

RAGA: Real Time Ray Traced Gaussian Shadow Casting for 3DGS Avatar-Scene Interaction

Aymen Mir^{1*}, Riza Alp Guler³, Jian Wang⁵, Peter Wonka^{4,5}, Bing Zhou⁵, and Gerard Pons-Moll^{1,2}

¹ Tübingen AI Center, University of Tübingen, Germany

² Max Planck Institute for Informatics, Saarland Informatics Campus, Germany

³ Imperial College London, UK

⁴ King Abdullah University of Science and Technology (KAUST), Saudi Arabia

⁵ Snap Inc., USA



Fig. 1: RAGA casts physically plausible shadows for animated 3DGS avatars composited into 3DGS scenes. Our ray-traced Gaussian shadow casting produces soft, temporally stable shadows entirely in Gaussian space, without any meshing.

Abstract. We study the problem of physically plausible shadow casting when animating 3D Gaussian Splatting (3DGS) avatars, either individually or in multi-avatar and object-interaction scenarios, within existing 3DGS scenes. In contrast to prior methods that rely on binary hit tests and mesh-based shadow casters, our method performs shadow computation entirely in Gaussian space, without requiring any mesh reconstruction. We introduce RAGA, a Ray-Traced Gaussian Shadow Casting formulation based on exact ray-Gaussian line integrals. For each occluding Gaussian, we integrate the opacity profile along the shadow ray and normalize by the theoretical maximum integral, producing a weight that captures *how* the ray traverses the occluder rather than merely whether an intersection occurred. To reduce temporal variance from clothing deformations in animated avatars, we further introduce an avatar proxy representation that stabilizes shadow casting while preserving visual fidelity. We implement

*Work partly done as intern at Snap Inc.

RAGA using custom CUDA kernels integrated with the NVIDIA OptiX framework; as such, our shadow tracer runs at rates of about 50 FPS. We evaluate on single-avatar, multi-avatar, and avatar–object interaction scenarios across multiple datasets, demonstrating substantially improved shadow realism, temporal stability, and scene coherence. Our project page is available at <https://miraymen.github.io/raga/>.

1 Introduction

Recent advances in 3D Gaussian Splatting (3DGS) [53] have demonstrated that, beyond high-quality scene reconstruction, Gaussian-based representations can support compelling dynamic content, including animatable 3D human avatars embedded into complex real-world scenes [75, 76, 78]. This opens the possibility of using 3DGS as a representation in content creation and simulation pipelines such as virtual production, CG compositing, and simulators for robotics and autonomous driving, which have traditionally relied on mesh-based assets.

Despite this progress, a fundamental limitation persists: physically plausible shadows are largely absent when animating 3DGS avatars interacting with static scenes or objects. As a result, inserted humans often appear visually detached from their environment, undermining realism and scene coherence, especially in dynamic settings where shadows provide critical spatial and temporal cues.

No existing method supports shadow casting at composition time (when an avatar is inserted into a new scene) in the volumetric setting where both caster (avatar) and receiver (scene) are 3D Gaussians. Tab. 1 summarizes these gaps.

Classical Shadow Mapping (Tab. 1, top block) [66, 90, 108] is not directly applicable in this setting because, unlike meshes, 3DGS is a volumetric, opacity-based representation that lacks a well-defined inside–outside structure or explicit surface geometry. Existing shadow mapping pipelines rely on surface intersections and depth tests, assumptions that do not hold for 3DGS.

Inverse rendering methods (Tab. 1, middle block) [7, 11, 20, 24, 47, 64, 111, 131] decompose scenes into materials and lighting, recovering per-scene self-shadows through baked occlusion volumes [64], ray-traced visibility [24], or learned shadow parameters [20, 131]. However, these are per-scene optimizations on static geometry: they cannot handle dynamic avatars, and compositing a new avatar into a scene does not update the baked or learned shadow representation.

Ray-based splatting methods (Tab. 1, bottom block) [10, 79, 110] perform ray–Gaussian intersection for rendering but rely on binary hit tests for shadow casting, ignoring how much of a Gaussian the ray actually traverses. Moreover, these methods assume mesh-based shadow casters; they do not address the fully volumetric setting where both caster and receiver are 3D Gaussians. Reconstructing meshes from 3DGS dynamic avatars is itself non-trivial [119] and often reduces visual fidelity.

RAGA. In this work, we address this gap by introducing a principled, fully Gaussian-based solution to shadow casting for animated 3DGS avatars and objects (Fig. 1). We propose RAGA, a Ray-Traced Gaussian Shadow Casting formulation that computes shadows using only Gaussian particles as rendering primitives, without requiring any mesh reconstruction or hybrid representation.

Our method constructs a ray-traced shadow map by shooting rays from scene Gaussians toward light sources and modeling light attenuation induced by occluding Gaussians. For each occluder, we compute exact entry and exit points where the shadow ray intersects the Gaussian’s support, integrate its opacity profile between those bounds, which admits a closed-form solution via the error function, and normalize by the theoretical maximum integral for that Gaussian.

Table 1: Capability Comparison

	Composable	3DGS cast.	3DGS recv.	Dyn. 3DGS avat.	Temp. stab.
Shadow maps [108]	✓	×	×	×	✓
Deep shadow [66]	✓	×	×	×	✓
GS-IR [64]	×	✓	✓	×	×
IRGS [24]	×	✓	✓	×	×
GS ³ [7]	×	✓	✓	×	×
3DGRT [†] [79]	✓	×	✓	×	×
RaySplat [†] [10]	✓	×	✓	×	×
Ours (RAGA)	✓	✓	✓	✓	✓

[†] Can be adapted for use with 3DGS shadow casters.

This normalization is the key design choice: the resulting weight captures *how* the ray traverses each occluder. A ray that merely grazes a Gaussian accumulates little opacity and receives a low weight; a ray passing through a long Gaussian along its major axis receives a high weight; and crucially, the same Gaussian crossed along its narrow direction receives a proportionally lower weight, correctly reflecting the reduced volumetric obstruction.

Adaptation of ray-based methods such as 3DGRT [79, 110], or RaySplat [10] for shadow casting with 3D Gaussian shadow casters fails: their binary hit tests treat all intersections equally regardless of traversal depth, and 3DGRT’s icosahedron proxy for each Gaussian introduces shadow artifacts due to the geometric mismatch with the smooth volumetric Gaussian. Our exact line-integral formulation eliminates these issues.

Temporal stability. To further reduce temporal variance caused by high-frequency clothing deformations in animated avatars, we introduce an avatar proxy representation that approximates pose-dependent changes in avatar Gaussians with isotropic Gaussians that preserve most of the geometric structure and reduce temporal variation in the avatar Gaussians. This proxy significantly reduces temporal jitter for moving avatars interacting with the environment.

We implement RAGA using custom CUDA kernels integrated with the NVIDIA OptiX framework, leveraging hardware ray tracing cores on modern GPUs; our shadow tracer alone runs at interactive rates of about 50 FPS. We evaluate our method on multiple datasets, including single-avatar, multi-avatar, and avatar–object interaction scenarios, demonstrating substantial improvements in shadow realism, stability, and scene integration. To summarize, our contributions are as follows:

- 1) We introduce a fully Gaussian-space shadow casting framework for 3DGS avatars animated in 3DGS scenes, without the need for mesh reconstruction.
- 2) We propose RAGA, a ray-traced shadow casting formulation based on exact ray–Gaussian line integrals and normalized attenuation, replacing the approximate proxy-geometry intersections used in prior work.

- 3) We introduce an avatar proxy strategy that stabilizes shadows under dynamic deformations.
- 4) We extensively evaluate on diverse avatar and interaction scenarios, demonstrating improved realism and temporal stability over existing approaches.

2 Related Work

Neural Rendering. NeRF [73] drove rapid progress in neural rendering [115] but remains costly even with accelerations [5, 6, 82, 103]. 3DGS [53] provides an efficient explicit alternative [56, 107], extended to dynamic scenes [57, 61, 68, 96, 109], SLAM [52], mesh extraction [25, 39], and sparse-view NVS [72]. Existing shadow casting for Gaussians [9] does not handle shadows *received* in 3DGS scenes, ray-traced variants [7, 10, 79, 110] do not cast onto scenes, and relighting methods [13, 23, 64] do not address dynamic avatars. Our method casts shadows from animated 3DGS avatars onto 3DGS scenes without mesh reconstruction.

