

# Gaussian Billboards: Expressive 2D Gaussian Splatting with Textures

Sebastian Weiss Derek Bradley  
DisneyResearch|Studios

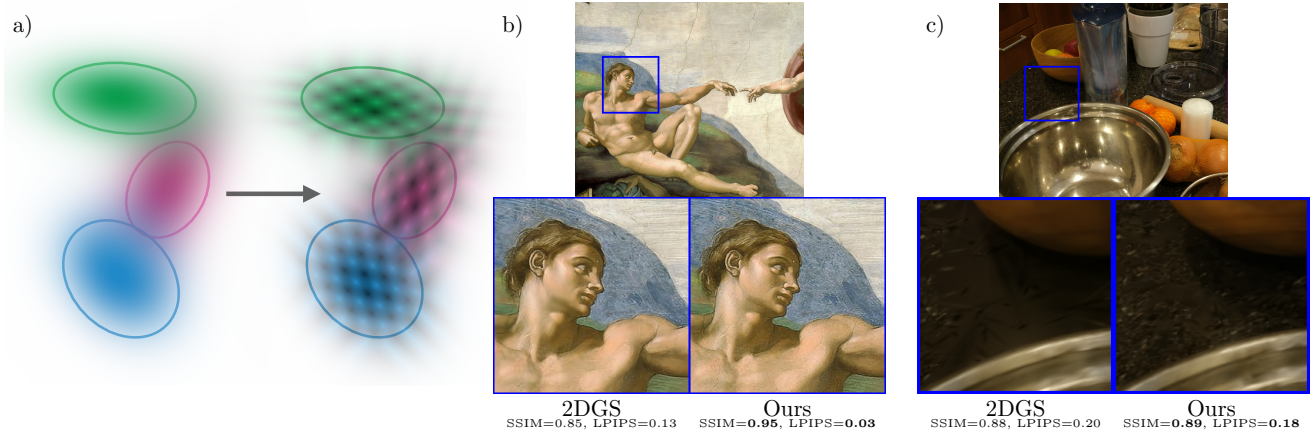


Figure 1. We propose an enhancement for 2D Gaussian Splatting that (a) replaces the per-primitive solid color with a small per-primitive texture. This allows the 2DGS to represent more details (b) when the number of primitives is fixed to, *e.g.*, 100000. Even with the standard densification strategy and identical hyperparameters, it improves the quality of 3D scene reconstruction (c), shown on a test view.

## Abstract

*Gaussian Splatting has recently emerged as the go-to representation for reconstructing and rendering 3D scenes. The transition from 3D to 2D Gaussian primitives has further improved multi-view consistency and surface reconstruction accuracy. In this work we highlight the similarity between 2D Gaussian Splatting (2DGS) and billboards from traditional computer graphics. Both use flat semi-transparent 2D geometry that is positioned, oriented and scaled in 3D space. However 2DGS uses a solid color per splat and an opacity modulated by a Gaussian distribution, where billboards are more expressive, modulating the color with a  $uv$ -parameterized texture. We propose to unify these concepts by presenting Gaussian Billboards, a modification of 2DGS to add spatially-varying color achieved using per-splat texture interpolation. The result is a mixture of the two representations, which benefits from both the robust scene optimization power of 2DGS and the expressiveness of texture mapping. We show that our method can improve the sharpness and quality of the scene representation in a wide range of qualitative and quantitative evaluations compared to the original 2DGS implementation.*

## 1. Introduction

Neural scene representations have become a standard solution to represent arbitrary 3D scenes from a collection of 2D images, allowing for novel view synthesis, relighting, animation playback or animation retargeting, just to name a few applications. Recently, 3D Gaussian Splatting (3DGS) [9] has established itself as the method of choice, outperforming previous representations like NeRF [14] or Mixture of Volumetric Primitives [13]. 3DGS’s core idea is to represent the scene explicitly as a collection of volumetric primitives, positioned and oriented in 3D, and then rasterized and rendered as 2D splats on the screen. To make this representation differentiable, the opacity of this volume is modeled via a 3D Gaussian distribution.

Since the introduction of 3DGS, many improvements to its rendering and representation quality have been presented. In particular, 2D Gaussian Splatting (2DGS) [6] replaces the volumetric primitives with oriented 2D primitives (disks). Still, representing highly detailed textured areas with potentially low geometric details, *i.e.* an image on a wall, remains challenging.

Looking back at traditional Computer Graphics, bill-

boards have been a common staple for rendering complicated, semi-transparent features like leaves, grass, or particle effects in general. We want to highlight here the similarities between 2DGS and billboards: in 2DGS, primitives are oriented 2D disks with a position, orientation and size in 3D, a constant color and opacity defined by a Gaussian distribution. Billboards are typically 2D squares with a position, orientation and size in 3D and a spatially-varying color and opacity using texture mapping.

To combine the best of both representations, we present *Gaussian Billboards*, an extension to 2DGS in which the color of the splat is spatially modulated using a small per-primitive texture, see Fig. 1a. We show that replacing the per-primitive solid color with a per-primitive texture improves the reconstruction capability when a fixed number of primitives is employed (see Fig. 1b), and also when the number of primitives is changed over the course of the optimization with densification and pruning strategies (see Fig. 1c). In short, our contributions are:

1. An extension to 2DGS with a spatially-varying color defined by a per-primitive texture,
2. A detailed ablation of the introduced hyperparameters,
3. An evaluation of 2D and 3D scene representations showing the improved capabilities of Gaussian Billboards in representing texture details.

It is worth mentioning that, while these contributions are novel, concurrent work by Rong *et al.* [16] named *GStex* also explores similar ideas to enhance 2DGS with spatially varying textures.

## 2. Related Work

We envision the proposed *Gaussian Billboards* as a method that is orthogonal to other techniques for improving the rendering quality of 3DGS and 2DGS.

Yu *et al.* [24] and Yan *et al.* [20] show how to overcome aliasing errors during rasterization if Gaussian splats become very small from far-away cameras. Improvements to the densification, splitting and training heuristics were presented by Zhang *et al.* [25], Ye *et al.* [22] and Fang&Wang [3]. Especially in the context of 3DGS, the volumetric splats are approximated as flat disks and then sorted by the depth of the center position for alpha blending. This approximation can lead to popping artifacts due to inconsistent sorting between views. Radl [15] show how to remedy these rendering errors using per-pixel sorting. If the input views are of particular low resolution, Feng *et al.* [4] show how to utilize sub-pixel constraints and 2D super-resolution networks to recover sharp details. Similarly, Seiskari *et al.* [17] show how to explicitly compensate for motion blur and rolling shutters if the input images come from a video sequence. Alternatives to the usually used Adam optimizer have been explored by Höllein *et al.* [7].

