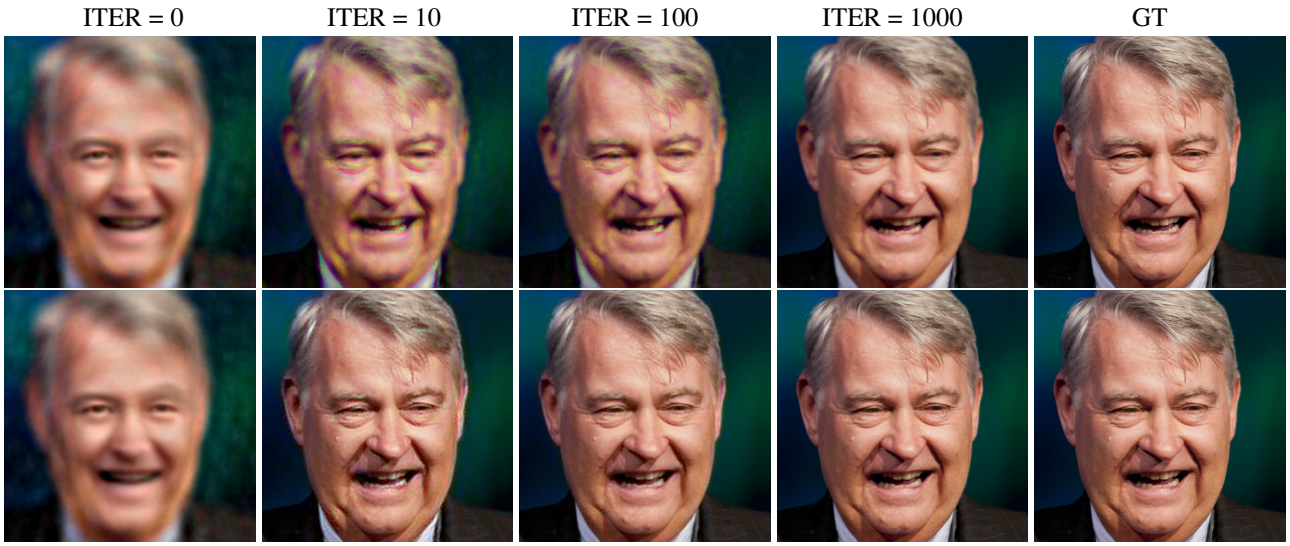

Figure VII. Another example of color space comparison on FFHQ dataset. Top: RGB color space reconstruction displays visible color bleeding during intermediate iterations. Bottom: YCbCr color space reconstruction maintains better color fidelity and stability, further supporting the advantages of separated luminance-chrominance processing in our method.

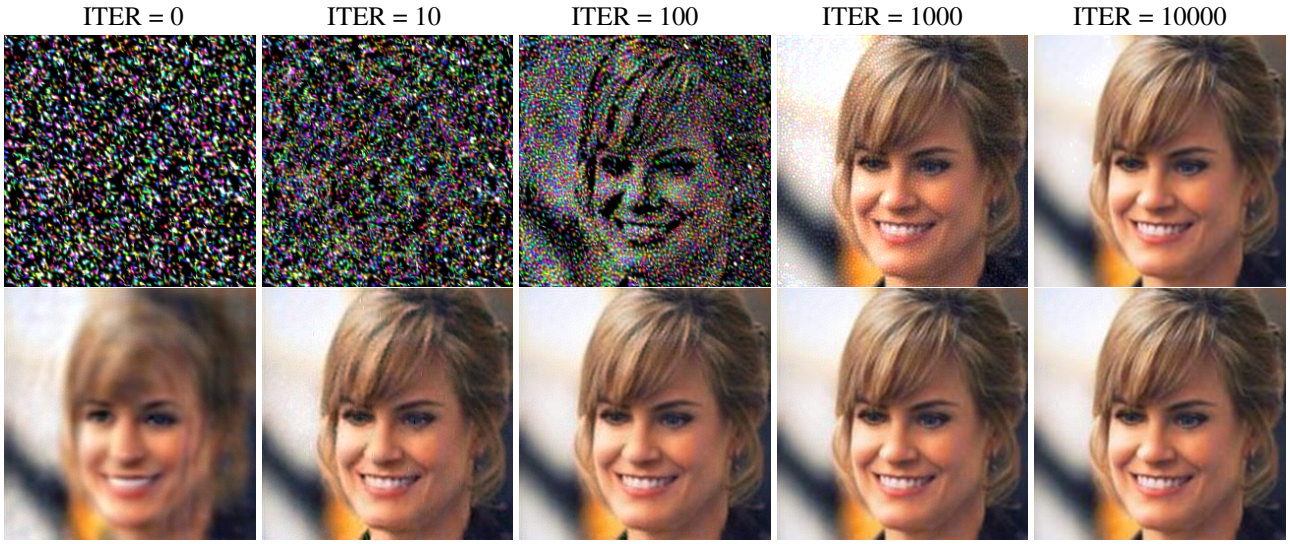

Figure VIII. Visual comparison between GaussianImage (top) and our method (bottom).