Inverse Rendering and Shadows in 3DGS. Several methods decompose 3DGS scenes into materials and lighting, recovering self-shadows via baked occlusion [64], precomputed visibility [23], environment shading [47], differentiable ray tracing [24], deferred shading [11], learned shadow parameters [20, 131], or screen-space ray tracing [111]. ShadowGS [69] extends this to satellite imagery. All are per-scene optimizations on static geometry that cannot handle dynamic avatars or composited objects. Our method operates at composition time.

Human Reconstruction and Rendering. Mesh templates [67, 86] recover shape and pose [8, 51] but lack photorealism [2, 3, 74]. Implicit representations [71, 85] model detailed clothed humans [4, 12, 17, 37, 40, 93] but struggle with re-posable rendering. Controllable NeRFs [26, 31, 42, 46, 60, 65, 87, 106, 116, 133] and 3DGS avatars [1, 49, 55, 58, 63, 80, 81, 84, 89, 117, 134] produce photorealistic humans but omit scene interactions. Joint human–environment methods [76, 118, 121] do not address consistent shadows. Our method fills this gap.

Humans and Scenes. Human–scene interaction spans affordance estimation [22, 29, 105], large-scale HSI datasets [16, 27, 30, 32, 33, 45, 70, 88, 95, 102, 126], joint human–object reconstruction [112–114, 123], and object-conditioned motion [19, 34, 77, 99, 126–129]. These rely on mesh representations. Our method operates entirely in 3DGS space.

Human Relighting. Portrait relighting uses light-stage data [44, 50, 54, 83, 100] and diffusion models [36, 38, 91, 97, 98]. Inverse-rendering approaches recover relightable avatars [14, 15, 28, 35, 43, 48, 92, 94, 104, 125], recently extended with Gaussian-based relighting [21, 59, 62, 122, 130] and subsurface scattering [18]. These study relighting in isolation; our method focuses on shadow casting for avatars composited into 3DGS scenes with temporal consistency.

3 Method

Given a 3DGS scene and one or more animated 3DGS avatars, our goal is to compute physically plausible shadows cast by the avatars onto the scene at each

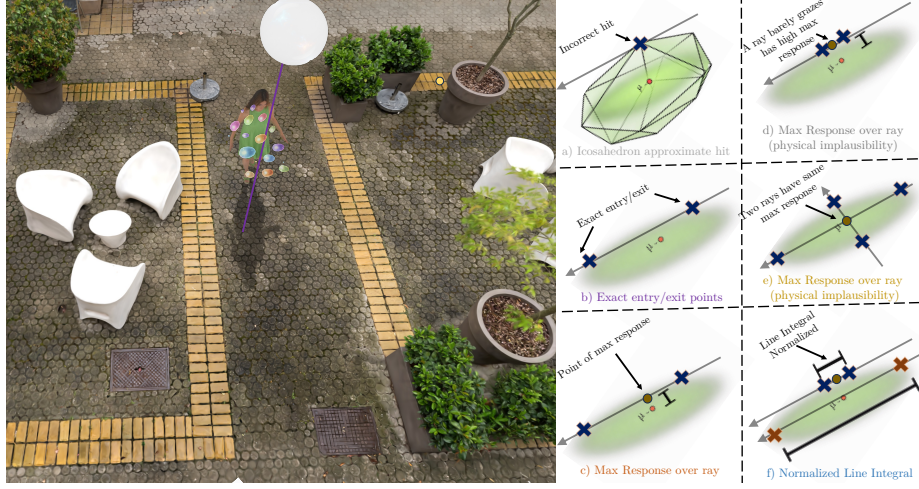


Fig. 2: Method overview. Left: a shadow ray is cast from a scene Gaussian toward the light source through an animated 3DGS avatar by accumulating ray transmittance over all Gaussians in the avatar. Right: comparison of ray–Gaussian intersection strategies. (a) The icosahedron proxy used by 3DGRT [79] produces incorrect hits outside the Gaussian support. (b) Our exact quadratic test (Sec. 3.2) finds precise entry/exit points. (c) Evaluating the maximum Gaussian response along the ray ignores traversal depth (Sec. 3.2). (d) The point of maximum response does not reflect total volumetric obstruction (Sec. 3.2). (e) Two rays with different traversals receive the same max response, and a grazing ray is overweighted (Sec. 3.2). (f) Our normalized line integral (Sec. 3.3, Eq. 17) correctly captures how much of each Gaussian the ray traverses.

frame, operating entirely in Gaussian space without mesh reconstruction. Fig. 2 provides an overview.

Our core idea is to formulate each Gaussian’s shadow contribution as a *normalized line integral* along the shadow ray (Fig. 2f). For each scene Gaussian (receiver), we cast a ray toward the light source and, for every occluding avatar Gaussian the ray encounters, we: (i) compute exact entry and exit points where the ray intersects the Gaussian’s bounded support, its χ^2 -confidence ellipsoid (Fig. 2b, Sec. 3.2), rather than relying on an icosahedron proxy that can produce incorrect hits (Fig. 2a); (ii) integrate the Gaussian’s opacity profile between those bounds, which admits a closed-form solution via the error function (Sec. 3.3); and (iii) normalize the result by the theoretical maximum integral, the value obtained when a ray passes through the center of the Gaussian along its longest axis (Sec. 3.4). This stands in contrast to simply evaluating the peak Gaussian response (Fig. 2c), which is direction-agnostic and fails for anisotropic Gaussians (Fig. 2d,e).

This normalization is the key design choice: the resulting factor $\eta \in [0, 1]$ captures *how* the ray traverses the Gaussian. A ray that merely grazes an occluder produces a small line integral and thus a low factor (Fig. 2d). A ray passing through a long, thin Gaussian along its major axis accumulates substantial opacity and receives a high factor. Crucially, a ray crossing the *same* Gaussian

along its narrow direction receives a low factor, correctly reflecting the reduced volumetric obstruction. The factor η then modulates the Gaussian’s opacity α when updating the shadow transmittance (Sec. 3.5). All ray–Gaussian queries are accelerated via a BVH built over conservative bounding boxes, enabling hardware ray tracing through NVIDIA OptiX (Sec. 3.6). Finally, to suppress temporal flickering from the pose-dependent Gaussian generator, we introduce an avatar proxy that ensures temporally coherent shadow casting (Sec. 3.8). All formal propositions and proofs supporting the derivations below are provided in the supplementary material.

3.1 Ray–Gaussian Shadowing

We represent each occluder as an anisotropic 3D Gaussian with center $\boldsymbol{\mu} \in \mathbb{R}^3$, rotation $\mathbf{R} \in \text{SO}(3)$, and per-axis scales $\mathbf{s} = (s_x, s_y, s_z)$, giving the covariance

$$\boldsymbol{\Sigma} = \mathbf{R} \text{diag}(s_x^2, s_y^2, s_z^2) \mathbf{R}^\top. \quad (1)$$

We use the unnormalized Gaussian kernel

$$g(\mathbf{x}) = \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right), \quad (2)$$

whose maximum is $g(\boldsymbol{\mu}) = 1$ (Lemma 1 in supp. mat.).