More closely related to our approach is 3D Half-

Gaussian Splatting by Li *et al.* [11] where the 3D volumetric primitive is split into two halves and a separate color and opacity can be assigned to each half. As one of the few works explicitly exploiting the *uv*-parametrization of 2DGS, Yu *et al.* [23] show how to modulate the Gaussian distribution of opacity with additional Hermite splines for sharper opacity borders. Texture-GS by Xu *et al.* [19] operates on 3DGS and utilizes a small MLP to perform a global UV-unwrapping that maps the 3D splat position to a *uv* position of a global texture atlas. Due to the smooth mapping, this method facilitates simple texture edits in texture space, but due to its global nature it is limited to simple, isolated objects.

Concurrent work called *GStex* by Rong *et al.* [16] also explores per-primitive color textures. Differences between *GStex* and our proposed method include that we assign a small texture to each primitive right from the start that is optimized from the beginning, whereas *GStex* first trains with per-primitive solid colors and only adds textures in later epochs with potentially different resolutions per primitive. Furthermore, we conduct a detailed ablation of the spatial extent of the texture within the primitive and a different optimum compared to the value reported by *GStex*.

## 3. Method

Before we present the proposed *Gaussian Billboards* method in Sec. 3.2, we first revisit the foundations of 3D and 2D Gaussian Splatting.

### 3.1. Gaussian Splatting Fundamentals

In the seminal work by Kerbel *et al.* [9], 3D Gaussian Splatting (3DGS) was introduced as a faster, higher quality 3D scene representation, compared to, *e.g.*, Neural Radiance Fields (NeRFs) [14]. Instead of representing the scene as a continuous, colored volume stored in an implicit neural network, 3DGS represents the scene as a collection of 3D Gaussian primitives that can be efficiently rasterized (*splatted*) to a 2D screen [27, 28]. In 3DGS, a single primitive  $k$  is defined by its location in 3D  $\mathbf{p}_k$ , its covariance matrix  $\Sigma_k$  parameterized via a scale  $\mathbf{s}_k$  and rotation quaternion  $\mathbf{q}_k$ , an RGB color  $\mathbf{c}_k$ , and an opacity  $\alpha_k$ . This representation can be viewed as an oriented and colored ellipsoid. To make this representation differentiable and optimizable, 3DGS turns these ellipsoids into 3D Gaussian primitives by multiplying the opacity with a 3D Gaussian distribution [9]:

$$\mathcal{G}(\mathbf{p}) = \exp\left(-\frac{1}{2}(\mathbf{p} - \mathbf{p}_k)^T \Sigma^{-1}(\mathbf{p} - \mathbf{p}_k)\right). \quad (1)$$

Using the techniques by Zwicker *et al.* [27], the perspective projection of a 3D Gaussian distribution can be analytically approximated giving rise to a 2D Gaussian distribution  $\mathcal{G}^{2D}$  in screen space. Then, all splats intersecting the camera ray



of a given pixel  $\mathbf{x}$  are blended together using front-to-back alpha compositing.

$$\mathbf{c}(\mathbf{x}) = \sum_{k=1}^K \mathbf{c}_k \alpha_k \mathcal{G}_k^{2D}(\mathbf{x}) \prod_{j=1}^{k-1} (1 - \alpha_j \mathcal{G}_j^{2D}(\mathbf{x})). \quad (2)$$

2D Gaussian Splatting [6] simplifies the above method by representing the scene as a collection of oriented, flat 2D primitives for better view consistency and improved surface reconstruction. Here, the geometry of a 2D Gaussian primitive is defined by a center position  $\mathbf{p}_k \in \mathbb{R}^3$ , two orthogonal tangent vectors  $\mathbf{t}_{u,k}, \mathbf{t}_{v,k} \in \mathbb{R}^3$  (derived from a rotation quaternion) and a 2D scaling vector  $S = (s_u, s_v) \in \mathbb{R}^2$ . This defines a local tangent plane parameterized by  $\mathbf{u} = (u, v)$ :

$$P(\mathbf{u}) = \mathbf{p}_k + s_u \mathbf{t}_{u,k} u + s_v \mathbf{t}_{v,k} v. \quad (3)$$

The 2D Gaussian distribution for modulating the opacity is now equally defined in this  $uv$ -space using

$$\mathcal{G}(\mathbf{u}) = \exp\left(-\frac{1}{2}(u^2 + v^2)\right). \quad (4)$$

### 3.2. Gaussian Billboards

In the original 2DGS implementation [6], each 2D primitive is rendered with a constant color  $\mathbf{c}_k$  across the entire 2D plane it spans. This ties both geometry and textural information together in the same representation, limited by the same quantity, *i.e.* the number of splats.

We observe, however, that in most scenes, the texture information is higher than geometry information. Therefore, we take inspiration from textured billboards of traditional Computer Graphics and introduce a spatially-varying color representation by bi-linearly interpolating a color grid based on the  $uv$ -parametrization of the 2D primitive, see Eq. (3). This enables modelling textures better with the same or fewer number primitives. Let  $N \in \mathbb{N}$  the resolution of the per-primitive color grid. Since the Gaussian primitive has infinite support, we have to define a region in the  $uv$ -plane in which the texture is defined. Let  $\sigma \in \mathbb{R}^+$  denote said spatial extent, then the texture grid is contained within  $[-\sigma, +\sigma]^2$  in the  $uv$ -plane. The pixel color of the current primitive is then given by

$$\mathbf{c}(u, v)_k = \text{bilinear}(C_k, \frac{u}{\sigma}, \frac{v}{\sigma}), \quad (5)$$

where  $\text{bilinear}(C, x, y)$  represents bilinear interpolation into a grid  $C$  of colors with shape  $N \times N$ , evaluated at normalized coordinates  $x, y \in [-1, +1]$ . Since the color grid is defined in the  $uv$ -parametrization, it naturally scales with  $S$  and rotates with  $\mathbf{t}_{u,k}, \mathbf{t}_{v,k}$ , see Eq. (3).