Given a receiver point $\mathbf{o}$ and a point light position $\boldsymbol{\ell}$, we cast a shadow ray

$$\mathbf{x}(\tau) = \mathbf{o} + \tau \mathbf{v}, \quad \tau \in [0, 1], \quad \mathbf{v} = \boldsymbol{\ell} - \mathbf{o}. \quad (3)$$

Gaussian-space transform. To simplify the intersection test, we first define the whitening transform

$$\mathbf{L} = \text{diag}(s_x^{-1}, s_y^{-1}, s_z^{-1}) \mathbf{R}^\top, \quad (4)$$

so that $\mathbf{y} = \mathbf{L}(\mathbf{x} - \boldsymbol{\mu})$ satisfies $\|\mathbf{y}\|_2^2 = (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})$. The ray becomes

$$\mathbf{y}(\tau) = \mathbf{o}' + \tau \mathbf{v}', \quad \mathbf{o}' = \mathbf{L}(\mathbf{o} - \boldsymbol{\mu}), \quad \mathbf{v}' = \mathbf{L}\mathbf{v}. \quad (5)$$

3.2 Exact Ellipsoid Intersection and Entry/Exit

We restrict contributions to the χ^2 -confidence ellipsoid (we use $\chi^2 = 9$, corresponding to a 3σ cutoff):

$$\|\mathbf{y}\|_2^2 \leq \chi^2, \quad (6)$$

which yields a quadratic constraint along the ray (Prop. 1 in supp. mat.):

$$\|\mathbf{o}' + \tau \mathbf{v}'\|_2^2 = A\tau^2 + 2B\tau + C, \quad A = \|\mathbf{v}'\|_2^2, \quad B = \mathbf{o}'^\top \mathbf{v}', \quad C = \|\mathbf{o}'\|_2^2. \quad (7)$$

The maximum-response point on the ray is the minimum of Eq. 7, attained at

$$\tau_c = -\frac{B}{A}, \quad (8)$$

and the minimum Mahalanobis distance squared to the ray is

$$d_{\perp}^2 = \|\mathbf{o}' + \tau_c \mathbf{v}'\|_2^2 = C - \frac{B^2}{A}. \quad (9)$$

If $d_{\perp}^2 > \chi^2$ the ray misses the ellipsoid; otherwise we define

$$\Delta = \chi^2 - d_{\perp}^2 \geq 0, \quad (10)$$

and the entry/exit times on the line are (Props. 2, Cor. 1 in supp. mat.)

$$\tau_{\text{in}} = \tau_c - \sqrt{\frac{\Delta}{A}}, \quad \tau_{\text{out}} = \tau_c + \sqrt{\frac{\Delta}{A}}. \quad (11)$$

We clamp to the ray segment,

$$\tilde{\tau}_{\text{in}} = \max(\tau_{\text{in}}, 0), \quad \tilde{\tau}_{\text{out}} = \min(\tau_{\text{out}}, 1), \quad (12)$$

and discard the Gaussian if $\tilde{\tau}_{\text{in}} \geq \tilde{\tau}_{\text{out}}$.

Why not evaluate at the peak? A simpler alternative to integration would be to use the Gaussian's response at the closest point τ_c (Eq. 8), i.e. $g(\mathbf{x}(\tau_c)) = \exp(-\frac{1}{2}d_{\perp}^2)$ (Fig. 2c). However, this value depends only on the perpendicular distance $d_{\perp}$ to the Gaussian center and is entirely *direction-agnostic*: two rays at the same distance but traversing vastly different distances through the Gaussian receive identical weights (Fig. 2e). Worse, a ray barely grazing a large Gaussian along its major axis can have a small $d_{\perp}$ and thus a high peak response, despite minimal volumetric obstruction (Fig. 2d). The line integral (Sec. 3.3) resolves both issues by accumulating opacity over the full traversal length, naturally producing lower weights for shorter paths through the Gaussian.

3.3 Truncated Line Integral (Closed Form)

Along the ray, Eq. 2 reduces to a 1D Gaussian in τ :

$$g(\mathbf{x}(\tau)) = \exp\left(-\frac{1}{2}\|\mathbf{o}' + \tau \mathbf{v}'\|_2^2\right) = \exp\left(-\frac{1}{2}d_{\perp}^2\right) \exp\left(-\frac{1}{2}A(\tau - \tau_c)^2\right). \quad (13)$$

We integrate over the ellipsoid intersection interval,

$$I_{\text{line}} = \int_{\tilde{\tau}_{\text{in}}}^{\tilde{\tau}_{\text{out}}} g(\mathbf{x}(\tau)) d\tau, \quad (14)$$

which admits a closed-form solution via the error function (Props. 3–4 in supp. mat.):

$$I_{\text{line}} = \exp\left(-\frac{1}{2}d_{\perp}^2\right) \sqrt{\frac{\pi}{2A}} \left[\text{erf}\left(\sqrt{\frac{A}{2}}(\tilde{\tau}_{\text{out}} - \tau_c)\right) - \text{erf}\left(\sqrt{\frac{A}{2}}(\tilde{\tau}_{\text{in}} - \tau_c)\right) \right]. \quad (15)$$

3.4 Normalization by Maximum Line Integral

To make the integral dimensionless and comparable across Gaussians of varying size, we normalize by the maximum attainable line integral within the same χ^2 bound: a center pass ($d_{\perp}^2 = 0$) through the longest axis $s_{\max} = \max(\mathbf{s})$ (proof in Sec. 2.5 of supp. mat.),

$$I_{\max} = \sqrt{2\pi} \frac{s_{\max}}{\|\mathbf{v}\|_2} \operatorname{erf}\left(\sqrt{\frac{\chi^2}{2}}\right), \quad (16)$$

where the factor $s_{\max}/\|\mathbf{v}\|_2$ converts between the segment parameter $\tau \in [0, 1]$ of Eq. 3 and world distance along the ray. We then define the normalized thickness factor

$$\eta = \operatorname{clamp}\left(\frac{I_{\text{line}}}{I_{\max}}, 0, 1\right). \quad (17)$$

The factor $\eta \in [0, 1]$ (Prop. 5 in supp. mat.) is near 0 for grazing intersections and approaches 1 for a center pass along the major axis.

3.5 From Opacity to Shadow Transmittance

Each occluder Gaussian carries an opacity parameter $\alpha \in (0, 1)$. We update the shadow transmittance T multiplicatively over all Gaussians intersecting the ray (found via BVH traversal and the exact test above). The normalized thickness factor η modulates each Gaussian’s effective attenuation. We consider three update rules.

Linear normalized model. Using η as a fractional coverage, we set

$$T \leftarrow T(1 - \alpha\eta). \quad (18)$$

Exponentiated normalized model. A more thickness-consistent variant treats η as a fraction of optical thickness,

$$T \leftarrow T(1 - \alpha)^\eta, \quad (19)$$

which is equivalent to attenuating with an effective opacity (Prop. 6 in supp. mat.)

$$\alpha_{\text{eff}} = 1 - (1 - \alpha)^\eta, \quad T \leftarrow T(1 - \alpha_{\text{eff}}). \quad (20)$$

Beer–Lambert integral model. Mapping opacity to an extinction coefficient $\kappa(\alpha) = -\ln(1 - \alpha) \geq 0$ gives the physically motivated update (Prop. 7 in supp. mat.)

$$T \leftarrow T \exp(-\kappa(\alpha) I_{\text{line}}). \quad (21)$$

We use the exponentiated model in all experiments (ablation in Sec. 4.5).

3.6 Acceleration via BVH over Conservative AABBs

To accelerate intersection queries, we build a BVH over per-Gaussian axis-aligned bounding boxes that conservatively enclose the χ^2 ellipsoid in world space. The ellipsoid projection radius along world axis $d \in \{x, y, z\}$ is

$$e_d = \sqrt{\chi^2 (s_x^2 \mathbf{R}_{d1}^2 + s_y^2 \mathbf{R}_{d2}^2 + s_z^2 \mathbf{R}_{d3}^2)}, \quad (22)$$

yielding AABB = $[\boldsymbol{\mu} - \mathbf{e}, \boldsymbol{\mu} + \mathbf{e}]$. Only BVH candidates invoke the exact quadratic test Eqs. 7–11 and integral Eq. 15, enabling efficient soft shadow estimation in dense 3DGS scenes. We implement the BVH using NVIDIA OptiX, leveraging hardware RT cores for near real-time performance (Sec. 4.6).

3.7 Shadow Rendering

After computing the transmittance $T_k \in [0, 1]$ for each scene Gaussian k , we modulate its color during 3DGS rasterization. The shadowed color of Gaussian k is $\mathbf{c}'_k = T_k \mathbf{c}_k$, where $T_k = 1$ indicates full illumination and $T_k \rightarrow 0$ indicates full shadow. The modified colors $\mathbf{c}'_k$ are then splatted using standard α -blending, producing the final shadowed rendering.