A visualization of the proposed spatial color can be seen in Fig. 2 for  $N = 2, 4, 8$  with a red-green-blue checkerboard texture. The darker area indicates the area of  $[-\sigma, +\sigma]^2$ .

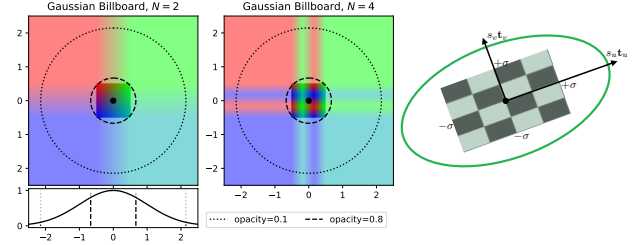


Figure 2. The color grid of Gaussian Billboards spans the  $uv$ -parametrization of the 2D primitive in the range of  $[-\sigma, +\sigma]$  with  $\sigma = 0.5$ . This way, the border of the grid where the interpolation starts being clamped roughly corresponds to the isoline of 0.8 opacity. On the right you can see how the texture within a primitive is rotated and scaled with  $s_u, s_v, \mathbf{t}_u, \mathbf{t}_v$ .

However, since the Gaussian primitive has infinite support, we border-clamp the  $uv$ -values in the bilinear interpolation function. The result of this clamping is visualized in the brighter shaded region. An intuitive choice for  $\sigma$  is to pick it such that the texture grid covers the majority of the visible area of a primitive to minimize the area where this clamping occurs, *e.g.*, at the isoline of 0.1 opacity (dotted line) with  $\sigma = 2.0$ . We found, however, that the value for  $\sigma$  has a significant impact on how much the proposed Gaussian Billboards can improve over traditional 2DGS. An ablation of this value can be found in Sec. 4.1, with the optimal value being  $\sigma = 0.5$ , or the full color grid residing within the isoline of 0.8 opacity (dotted line), shown in Fig. 2. For all of our results, if not otherwise mentioned, we use a value of  $\sigma = 0.5$ .

### 3.3. Training

For training the scene representation, we follow the standard practices of 3DGS and 2DGS. We initialize the primitives using structure-from-motion if available and add slight color variations to the color grid. We use Adam [10] as the optimizer and train only with a photometric loss. Since we envision *Gaussian Billboards* as an orthogonal method to other improvement techniques, we opted to keep all other training hyperparameters at the default values used in traditional 2DGS to isolate the effects of introducing spatially-varying colors.

For optimal performance, we integrated Gaussian Billboards directly into the fused CUDA kernels of *gsplat* [21] with gradients being propagated to all parameters, including the color grid and the  $uv$  coordinate, thus also to the position, scale and orientation of the primitive.

To reduce global memory bandwidth in the forward and backward pass of the rasterization kernels, *gsplat* first loads and accumulates the per-primitive parameters and their gradients in shared memory for a batch of pixels and primi-

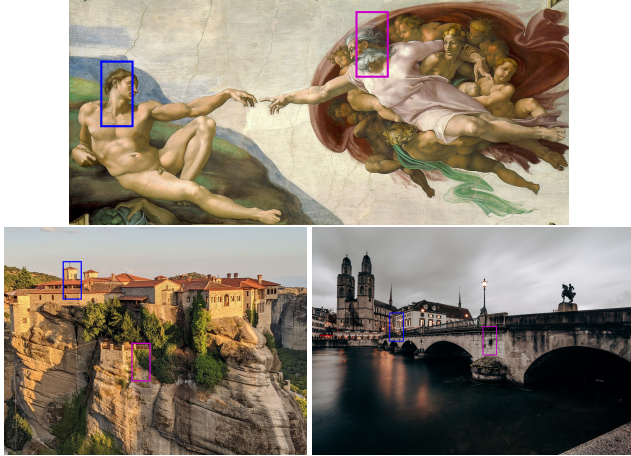


Figure 3. The three images used for image fitting tests and hyper-parameter ablations: *Creation of Adam*, *Meteora*, and *Zurich*. The marked areas indicate the crops visualized in Fig. 6.

tives called a tile, before writing them out to global memory. Shared memory is bound by hardware to 32kB to 100kB depending on the GPU architecture and sets an upper bound of the per-primitive texture resolution. For  $N = 2$  and  $N = 4$ , the additional parameters fit into shared memory without changes. For  $N = 8$ , however, the required shared memory exceeds the available memory in hardware (tested on an A6000 and RTX3090), and the tile size must be reduced from  $16 \times 16$  to  $8 \times 8$ , drastically increasing the rendering and training times (see Sec. 4.1 and Fig. 6). This also makes larger resolutions of, e.g.,  $N = 16$  infeasible.

## 4. Results

We first evaluate the proposed method and design decisions on the task of fitting a 2D image, see Sec. 4.1. Then we qualitatively and quantitatively evaluate Gaussian Billboards on a range of 3D scene reconstruction tasks using the NeRF360 dataset [1] and human faces, see Sec. 4.2.

### 4.1. Ablations on Image Overfitting

For overfitting a single image, we define one camera focussed on the image in +Z direction and restrict the rotation of the primitives to only rotate around the Z-axis. This way, the plane of the 2D primitives aligns with the image plane. The primitives are initialized with random positions in the image plane, random orientation, size, and color. We perform these tests on three images (see Fig. 3), *Creation of Adam* by Michelangelo of size  $1920 \times 871^*$ , a photograph of the *Meteora* monastery in Greece with resolution  $2016 \times 1512$ , and a photograph of the Münsterbrücke in *Zurich*, Switzerland, with resolution  $1833 \times 2000$ . All results shown are trained for 20000 epochs with an MSE loss.

\**Creation of Adam* is taken from Wikicommons, public domain.