3.8 Avatar Proxy for Temporally Stable Shadows

While the pose-dependent 3DGS avatar model produces high-fidelity renderings, using the raw primitives for shadow casting introduces temporal instability. The pose-to-geometry generator $\mathcal{H}(\boldsymbol{\theta}_t)$, a neural network that maps body pose $\boldsymbol{\theta}_t$ to 3D Gaussian primitives, produces outputs that fluctuate in topology, opacity, and anisotropic orientation to satisfy view-dependent photometric losses. When used as shadow occluders, these fluctuations cause the shadow silhouette to flicker and oscillate rapidly between frames.

To address this, we decouple the render-ready geometry from the shadow-casting geometry by constructing a **stable avatar proxy** (Fig. 3), a regularized derivative of the avatar that guarantees temporal coherence.

Canonical topology locking. To eliminate topological noise, we freeze the generative process at a canonical reference pose $\boldsymbol{\theta}_{\text{ref}}$ (e.g., a T-pose), caching the configuration once:

$$\mathcal{P}_{\text{proxy}} = \{(\boldsymbol{\mu}_k^{(0)}, \mathbf{s}_k^{(0)}, \alpha_k^{(0)})\}_{k=1}^N = \mathcal{H}(\boldsymbol{\theta}_{\text{ref}}). \quad (23)$$

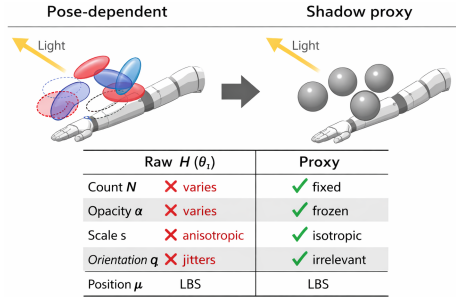


Fig. 3: Avatar proxy. We replace the pose-dependent anisotropic Gaussians (left) with a stable isotropic proxy (right) that eliminates temporal flickering in shadows.

Table 2: Shadow quality vs. pseudo-GT on ScanNet++. Averaged over 5 scenes. Mesh Shadow is an oracle.

Method	SAE↓	SM-IoU↑	BF↑
Mesh Shadow (oracle)	0.000	1.000	1.000
Ours (RAGA)	0.031	0.847	0.812
3DGRT [79] (mod.)	0.058	0.741	0.693
RaySplat [10] (mod.)	0.062	0.728	0.679

Table 3: Temporal consistency (TSC↓). Averaged over 5 sequences. Lower = more stable.

Method	TSC↓
Ours (full proxy)	0.0018
Ours w/o proxy	0.0061
3DGRT [79] (mod.)	0.0114
RaySplat [10] (mod.)	0.0121

By fixing the number of primitives N and their initial attributes, we ensure a persistent set of occluders, preventing artifacts caused by primitives appearing and disappearing between frames.

Isotropic regularization. To enforce rotational invariance, we override the anisotropic scale with a conservative isotropic radius $r_k = \max(s_{k,x}^{(0)}, s_{k,y}^{(0)}, s_{k,z}^{(0)})$, transforming the covariance into $\Sigma'_k = r_k^2 \mathbf{I}$. Since a sphere’s silhouette is rotation-invariant, the shadow projection becomes independent of the primitive’s orientation.

Simplified kinematics. During inference, we articulate the proxy using Linear Blend Skinning (LBS) applied strictly to positions. Given the canonical center $\mu_k^{(0)}$ and joint transforms $\mathbf{M}_j(\theta_t)$:

$$\mu_k^{(t)} = \sum_{j \in \mathcal{J}} w_{kj} \mathbf{M}_j(\theta_t) \mu_k^{(0)}, \quad (24)$$

where $\mathcal{J}$ denotes the skeleton joints and $w_{kj} \geq 0$ are normalized skinning weights ($\sum_{j \in \mathcal{J}} w_{kj} = 1$). Crucially, we do **not** update the orientation of the proxies. Since the primitives are isotropic spheres, their rotation is irrelevant to the ray intersection test, guaranteeing that the shadow signal is continuous with respect to the pose θ_t .

4 Experiments

We evaluate RAGA along two axes: *shadow quality*, how accurately do our ray-traced Gaussian shadows approximate mesh-based ground truth? (Sec. 4.2); and *temporal stability*, does the avatar proxy reduce flickering compared to raw Gaussian shadow casting? (Sec. 4.3); in both evaluations we compare with two modified ray-traced methods, 3DGRT [79] and RaySplat [10]. We further report *ablations* on each pipeline component (Sec. 4.5) and a timing analysis (Sec. 4.6).

4.1 Experimental Setup

Pseudo-GT Protocol. We obtain pseudo-GT shadow maps $S^\dagger(p)$ via classical mesh-based rendering in two settings:

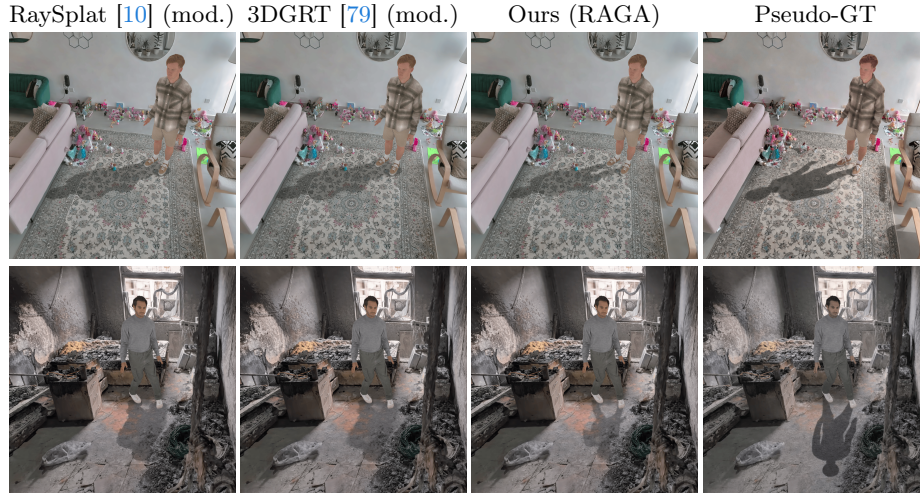


Fig. 4: Qualitative comparison against baselines on SuperSplat [101] scenes. Pseudo-GT is obtained using a proxy floor mesh and posed SMPL mesh via classical shadow mapping. Our method produces soft, physically plausible shadows that closely match the pseudo-GT. Both 3DGRT [79] and RaySplat [10] lose shadow detail due to their implausible attenuation models.

- **ScanNet++ [120] scenes:** Since ScanNet++ provides both a 3DGS and a mesh reconstruction, we replace the avatar with the posed SMPL mesh and cast shadows onto the scene mesh using classical shadow mapping. We render the shadow attenuation onto a transparent surface and composite it with the 3DGS scene rendering. We use 5 ScanNet++ scenes for quantitative evaluation.
- **General 3DGS scenes (e.g. SuperSplat [101]):** For scenes without a mesh reconstruction, we construct an approximate floor plane from the 3DGS point cloud, replace the avatar with the posed SMPL mesh, and cast shadows onto the floor proxy via classical rendering. The floor is set to invisible and the resulting shadow map is composited with the 3DGS rendering.

Why not use mesh-based shadow casting? While the pseudo-GT protocol above uses meshes to establish a reference, mesh-based shadows are not a practical solution for general 3DGS scenes:

- **Scene mesh extraction is lossy.** Obtaining a high-quality mesh from a 3DGS scene is not straightforward and leads to significant loss of detail [119]. While we use a floor proxy for pseudo-GT, this discards all wall- and furniture-based shadow receivers; ScanNet++ is an exception where dense meshes happen to be available.
- **Mesh proxies lose detail.** Mesh-based shadow casting requires replacing occluders with mesh proxies such as SMPL for humans, which discards clothing and hair detail from the shadow silhouette. While our avatar proxy

also smooths fine detail, it retains significantly more geometric fidelity than the bare SMPL mesh because it operates on the actual Gaussian point cloud rather than a parametric body model.

- **No support for Gaussian objects.** Mesh proxies do not generalize to arbitrary 3DGS objects inserted into scenes. Our method operates entirely in Gaussian space and naturally supports avatar–object interactions, as shown in the qualitative results (Sec. 4.7).