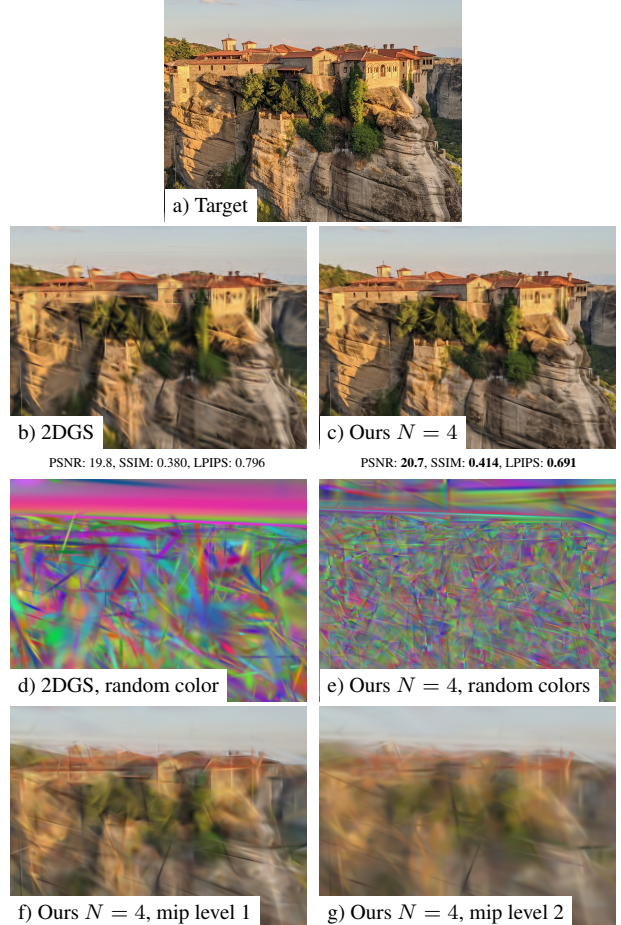


Figure 4. Overfitting a single image (a) with a fixed number of 1000 splats. For that limited number of splats, Ours with  $N = 4$  (c) achieves better quality than Traditional 2DGS (b) in terms of quantitative statistics and image sharpness. If we assign random colors, the individual splats and the bilinear interpolation of the color grid become visible (d,e). To verify that the optimization utilizes the additional color information, we downsample the color grid to a resolution of  $2 \times 2$  (f) and a single color (g).

In a first test, we fix the number of splats to 1000 and compare traditional 2DGS with our proposed method, see Fig. 4, and  $N = 4$ . For this limited number of splats, our method achieves visually a sharper and more texture-rich image (c vs b) and improved quantitative statistics. To visualize the individual splats, we assign random colors to each splat (d and e). Finally, to verify that the optimization utilizes the additional color parameters, we render the image with the color grid downsampled to a size of  $2 \times 2$  (f) or a single color (g) using a box filter. As one can see, the image quality drastically decreases, highlighting that the color grid is now crucial for the image quality.

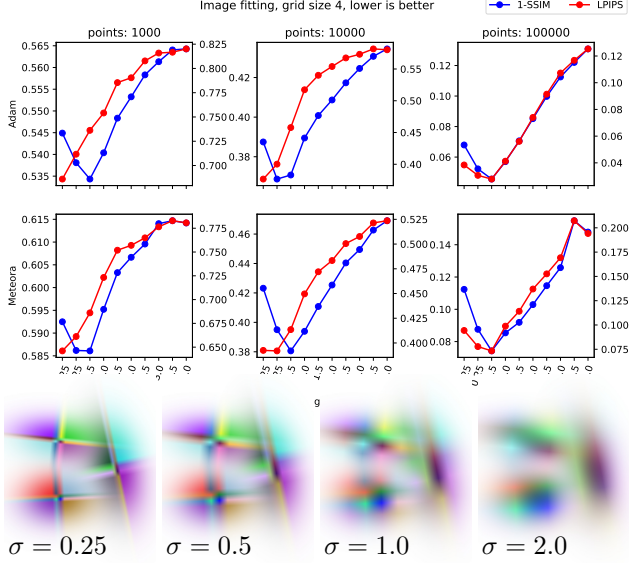


Figure 5. Ablation of the spatial extent of the grid in the  $uv$ -plane,  $\sigma$ , for overfitting of *Creation of Adam* and *Meteora* (see Fig. 3) with  $N = 4$ . With a lower value for  $\sigma$ , the texture information is concentrated in the area of the primitive with the highest opacity. As can be seen in the plot, the optimal value for  $\sigma$  is achieved at around 0.5. For a visualization of what different values for  $\sigma$  mean, the bottom row shows three Gaussian primitives with a random texture rendered with four values for  $\sigma$ .

**Ablation of the Grid Extent  $\sigma$ .** Next, we evaluate the influence of the hyperparameter  $\sigma$ , the spatial extent of the grid in the  $uv$ -parametrization, on an image fitting task with a fixed number of primitives (see Fig. 5). A lower value for  $\sigma$  indicates that the texture details are concentrated in the center area of the Gaussian primitive where the opacity is highest. As can be seen in the statistics of Fig. 5, across two different test images and different number of primitives, the reconstruction quality measured in SSIM [18] and LPIPS [26] scores, improves with lower values for  $\sigma$ . This reaches an optimal point of around  $\sigma = 0.5$ . We argue that for even smaller values, the texture essentially collapses to a binary quadrant image and cannot make use of the full resolution of a  $4 \times 4$  grid in this case anymore.

**Ablation of the Texture Resolution  $N$ .** With a fixed value of  $\sigma = 0.5$  we now turn to analyzing the texture resolution  $N$  and its impact on the reconstruction quality and training speed compared to traditional 2DGS without spatially varying colors. We perform this experiment on all three test images with a fixed number of primitives of 10,000 and 100,000. See Fig. 6 for the results.

Qualitatively, the difference can be seen best for 10,000 primitives (upper row of Fig. 6 per image). With  $N = 1$  (traditional 2DGS), many details are either blurred out in

*Creation of Adam* and *Zurich* or approximated with elongated splats in *Meteora*. From  $N = 2$  to  $N = 4$ , the reconstructed image significantly improves. Texture detail of the paint becomes visible in *Creation of Adam* and details in the plants and bricks become visible in *Meteora* and *Zurich*. Continuing on to  $N = 8$  still shows improvements, but less noticeable.

Quantitatively, increasing the texture resolution strictly improves the PSNR, SSIM, and LPIPS scores, evaluated over the entire image. For *Creation of Adam* and 100,000 primitives, the SSIM score increases, for example, from 0.849 for  $N = 1$  to 0.988 for  $N = 8$ , matching the input image almost pixel by pixel.

Enhancing the Gaussian primitives with spatially-varying textures, however, does come with some overhead in computation, but especially with an overhead of additional memory transfer and reduced tile size (see Sec. 3.3) due to the additional parameters that must be considered in the splatting kernel. We, therefore, look at the timings shown in Fig. 6 now, all tests trained for 20000 epochs on an A6000 GPU. For  $N = 2$  and  $N = 4$ , the training time usually increases between 10% and 30% compared to traditional 2DGS. However, for  $N = 8$ , the tile size has to be reduced and the training time increases dramatically for all test cases, *e.g.*, by a factor of almost  $4\times$  for *Meteora*. Despite the improved quality with  $N = 8$ , we found the increase in computation time to not be worth it. We, therefore, recommend  $N = 4$  and use that value for all following tests.

## 4.2. Evaluation on 3D Scene Reconstruction

In a 3D scene reconstruction task, the surface color can change depending on the view direction due to specular effects. We can model this as Spherical Harmonics as traditionally done in 3DGS and 2DGS [6, 9] and directly extend this to a spatially varying color grid by storing  $M$  spherical harmonics coefficients for each texel instead of a single RGB color. This, however, quickly becomes expensive in the number of parameters for high spherical harmonics degrees. Alternatively, we can treat the color grid as the albedo of the surface and then model the view-dependent appearance with explicit BSDF models [8]. In this work, however, we opt to perform a diffuse-specular split. We use the proposed color grid as the view-independent Lambertian part and model the residual specular components as an additive term using a single set of spherical harmonics coefficients – not spatially varying – per primitive.

Results of the 3D scene reconstruction task on the NeRF360 dataset [1], downsampled by a factor of 2 to a resolution of approximately  $2500 \times 1650$ , can be found in Fig. 8. We train for 30000 epochs with default densification and pruning enabled, initialized with SfM points, and a loss of  $MSE + 0.2 * SSIM$ . Quantitatively, our method





Figure 6. Evaluation of the texture resolution  $N$  for image fitting with a fixed number of primitives, per scene with 10,000 primitives (top row) and 100,000 primitives (bottom row). With an increase in texture resolution, the Gaussian Billboards consistently lead to better fittings.

( $N = 4, \sigma = 0.5$ ) consistently outperforms traditional 2DGS in terms of LPIPS statistics on validation images. For SSIM, *Ours* outperforms 2DGS in 5 out of 7 cases with the two outliers being very close between both methods.

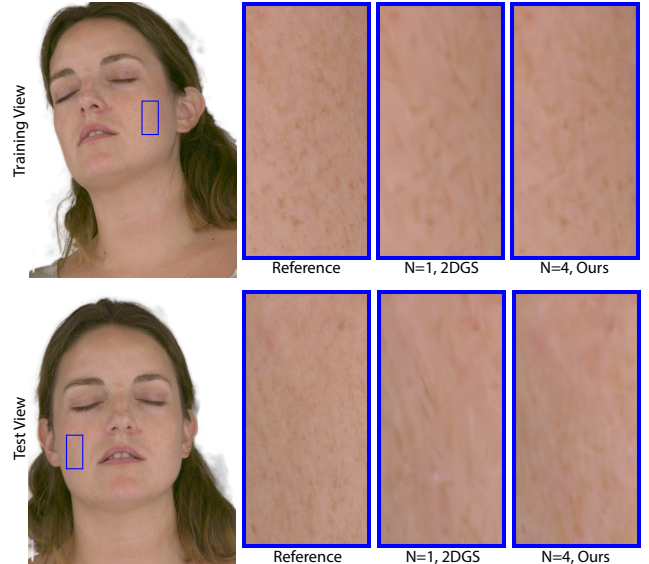
Qualitatively, we observe that *Gaussian Billboards* improves the visual quality especially in areas where the geometry is fairly flat but shows a lot of texture variation. One case can be seen in the counter scene (third row) where the marble texture of the counter is represented better in *Ours* (blue inset). Another example of the same behavior can be seen in the shovel of the toy bulldozer. Even for areas of high spatial variation, *e.g.*, the vegetation in the bicycle scene of row one, *Ours* visually reconstructs more detail.

**Face Reconstruction.** For a last test we evaluate *Gaussian Billboards* on 3D face reconstruction from a multi-view camera setup with 14 training cameras arranged in a hemisphere around the front of the subject and one central test camera not used for training. We use camera calibration with fiducial markers [5] and a 3D face scan [2] to obtain a point cloud for initialization. The background of the camera images with resolution  $2047 \times 1535$  was removed using Lin & Ryabtsev *et al.* [12]. All training parameters remain the same as described above.

The results can be found in Fig. 7. Even while quantitative metrics remain almost equal, with only the LPIPS score being noticeably improved from traditional 2DGS ( $N = 1$ ) to our method ( $N = 4$ ), differences are visible in the qualitative comparison. Our result preserves more details in both training and testing views while having a more uniform appearance. In the baseline with  $N = 1$ , artifacts produced by elongated splats are especially visible. Furthermore, with standard densification and pruning enabled, our method achieves better results with fewer primitives ( $\approx 90000 \rightarrow \approx 64000$ ) compared to the baseline. This also reflects in the time to render the test view, which decreases slightly ( $0.019s \rightarrow 0.015s$ ). The training time, however, increases slightly due to the additional memory transfer in the backward pass (see Sec. 3.3) from 80 minutes to 95 minutes for 30000 epochs.

## 5. Conclusion

In this work we presented an improvement over 2D Gaussian Splatting that replaces the per-primitive solid color with a small color texture that is jointly optimized together with all other 2DGS parameters. This increases the power to express texture details, leading to an improved image reconstruction for the same number of Gaussian primitives. Even if the number of primitives can be freely optimized due to densification, splitting, and pruning strategies, our proposed method improves the reconstruction and novel view synthesis over traditional 2DGS on 3D reconstruction tasks.



		PSNR	SSIM	LPIPS
Train	Baseline 2DGS	35.9	0.928	0.159
	$N = 4$ (ours)	35.2	0.924	<b>0.151</b>
Test	Baseline 2DGS	27.6	0.866	0.273
	$N = 4$ (ours)	27.5	0.865	<b>0.263</b>

Figure 7. 3D Reconstruction of a human face from 14 training cameras and one center test camera. Despite very similar quantitative metrics, our method produces sharper textures on the face, even with less Gaussian primitives, 90566 (Baseline 2DGS) to 64461 (Ours). Images best viewed zoomed-in.