Avatars. We use 3DGS avatars from ActorsHQ [41] (48 cameras) and Avatar-ReX [132], spanning a range of reconstruction quality. For multi-avatar and avatar-object scenarios, we additionally use NeuralDome [124].

Baselines. Since no prior method addresses the same problem (shadow casting for animated 3DGS avatars in 3DGS scenes), we construct the strongest possible Gaussian-based baselines by modifying two recent ray-tracing methods for shadow casting:

- **3DGRT** [79] (modified): Following 3DGRT, we first detect ray–Gaussian intersections using an icosahedron approximation around each Gaussian, then for detected intersections we accumulate opacity values to compute ray attenuation. This formulation does not use our line integral or avatar proxy.
- **RaySplat** [10] (modified): We detect ray–Gaussian intersections using the exact analytic test from RaySplat, then accumulate opacity for each detected Gaussian. As with 3DGRT, this formulation uses neither our line integral nor avatar proxy.
- **Mesh Shadow (oracle)**: We render a shadow-only pass by casting shadows from the posed SMPL mesh onto a transparent mesh receiver via classical shadow mapping, then alpha-blend the result onto the 3DGS scene rendering.

4.2 Shadow Quality Evaluation

We evaluate shadow maps against the mesh-based pseudo-GT using three pixel-space metrics computed in an avatar-centric region of interest:

- **SAE** (Shadow Attenuation Error): Mean absolute error between predicted and pseudo-GT shadow within the shadow region Ω_s (lower is better).
- **SM-IoU** (Shadow Matte IoU): Intersection-over-union of binarized shadow masks (threshold 0.1) (higher is better).
- **BF** (Boundary F-measure): F-score between shadow boundaries with 2 px tolerance (higher is better).

Tab. 2 reports shadow quality averaged over 5 ScanNet++ scenes. Our method achieves the best results among all Gaussian-based approaches, approaching the mesh oracle. 3DGRT’s binary hit test based on its icosahedron approximation produces hard shadow boundaries without penumbrae (Fig. 2a), leading to poor SAE and BF scores. RaySplat’s collision-based approach reduces to a max-response evaluation (Fig. 2c–e) that lacks proper attenuation modeling. Fig. 4 shows a visual comparison.

Table 4: Ablation studies. Evaluated against pseudo-GT on ScanNet++. Best in each group is **bold**.

Metric	(A) Attenuation model			(B) Normalization	
	Exp. (Ours)	Linear	Beer-L.	$I_{\max}$ (Ours)	w/o norm.
SAE ↓	0.031	0.035	0.039	0.031	0.047
SM-IoU ↑	0.847	0.831	0.819	0.847	0.793
BF ↑	0.812	0.796	0.781	0.812	0.754
TSC ↓	0.0018	0.0021	0.0024	0.0018	0.0022

Table 5: Avatar proxy ablation. Temporal stability (TSC↓). Lower = more stable.

Variant	TSC↓
Full proxy (Ours)	0.0018
w/o topology lock	0.0037
w/o isotropic reg.	0.0032
No proxy	0.0061


Fig. 5: Visual ablations. Without $I_{\max}$ normalization (left), large anisotropic Gaussians cast disproportionately dark shadows. The linear attenuation model (middle) produces harder shadow falloff than our full exponentiated model (right).

4.3 Temporal Stability

A key contribution of RAGA is the avatar proxy (Sec. 3.8) that ensures temporally stable shadows under animation. We measure temporal shadow consistency (TSC) as the mean per-pixel absolute difference between consecutive shadow frames: $TSC = \frac{1}{F-1} \sum_{t=1}^{F-1} \frac{1}{|\Omega|} \sum_{p \in \Omega} |S_t(p) - S_{t-1}(p)|$, where S_t is the shadow map at frame t , F is the total number of frames, and Ω is the avatar-centric ROI. Lower TSC indicates more stable shadows. Tab. 3 shows that the avatar proxy substantially reduces temporal flickering. Fig. 7 shows temporal difference maps ($|S_t - S_{t-1}|$) confirming that the avatar proxy reduces shadow flickering.

4.4 Perceptual Study

We conduct a two-alternative forced-choice (2AFC) study with 12 naive raters across 5 ScanNet++ scenes, comparing **Ours (RAGA)** against a **No Shadow** baseline. Raters are shown 5 video clips (10 s at 30 fps) per scene with randomized method order and placement, and asked which rendering looks more realistic. On average, raters prefer RAGA 73% of the time.

4.5 Ablation Studies

We ablate each component of RAGA in Tab. 4. For the attenuation model, we compare linear, exponentiated, and Beer-Lambert transmittance updates

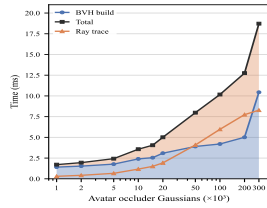


Fig. 6: Runtime vs. avatar Gaussians.



Fig. 7: Proxy ablation. $|S_t - S_{t-1}|$ maps: bright = large change. The proxy stabilizes shadows.

(Eqs. 18–21); the exponentiated model provides the best trade-off. For line integral normalization, we compare the raw line integral I_{line} against our normalized thickness factor $\eta = I_{\text{line}}/I_{\text{max}}$ (Eq. 17); without normalization, large anisotropic Gaussians cast disproportionately dark shadows (Fig. 5). For the avatar proxy, we ablate canonical topology locking and isotropic regularization (Tab. 5).

4.6 Runtime Analysis

We analyze the runtime of RAGA as a function of avatar complexity on an NVIDIA H100 GPU with $\sim 500\text{K}$ scene Gaussians. Fig. 6 reports per-frame shadow computation time as the number of avatar Gaussians varies from 20K to 400K. Only our shadow tracer, implemented as a CUDA kernel, runs at interactive rates of about 50 FPS; the end-to-end pipeline (including 3DGS rasterization) achieves near real-time performance even for dense avatars thanks to OptiX hardware ray tracing and BVH acceleration.

4.7 Qualitative Results

Fig. 8 shows results across diverse scenarios: single-avatar animation, multi-avatar motion, and avatar-object interaction in various ScanNet++ and SuperSplat [101] scenes. Our method produces soft, physically plausible shadows that follow the avatar’s motion and respect scene geometry. We demonstrate results for multiple avatars interacting with scenes or avatars interacting with 3DGS objects placed in 3DGS scenes.

5 Conclusion

We presented RAGA, a principled and efficient framework for physically plausible shadow casting in animated 3DGS avatar-scene interaction scenarios. By operating entirely in Gaussian space, our approach overcomes a fundamental limitation of prior work, which relies on mesh-based shadow computation or approximate ray–Gaussian interactions that break down in dynamic settings.

At the core of RAGA is a ray-traced Gaussian shadow casting formulation that models light attenuation as a set of independent absorption events along rays cast toward light sources. Leveraging the closed-form solution of the ray–Gaussian line



Fig. 8: Qualitative results: our method casts plausible soft shadows for various 3DGS avatars animated alone and with objects across diverse scenes.

integral and analytically computable entry and exit points, our method enables stable and physically grounded aggregation of shadow contributions from multiple volumetric occluders. This principled formulation avoids the approximations commonly used in existing ray tracing approaches for 3DGS and results in significantly improved temporal stability and visual coherence.

To further address shadow instability in animated avatars, we introduced an avatar proxy that reduces variance caused by non-rigid deformations, leading to more consistent shadows during motion without sacrificing visual quality. Our GPU implementation integrates custom CUDA kernels with NVIDIA OptiX, exploiting hardware ray tracing to achieve near real-time performance. Extensive experiments across single-avatar, multi-avatar, and avatar-object interaction scenarios demonstrate that RAGA produces realistic, stable shadows.

While our method advances shadow casting for 3DGS avatars, several limitations remain. We model point and directional light but do not support area lights. Lighting on the avatar and scene remains baked into the 3DGS representation; we modulate scene Gaussian colors to create shadows but do not relight the avatar or scene under novel illumination. Similarly, we do not estimate light source positions from the scene; the user must specify them manually. Future work includes automatic light source estimation from 3DGS scenes, fully relightable avatars that respond to dynamic illumination, and area light support.