We envision the proposed approach of *Gaussian Billboards* as an orthogonal approach to many related methods that also tackle the problem of how to enhance the representation quality of 3DGS or 2DGS itself. In future work, we want to explore the combination of these methods presented recently. Most promisingly, Yu *et al.* [23] introduce a spatial modulation of the opacity that could show good synergies with our proposed spatial color modulation.

Since the 2DGS primitive attributes are cached in shared memory during the forward and adjoint rasterization stage, the maximal supported color grid size is limited by the available shared memory. In our implementation that extends the CUDA kernels of *gsplat*, this leads to a maximal texture resolution of  $8 \times 8$  texels. In future, we plan to investigate how to lift this technical restriction.

## References

- [1] Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-NeRF 360: Unbounded anti-aliased neural radiance fields. *arXiv [cs.CV]*, 2021. 4, 5, 8
- [2] Thabo Beeler, Bernd Bickel, Paul Beardsley, Bob Sumner,



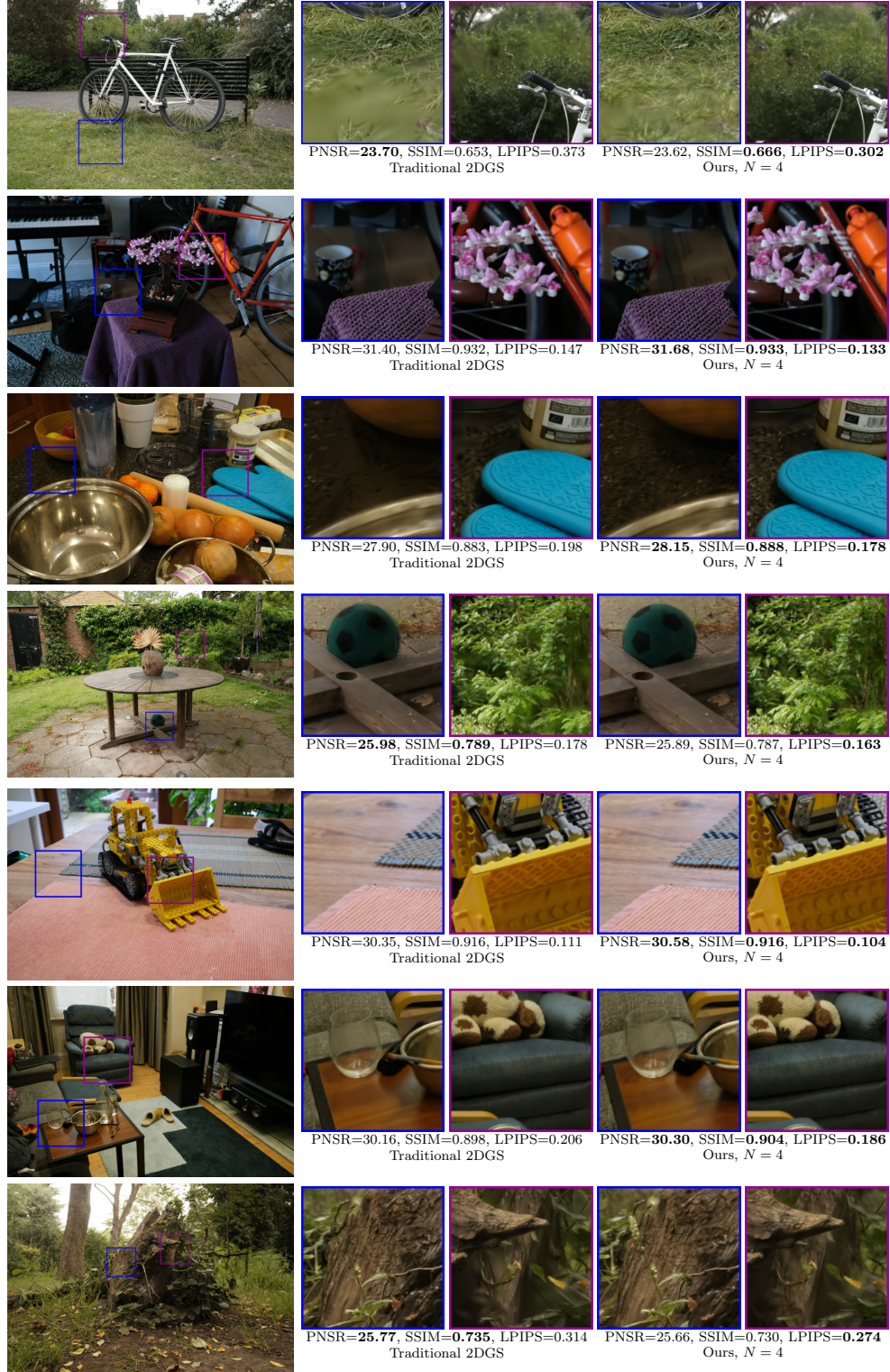


Figure 8. Evaluation of traditional 2DGS compared to Gaussian Billboards (*Ours*) on the NeRF360 validation images [1]. With the proposed method, areas that are seen from few images and areas that show texture variation with little 3D structure are reconstructed better, see *e.g.* the grass and leaf areas in row 1 or the inside of the shovel of the toy bulldozer in row 5. The reported PSNR, SSIM and LPIPS statistics are averaged over all validation images. The left column shows the reference image. Images best viewed zoomed-in.



- and Markus Gross. High-quality single-shot capture of facial geometry. *ACM Trans. Graph.*, 29(4):1, 2010. 7
- [3] Guangchi Fang and Bing Wang. Mini-splatting: Representing scenes with a constrained number of gaussians. *arXiv [cs.CV]*, 2024. 2
- [4] Xiang Feng, Yongbo He, Yubo Wang, Yan Yang, Zhenzhong Kuang, Yu Jun, Jianping Fan, and Jiajun Ding. SRGS: Super-resolution 3D gaussian splatting. *arXiv [cs.CV]*, 2024. 2
- [5] S Garrido-Jurado, R Muñoz-Salinas, F J Madrid-Cuevas, and M J Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognit.*, 47(6):2280–2292, 2014. 7
- [6] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 Conference Papers*, New York, NY, USA, 2024. Association for Computing Machinery. 1, 3, 5
- [7] Lukas Höllein, Aljaž Božič, Michael Zollhöfer, and Matthias Nießner. 3DGS-LM: Faster gaussian-splatting optimization with levenberg-marquardt. *arXiv [cs.CV]*, 2024. 2
- [8] Yingwenqi Jiang, Jiadong Tu, Yuan Liu, Xifeng Gao, Xiaoxiao Long, Wenping Wang, and Yuexin Ma. GaussianShader: 3D gaussian splatting with shading functions for reflective surfaces. *arXiv [cs.CV]*, 2023. 5
- [9] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkuehler, and George Drettakis. 3D gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):1–14, 2023. 1, 2, 5
- [10] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv [cs.LG]*, 2014. 3
- [11] Haolin Li, Jinyang Liu, Mario Sznaiier, and Octavia Camps. 3D-HGS: 3D half-gaussian splatting. *arXiv [cs.CV]*, 2024. 2
- [12] Shanchuan Lin, Andrey Ryabtsev, Soumyadip Sengupta, Brian L Curless, Steven M Seitz, and Ira Kemelmacher-Shlizerman. Real-time high-resolution background matting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8762–8771, 2021. 7
- [13] Stephen Lombardi, Tomas Simon, Gabriel Schwartz, Michael Zollhoefer, Yaser Sheikh, and Jason Saragih. Mixture of volumetric primitives for efficient neural rendering. *ACM Trans. Graph.*, 40(4):1–13, 2021. 1
- [14] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: representing scenes as neural radiance fields for view synthesis. *Commun. ACM*, 65(1):99–106, 2021. 1, 2
- [15] Lukas Radl, Michael Steiner, Mathias Parger, Alexander Weinrauch, Bernhard Kerbl, and Markus Steinberger. StopThePop: Sorted gaussian splatting for view-consistent real-time rendering. *arXiv [cs.GR]*, 2024. 2
- [16] Victor Rong, Jingxiang Chen, Sherwin Bahmani, Kiriakos N Kutulakos, and David B Lindell. GStex: Per-primitive texturing of 2D gaussian splatting for decoupled appearance and geometry modeling. *arXiv [cs.CV]*, 2024. 2
- [17] Otto Seiskari, Jerry Ylilammi, Valtteri Kaatrasalo, Pekka Rantalankila, Matias Turkulainen, Juho Kannala, Esa Rahtu, and Arno Solin. Gaussian splatting on the move: Blur and rolling shutter compensation for natural camera motion. *arXiv [cs.CV]*, 2024. 2
- [18] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.*, 13(4):600–612, 2004. 5
- [19] Tian-Xing Xu, Wenbo Hu, Yu-Kun Lai, Ying Shan, and Song-Hai Zhang. Texture-GS: Disentangling the geometry and texture for 3D gaussian splatting editing. *arXiv [cs.CV]*, 2024. 2
- [20] Zhiwen Yan, Weng Fei Low, Yu Chen, and Gim Hee Lee. Multi-scale 3D gaussian splatting for anti-aliased rendering. *arXiv [cs.CV]*, 2023. 2
- [21] Vickie Ye, Ruilong Li, Justin Kerr, Matias Turkulainen, Brent Yi, Zhuoyang Pan, Otto Seiskari, Jianbo Ye, Jeffrey Hu, Matthew Tancik, and Angjoo Kanazawa. gsplat: An open-source library for Gaussian splatting. *arXiv preprint arXiv:2409.06765*, 2024. 3, 1
- [22] Zongxin Ye, Wenyu Li, Sidun Liu, Peng Qiao, and Yong Dou. AbsGS: Recovering fine details for 3D gaussian splatting. *arXiv [cs.CV]*, 2024. 2
- [23] Ruihan Yu, Tianyu Huang, Jingwang Ling, and Feng Xu. 2DGH: 2D gaussian-hermite splatting for high-quality rendering and better geometry reconstruction. *arXiv [cs.CV]*, 2024. 2, 7
- [24] Zehao Yu, Anpei Chen, Binbin Huang, Torsten Sattler, and Andreas Geiger. Mip-splatting: Alias-free 3D gaussian splatting. *arXiv [cs.CV]*, 2023. 2
- [25] Jiahui Zhang, Fangneng Zhan, Muyu Xu, Shijian Lu, and Eric Xing. FreGS: 3D gaussian splatting with progressive frequency regularization. *arXiv [cs.CV]*, 2024. 2
- [26] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. *CVPR*, 2018. 5
- [27] M Zwicker, H Pfister, J van Baar, and M Gross. EWA volume splatting. In *Proceedings Visualization, 2001. VIS '01.*, pages 29–538. IEEE, 2001. 2
- [28] M Zwicker, H Pfister, J van Baar, and M Gross. EWA splatting. *IEEE Trans. Vis. Comput. Graph.*, 8(3):223–238, 2002. 2

# Gaussian Billboards: Expressive 2D Gaussian Splatting with Textures

## Supplementary Material

### 6. Algorithm Details

In Sec. 3.2 we introduce the basic algorithm of *Gaussian Billboards*, enhancing 2DGS with spatially-varying textures. Here we give additional implementation details on how to integrate this into the *gsplat* kernel [21].

In Algorithm 1 we show where the bilinear interpolation fits into the rasterization kernels of *gsplat* with the changes to the original kernel highlighted in **green**. The main changes in the forward algorithm is twofold. First, the parameter fetching is extended from a single color to a grid of color. Second, the bilinear interpolation kernel is invoked per primitive and pixel to obtain the per-pixel color given the current  $u, v$ . The central change to the backward algorithm is that the adjoint variables for the  $uv$ -parameterization,  $\hat{u}, \hat{v}$  are now also influenced by the adjoint of the color interpolation, instead of just the Gaussian opacity as in traditional 2DGS.

In Algorithm 2 we give the detailed pseudocode of the forward and backward implementation of bilinear interpolation. We manually derive the adjoint code – the derivatives of the output with respect to all inputs – from the bilinear interpolation code. This allows for direct integration of the proposed changes into the fused rasterization kernels of *gsplat*.