Acknowledgements. We gratefully acknowledge support from the hessian.AI Service Center (funded by the Federal Ministry of Research, Technology and Space, BMFTR, grant no. 16IS22091) and the hessian.AI Innovation Lab (funded by the Hessian Ministry for Digital Strategy and Innovation, grant no. S-DIW04/0013/003). We especially thank Patrick Blauth at the hessian.AI Service Center for technical assistance. This work was further supported by the German Federal Ministry of Education and Research (BMBF): Tübingen AI Center, FKZ: 01IS18039A. GPM is a member of the Machine Learning Cluster of Excellence, EXC number 2064/1 – Project number 390727645. The work was also partially supported by funding from KAUST – Center of Excellence for Generative AI, under award number 5940.

References

1. Abdal, R., Yifan, W., Shi, Z., Xu, Y., Po, R., Kuang, Z., Chen, Q., Yeung, D.Y., Wetzstein, G.: Gaussian shell maps for efficient 3d human generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
2. Alldieck, T., Magnor, M., Bhatnagar, B.L., Theobalt, C., Pons-Moll, G.: Learning to reconstruct people in clothing from a single RGB camera. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
3. Alldieck, T., Magnor, M., Xu, W., Theobalt, C., Pons-Moll, G.: Video based reconstruction of 3d people models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
4. Alldieck, T., Xu, H., Sminchisescu, C.: imGHUM: Implicit generative models of 3d human shape and articulated pose. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2021)
5. Barron, J.T., Mildenhall, B., Verbin, D., Srinivasan, P.P., Hedman, P.: Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022)
6. Barron, J.T., Mildenhall, B., Verbin, D., Srinivasan, P.P., Hedman, P.: Zip-nerf: Anti-aliased grid-based neural radiance fields. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2023)
7. Bi, Z., Zeng, Y., Zeng, C., Pei, F., Feng, X., Zhou, K., Wu, H.: Gs³: Efficient relighting with triple gaussian splatting. In: ACM SIGGRAPH Asia Conference Papers (2024)
8. Bogo, F., Kanazawa, A., Lassner, C., Gehler, P., Romero, J., Black, M.J.: Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. In: European Conference on Computer Vision (ECCV) (2016)
9. Bolanos, L., Su, S.Y., Rhodin, H.: Gaussian shadow casting for neural characters. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
10. Byrski, K., Mazur, M., Tabor, J., Dziarmaga, T., Kądziołka, M., Baran, D., Spurek, P.: Raysplats: Ray tracing based gaussian splatting. arXiv preprint arXiv:2501.19196 (2025)
11. Chen, H., Lin, Z., Zhang, J.: GI-GS: Global illumination decomposition on gaussian splatting for inverse rendering. In: International Conference on Learning Representations (ICLR) (2025)
12. Chen, X., Zheng, Y., Black, M.J., Hilliges, O., Geiger, A.: Snarf: Differentiable forward skinning for animating non-rigid neural implicit shapes. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2021)
13. Chen, Y., Chen, Z., Zhang, C., Wang, F., Yang, X., Wang, Y., Cai, Z., Yang, L., Liu, H., Lin, G.: Gaussianeditor: Swift and controllable 3d editing with gaussian splatting. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
14. Chen, Z., Liu, Z.: Relighting4d: Neural relightable human from videos. In: European Conference on Computer Vision (ECCV) (2022)
15. Chen, Z., et al.: Urhand: Universal relightable hands. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)

16. Cheng, W., Chen, R., Yin, W., Fan, S., Chen, K., He, H., Luo, H., Cai, Z., Wang, J., Gao, Y., Yu, Z., Lin, Z., Ren, D., Yang, L., Liu, Z., Loy, C.C., Qian, C., Wu, W., Lin, D., Dai, B., Lin, K.Y.: Dna-rendering: A diverse neural actor repository for high-fidelity human-centric rendering. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2023)
17. Deng, B., Lewis, J.P., Jeruzalski, T., Pons-Moll, G., Hinton, G., Norouzi, M., Tagliasacchi, A.: Nasa neural articulated shape approximation. In: European Conference on Computer Vision (ECCV) (2020)
18. Dihlmann, J.N., Majumdar, A., Engelhardt, A., Braun, R., Lensch, H.P.: Sub-surface scattering for 3D gaussian splatting. In: Advances in Neural Information Processing Systems (NeurIPS) (2024)
19. Diller, C., Dai, A.: Cg-hoi: Contact-guided 3d human-object interaction generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
20. Du, K., Liang, Z., Shen, Y., Wang, Z.: GS-ID: Illumination decomposition on gaussian splatting via adaptive light aggregation and diffusion-guided material priors. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2025)
21. Fan, J., Luan, F., Yang, J., Hasan, M., Wang, B.: Rng: Relightable neural gaussians. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2025)
22. Fouhey, D.F., Delaitre, V., Gupta, A., Efros, A.A., Laptev, I., Sivic, J.: People watching: Human actions as a cue for single view geometry. *International journal of computer vision* **110**(3) (2014)
23. Gao, J., Gu, C., Lin, Y., Li, Z., Zhu, H., Cao, X., Zhang, L., Yao, Y.: Relightable 3d gaussians: Realistic point cloud relighting with brdf decomposition and ray tracing. In: European Conference on Computer Vision (ECCV) (2024)
24. Gu, C., Wei, X., Zeng, Z., Yao, Y., Zhang, L.: IRGS: Inter-reflective gaussian splatting with 2D gaussian ray tracing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2025)
25. Guédon, A., Lepetit, V.: Sugar: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
26. Guo, C., Jiang, T., Chen, X., Song, J., Hilliges, O.: Vid2avatar: 3d avatar reconstruction from videos in the wild via self-supervised scene decomposition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2023)
27. Guo, C., Zou, S., Zuo, X., Wang, S., Ji, W., Li, X., Cheng, L.: Generating diverse and natural 3d human motions from text. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022)
28. Guo, K., Lincoln, P., Davidson, P., Busch, J., Yu, X., Whalen, M., Harvey, G., Orts Escolano, S., Pandey, R., Dourgarian, J., Tang, D., Tkach, A., Kowdle, A., Cooper, E., Dou, M., Fanello, S., Fyffe, G., Rhemann, C., Taylor, J., Debevec, P., Izadi, S.: The relightables: Volumetric performance capture of humans with realistic relighting. *ACM Transactions on Graphics* (2019)
29. Gupta, A., Satkin, S., Efros, A.A., Hebert, M.: From 3d scene geometry to human workspace. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2011)

30. Guзов, V., Mir, A., Sattler, T., Pons-Moll, G.: Human poseitioning system (hps): 3d human pose estimation and self-localization in large scenes from body-mounted sensors. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021)
31. Habermann, M., Liu, L., Xu, W., Pons-Moll, G., Zollhoefer, M., Theobalt, C.: Hdhumans: A hybrid approach for high-fidelity digital humans. *Proceedings of the ACM on Computer Graphics and Interactive Techniques* **6**(3) (2023)
32. Hassan, M., Ceylan, D., Villegas, R., Saito, J., Yang, J., Zhou, Y., Black, M.: Stochastic Scene-Aware motion prediction. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2021)
33. Hassan, M., Choutas, V., Tzionas, D., Black, M.J.: Resolving 3D human pose ambiguities with 3D scene constraints. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2019)
34. Hassan, M., Ghosh, P., Tesch, J., Tzionas, D., Black, M.J.: Populating 3D scenes by learning human-scene interaction. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021)
35. Hasselgren, J., Hofmann, N., Munkberg, J.: Shape, light, and material decomposition from images using monte carlo rendering and denoising. In: *Advances in Neural Information Processing Systems (NeurIPS)* (2022)
36. He, M., Clausen, P., Taşel, A.L., Ma, L., Pilarski, O., Xian, W., Rikker, L., Yu, X., Burgert, R., Yu, N., Debevec, P.: Diffrelight: Diffusion-based facial performance relighting. In: *ACM SIGGRAPH Asia Conference Papers* (2024)
37. He, T., Xu, Y., Saito, S., Soatto, S., Tung, T.: Arch++: Animation-ready clothed human reconstruction revisited. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2021)
38. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: *Advances in Neural Information Processing Systems (NeurIPS)* (2020)
39. Huang, B., Yu, Z., Chen, A., Geiger, A., Gao, S.: 2d gaussian splatting for geometrically accurate radiance fields. In: *ACM SIGGRAPH Conference Proceedings* (2024)
40. Huang, Z., Xu, Y., Lassner, C., Li, H., Tung, T.: Arch: Animatable reconstruction of clothed humans. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2020)
41. Işık, M., Rünz, M., Georgopoulos, M., Khakhulin, T., Starck, J., Agapito, L., Nießner, M.: Humanrf: High-fidelity neural radiance fields for humans in motion. *ACM Transactions on Graphics* **42**(4) (2023)
42. Iqbal, U., Caliskan, A., Nagano, K., Khamis, S., Molchanov, P., Kautz, J.: Rana: Relightable articulated neural avatars. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2023)
43. Iwase, S., Saito, S., Simon, T., Lombardi, S., Bagautdinov, T., Joshi, R., Prada, F., Shiratori, T., Sheikh, Y., Saragih, J.: Relightablehands: Efficient neural relighting of articulated hand models. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2023)
44. Ji, C., Yu, T., Guo, K., Liu, J., Liu, Y.: Geometry-aware single-image full-body human relighting. In: *European Conference on Computer Vision (ECCV)* (2022)
45. Jiang, N., Zhang, Z., Li, H., Ma, X., Wang, Z., Chen, Y., Liu, T., Zhu, Y., Huang, S.: Scaling up dynamic human-scene interaction modeling. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024)