---

**Algorithm 1** Integration of bilinear color interpolation into the rasterization kernel (Simplified)

---

```
1: procedure RASTERIZE_FWD ▷ forward code
   Each CUDA block processes a range of 2DGS primitives.
   Each thread per block outputs one pixel.
2:   Compute the pixel location of the current thread  $p_x, p_y$ 
3:   Fetch all primitive parameters of the current range, with the color extended to a color grid
4:    $T = 1.0$  ▷ The transmittance
5:    $\mathbf{c}_{\text{acc}} = (0, 0, 0)$  ▷ Accumulated color
6:   for all primitives  $k$  in the sorted range do
7:     Intersect the primitive with the view ray  $\rightarrow u, v$ 
8:     Evaluate the color  $\mathbf{c} = \text{bilinear}(u, v, N, \sigma, C_k)$ 
9:     Compute the opacity  $\alpha = \alpha_k * \exp(-\frac{u^2+v^2}{2})$ 
10:     $\mathbf{c}_{\text{acc}} += \mathbf{c}\alpha T$  ▷ alpha-blending
11:     $T *= (1 - \alpha)$ 
12:   end for
13:   return  $\mathbf{c}_{\text{acc}}$ 
14: end procedure

15: procedure RASTERIZE_BWD ▷ backward code
   Each CUDA block processes a range of 2DGS primitives.
   Each thread per block outputs one pixel.
16:   Compute the pixel location of the current thread  $p_x, p_y$ 
17:   Fetch all primitive parameters of the current range, with the color extended to a color grid
18:   Fetch transmittance  $T$  after the last Gaussian
19:   Fetch adjoint of the output color  $\hat{\mathbf{c}}_{\text{acc}}$ 
20:    $\mathbf{c}_{\text{acc}} = (0, 0, 0)$  ▷ Accumulated color
21:   for all primitives  $k$  in the sorted range do
22:     Intersect the primitive with the view ray  $\rightarrow u, v$ 
23:     Compute the opacity  $\alpha = \alpha_k * \exp(-\frac{u^2+v^2}{2})$ 
24:     Evaluate the color  $\mathbf{c} = \text{bilinear}(u, v, N, \sigma, C_k)$ 
   Adjoint computations:
25:     Adjoint of color with visibility  $\hat{\mathbf{c}}_k = (\alpha * T)\hat{\mathbf{c}}_{\text{acc}}$ 
26:     Color-adjoint:  $\hat{C}, \hat{u}, \hat{v} = \text{adj-bilinear}(u, v, C_k, \hat{\mathbf{c}}_k)$ 
27:     Adjoint of alpha:  $\hat{\alpha}$ , requires  $\mathbf{c}$ 
28:     Adjoint of Gaussian opacity:  $\hat{u}, \hat{v} += \dots$ 
29:     Adjoint of ray intersection:  $\hat{\mathbf{p}}, \hat{\mathbf{t}}_u, \hat{\mathbf{t}}_v, \hat{s}_u, \hat{s}_v$ 
30:     Accumulate gradients into global memory, with the color extended to a color grid
31:   end for
32: end procedure
```

---



---

**Algorithm 2** Forward and backward algorithm of bilinear interpolation at  $u, v$ , given resolution  $N$ , spatial extend  $\sigma$ , and an input color grid  $C \in \mathbb{R}^{N,N,3}$ .

---

```

1: function BILINEAR( $u, v, N, \sigma, C$ )
2:    $u' = \text{clamp}(\frac{(N-1)*(u+\sigma)}{2*\sigma}, 0, N-1)$  ▷ From  $[-\sigma, +\sigma]$  to  $[0, N-1]$ 
3:    $v' = \text{clamp}(\frac{(N-1)*(v+\sigma)}{2*\sigma}, 0, N-1)$ 
4:    $i_u = \text{floor}(u)$  ,  $f_u = u' - i_u$  ▷ Integer index and fractional part
5:    $i_v = \text{floor}(v)$  ,  $f_v = v' - i_v$ 
6:    $\mathbf{c} =$  ▷ Interpolation
7:      $C[i_u, i_v, :] * (1 - f_u) * (1 - f_v) +$ 
8:      $C[i_u + 1, i_v, :] * f_u * (1 - f_v) +$ 
9:      $C[i_u, i_v + 1, :] * (1 - f_u) * f_v +$ 
10:     $C[i_u + 1, i_v + 1, :] * f_u * f_v$ 
11:   return  $\mathbf{c}$ 
12: end function
13: function ADJ-BILINEAR( $u, v, N, \sigma, C, \hat{\mathbf{c}}$ )
     $\hat{\mathbf{c}}$  indicates the adjoint variable of variable  $\mathbf{c}$ 
14:    $u' = \text{clamp}(\frac{(N-1)*(u+\sigma)}{2*\sigma}, 0, N-1)$  ▷ From  $[-\sigma, +\sigma]$  to  $[0, N-1]$ 
15:    $v' = \text{clamp}(\frac{(N-1)*(v+\sigma)}{2*\sigma}, 0, N-1)$ 
16:    $i_u = \text{floor}(u)$  ,  $f_u = u' - i_u$  ▷ Integer index and fractional part
17:    $i_v = \text{floor}(v)$  ,  $f_v = v' - i_v$ 
    Adjoint for color:
18:    $\hat{C}[i_u, i_v] = \hat{\mathbf{c}} * (1 - f_u) * (1 - f_v)$ 
19:    $\hat{C}[i_u + 1, i_v] = \hat{\mathbf{c}} * f_u * (1 - f_v)$ 
20:    $\hat{C}[i_u, i_v + 1] = \hat{\mathbf{c}} * (1 - f_u) * f_v$ 
21:    $\hat{C}[i_u + 1, i_v + 1] = \hat{\mathbf{c}} * f_u * f_v$ 
    Adjoint for fractional grid position  $u, v$  (uses dot-product):
22:    $\hat{f}_u = \hat{\mathbf{c}} \cdot ((C[i_u + 1, i_v, :] - C[i_u, i_v, :]) * (1 - f_v) + (C[i_u + 1, i_v + 1, :] - C[i_u, i_v + 1, :]) * f_v)$ 
23:    $\hat{f}_v = \hat{\mathbf{c}} \cdot ((C[i_u, i_v + 1, :] - C[i_u, i_v, :]) * (1 - f_u) + (C[i_u + 1, i_v + 1, :] - C[i_u + 1, i_v, :]) * f_u)$ 
    Adjoint for  $u, v$  coordinates:
24:    $\hat{u} = (u > -\sigma \text{ and } u < +\sigma) ? \left( \frac{(N-1)}{2\sigma} \hat{f}_u \right) : 0$ 
25:    $\hat{v} = (v > -\sigma \text{ and } v < +\sigma) ? \left( \frac{(N-1)}{2\sigma} \hat{f}_v \right) : 0$ 
26:   return  $\hat{C}, \hat{u}, \hat{v}$ 
27: end function

```

---