46. Jiang, W., Yi, K.M., Samei, G., Tuzel, O., Ranjan, A.: Neuman: Neural human radiance field from a single video. In: European Conference on Computer Vision (ECCV) (2022)
47. Jiang, Y., Tu, J., Liu, Y., Gao, X., Long, X., Wang, W., Ma, Y.: GaussianShader: 3D gaussian splatting with shading functions for reflective surfaces. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
48. Jin, H., Liu, I., Xu, P., Zhang, X., Han, S., Bi, S., Zhou, X., Xu, Z., Su, H.: Tensor: Tensorial inverse rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2023)
49. Junkawitsch, H., Sun, G., Zhu, H., Theobalt, C., Habermann, M.: Eva: Expressive virtual avatars from multi-view videos. In: ACM SIGGRAPH Conference Proceedings (2025)
50. Kanamori, Y., Endo, Y.: Relighting humans: Occlusion-aware inverse rendering for full-body human images. *ACM Transactions on Graphics* **37**(6) (2018)
51. Kanazawa, A., Black, M.J., Jacobs, D.W., Malik, J.: End-to-end recovery of human shape and pose. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
52. Keetha, N., Karhade, J., Jatavallabhula, K.M., Yang, G., Scherer, S., Ramanan, D., Luiten, J.: Splatam: Splat, track & map 3d gaussians for dense rgb-d slam. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
53. Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G.: 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics* **42**(4) (2023)
54. Kim, H., Jang, M., Yoon, W., Lee, J., Na, D., Woo, S.: Switchlight: Co-design of physics-driven architecture and pre-training framework for human portrait relighting. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
55. Kocabas, M., Chang, J.H.R., Gabriel, J., Tuzel, O., Ranjan, A.: HUGS: Human gaussian splats. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
56. Lassner, C., Zollhöfer, M.: Pulsar: Efficient sphere-based neural rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021)
57. Lee, J., Won, C., Jung, H., Bae, I., Jeon, H.G.: Fully explicit dynamic gaussian splatting. In: Advances in Neural Information Processing Systems (NeurIPS) (2024)
58. Lei, J., Wang, Y., Pavlakos, G., Liu, L., Daniilidis, K.: GART: Gaussian articulated template models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
59. Li, J., Cao, C., Schwartz, G., Khirrodar, R., Richardt, C., Simon, T., Sheikh, Y., Saito, S.: Uravatar: Universal relightable gaussian codec avatars. In: ACM SIGGRAPH Asia Conference Papers (2024)
60. Li, R., Tanke, J., Vo, M., Zollhofer, M., Gall, J., Kanazawa, A., Lassner, C.: Tava: Template-free animatable volumetric actors. In: European Conference on Computer Vision (ECCV) (2022)
61. Li, Z., Chen, Z., Li, Z., Xu, Y.: Spacetime gaussian feature splatting for real-time dynamic view synthesis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)

62. Li, Z., Sun, Y., Zheng, Z., Wang, L., Zhang, S., Liu, Y.: Animatable and relightable gaussians for high-fidelity human avatar modeling. *arXiv preprint arXiv:2311.16096* (2024)
63. Li, Z., Zheng, Z., Wang, L., Liu, Y.: Animatable gaussians: Learning pose-dependent gaussian maps for high-fidelity human avatar modeling. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024)
64. Liang, Z., Zhang, Q., Feng, Y., Shan, Y., Jia, K.: GS-IR: 3D gaussian splatting for inverse rendering. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024)
65. Liu, L., Habermann, M., Rudnev, V., Sarkar, K., Gu, J., Theobalt, C.: Neural actor: Neural free-view synthesis of human actors with pose control. *ACM Transactions on Graphics* (2021)
66. Lokovic, T., Veach, E.: Deep shadow maps. In: *ACM SIGGRAPH Conference Proceedings* (2000)
67. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: SMPL: A skinned multi-person linear model. *ACM Transactions on Graphics* **34**(6) (2015)
68. Luiten, J., Kopanas, G., Leibe, B., Ramanan, D.: Dynamic 3d gaussians: Tracking by persistent dynamic view synthesis. In: *International Conference on 3D Vision (3DV)* (2024)
69. Luo, F., Pan, H., Yang, X., Jiang, B., Liu, F., Huang, T.: ShadowGS: Shadow-aware 3D gaussian splatting for satellite imagery. *arXiv preprint arXiv:2601.00939* (2026)
70. Mahmood, N., Ghorbani, N., Troje, N.F., Pons-Moll, G., Black, M.J.: AMASS: Archive of motion capture as surface shapes. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2019)
71. Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., Geiger, A.: Occupancy networks: Learning 3d reconstruction in function space. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019)
72. Mihajlovic, M., Prokudin, S., Tang, S., Maier, R., Bogo, F., Tung, T., Boyer, E.: SplatFields: Neural gaussian splats for sparse 3d and 4d reconstruction. In: *European Conference on Computer Vision (ECCV)* (2024)
73. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. In: *European Conference on Computer Vision (ECCV)* (2020)
74. Mir, A., Alldieck, T., Pons-Moll, G.: Learning to transfer texture from clothing images to 3d humans. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2020)
75. Mir, A., Guler, R.A., Tang, X., Wonka, P., Pons-Moll, G.: Ahoy! animatable humans under occlusion from youtube videos with gaussian splatting and video diffusion priors. In: *European Conference on Computer Vision (ECCV)* (2026)
76. Mir, A., Moreau, A., Dharmo, H., Zhang, Z., Pons-Moll, G., Pérez-Pellitero, E.: Gaspacho: Gaussian splatting for controllable humans and objects. *arXiv preprint arXiv:2503.09342* (2025)
77. Mir, A., Puig, X., Kanazawa, A., Pons-Moll, G.: Generating continual human motion in diverse 3d scenes. In: *International Conference on 3D Vision (3DV)* (2024)
78. Mir, A., Wang, J., Guler, R.A., Guo, C., Pons-Moll, G., Zhou, B.: Aha! animating human avatars in diverse scenes with gaussian splatting. *arXiv preprint arXiv:2511.09827* (2025)

79. Moenne-Loccoz, N., Mirzaei, A., Perel, O., de Lutio, R., Esturo, J.M., State, G., Fidler, S., Sharp, N., Gojcic, Z.: 3d gaussian ray tracing: Fast tracing of particle scenes. *ACM Transactions on Graphics* (2024)
80. Moon, G., Shiratori, T., Saito, S.: Expressive whole-body 3D gaussian avatar. In: *European Conference on Computer Vision (ECCV)* (2024)
81. Moreau, A., Song, J., Dhano, H., Shaw, R., Zhou, Y., Pérez-Pellitero, E.: Human gaussian splatting: Real-time rendering of animatable avatars. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024)
82. Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics* **41**(4) (2022)
83. Pandey, R., Orts-Escolano, S., LeGendre, C., Haene, C., Bouaziz, S., Rhemann, C., Debevec, P., Fanello, S.: Total relighting: Learning to relight portraits for background replacement. *ACM Transactions on Graphics* **40**(4) (2021)
84. Pang, H., Zhu, H., Kortylewski, A., Theobalt, C., Habermann, M.: Ash: Animatable gaussian splats for efficient and photoreal human rendering. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024)
85. Park, J.J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S.: DeepSDF: Learning continuous signed distance functions for shape representation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019)
86. Pavlakos, G., Choutas, V., Ghorbani, N., Bolkart, T., Osman, A.A.A., Tzionas, D., Black, M.J.: Expressive body capture: 3D hands, face, and body from a single image. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019)
87. Peng, S., Dong, J., Wang, Q., Zhang, S., Shuai, Q., Zhou, X., Bao, H.: Animatable neural radiance fields for modeling dynamic human bodies. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2021)
88. Punnakkal, A.R., Chandrasekaran, A., Athanasiou, N., Quiros-Ramirez, A., Black, M.J.: BABEL: Bodies, action and behavior with english labels. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021)
89. Qian, Z., Wang, S., Mihajlovic, M., Geiger, A., Tang, S.: 3dgs-avatar: Animatable avatars via deformable 3d gaussian splatting. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024)
90. Reeves, W.T., Salesin, D.H., Cook, R.L.: Rendering antialiased shadows with depth maps. In: *ACM SIGGRAPH Conference Proceedings* (1987)
91. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2022)
92. Saito, S., Schwartz, G., Simon, T., Li, J., Nam, G.: Relightable gaussian codec avatars. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024)
93. Saito, S., Simon, T., Saragih, J., Joo, H.: PIFuHD: Multi-level pixel-aligned implicit function for high-resolution 3D human digitization. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2020)
94. Sarkar, K., Buehler, M.C., Li, G., Wang, D., Vicini, D., Riviere, J., Zhang, Y., Orts-Escolano, S., Gotardo, P., Beeler, T., Meka, A.: LitNeRF: Intrinsic radiance

- decomposition for high-quality view synthesis and relighting of faces. In: ACM SIGGRAPH Asia Conference Papers (2023)
95. Savva, M., Chang, A.X., Hanrahan, P., Fisher, M., Nießner, M.: PiGraphs: Learning Interaction Snapshots from Observations. *ACM Transactions on Graphics* **35**(4) (2016)
 96. Shaw, R., Nazarczuk, M., Song, J., Moreau, A., Catley-Chandar, S., Dharmo, H., Perez-Pellitero, E.: SWinGS: Sliding windows for dynamic 3D gaussian splatting. In: European Conference on Computer Vision (ECCV) (2024)
 97. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. In: International Conference on Learning Representations (ICLR) (2021)
 98. Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B.: Score-based generative modeling through stochastic differential equations. In: International Conference on Learning Representations (ICLR) (2021)
 99. Starke, S., Zhang, H., Komura, T., Saito, J.: Neural state machine for character-scene interactions. *ACM Transactions on Graphics* **38**(6) (2019)
 100. Sun, T., Barron, J.T., Tsai, Y.T., Xu, Z., Yu, X., Fyffe, G., Rhemann, C., Busch, J., Debevec, P.E., Ramamoorthi, R.: Single image portrait relighting. *ACM Transactions on Graphics* **38**(4) (2019)
 101. SuperSplat: Supersplat. <https://superspl.at/>, accessed: 2025-09-17
 102. Taheri, O., Ghorbani, N., Black, M.J., Tzionas, D.: GRAB: A dataset of whole-body human grasping of objects. In: European Conference on Computer Vision (ECCV) (2020)
 103. Tancik, M., Weber, E., Ng, E., Li, R., Yi, B., Kerr, J., Wang, T., Kristoffersen, A., Austin, J., Salahi, K., Ahuja, A., McAllister, D., Kanazawa, A.: Nerfstudio: A modular framework for neural radiance field development. In: ACM SIGGRAPH Conference Proceedings (2023)
 104. Wang, S., Simon, T., Santesteban, I., Bagautdinov, T., Li, J., Agrawal, V., Prada, F., Yu, S.I., Nalbone, P., Gramlich, M., Lubachersky, R., Wu, C., Romero, J., Saragih, J., Zollhoefer, M., Geiger, A., Tang, S., Saito, S.: Relightable full-body gaussian codec avatars. In: ACM SIGGRAPH Conference Proceedings (2025)
 105. Wang, X., Girdhar, R., Gupta, A.: Binge watching: Scaling affordance learning from sitcoms. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2017)
 106. Weng, C.Y., Curless, B., Srinivasan, P.P., Barron, J.T., Kemelmacher-Shlizerman, I.: HumanNeRF: Free-viewpoint rendering of moving people from monocular video. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022)
 107. Westover, L.A.: Splatting: a parallel, feed-forward volume rendering algorithm. Ph.D. thesis, University of North Carolina at Chapel Hill (1992)
 108. Williams, L.: Casting curved shadows on curved surfaces. In: ACM SIGGRAPH Conference Proceedings (1978)
 109. Wu, G., Yi, T., Fang, J., Xie, L., Zhang, X., Wei, W., Liu, W., Tian, Q., Wang, X.: 4d gaussian splatting for real-time dynamic scene rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
 110. Wu, Q., Martinez Esturo, J., Mirzaei, A., Moenne-Loccoz, N., Gojcic, Z.: 3dgut: Enabling distorted cameras and secondary rays in gaussian splatting. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2025)

111. Wu, Z., Chen, J., Li, L., Wu, S., Zhu, Z., Xu, K., Oswald, M.R., Song, J.: 3D gaussian inverse rendering with approximated global illumination. *arXiv preprint arXiv:2504.01358* (2025)
112. Xie, X., Bhatnagar, B.L., Lenssen, J.E., Pons-Moll, G.: Template free reconstruction of human-object interaction with procedural interaction generation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024)
113. Xie, X., Bhatnagar, B.L., Pons-Moll, G.: Chore: Contact, human and object reconstruction from a single rgb image. In: *European Conference on Computer Vision (ECCV)* (2022)
114. Xie, X., Bhatnagar, B.L., Pons-Moll, G.: Visibility aware human-object interaction tracking from single rgb camera. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2023)
115. Xie, Y., Takikawa, T., Saito, S., Litany, O., Yan, S., Khan, N., Tombari, F., Tompkin, J., Sitzmann, V., Sridhar, S.: Neural fields in visual computing and beyond. *Computer Graphics Forum* (2022)
116. Xu, H., Alldieck, T., Sminchisescu, C.: H-nerf: Neural radiance fields for rendering and temporal reconstruction of humans in motion. In: *Advances in Neural Information Processing Systems (NeurIPS)* (2021)
117. Xu, Y., Chen, B., Li, Z., Zhang, H., Wang, L., Zheng, Z., Liu, Y.: Gaussian head avatar: Ultra high-fidelity head avatar via dynamic gaussians. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024)
118. Xue, L., Guo, C., Zheng, C., Wang, F., Jiang, T., Ho, H.I., Kaufmann, M., Song, J., Hilliges, O.: HSR: holistic 3d human-scene reconstruction from monocular videos. In: *European Conference on Computer Vision (ECCV)* (2024)
119. Ye, C., Nie, Y., Chang, J., Chen, Y., Zhi, Y., Han, X.: Gaustudio: A modular framework for 3d gaussian splatting and beyond. *arXiv preprint arXiv:2403.19632* (2024)
120. Yeshwanth, C., Liu, Y.C., Nießner, M., Dai, A.: Scannet++: A high-fidelity dataset of 3d indoor scenes. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2023)
121. Zhan, Y., Zhu, Q., Niu, M., Ma, M., Zhao, J., Zhong, Z., Sun, X., Qiao, Y., Zheng, Y.: Tomie: Towards explicit exoskeleton for the reconstruction of complicated 3d human avatars. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2025)
122. Zhan, Y., Shao, T., Wang, H., Yang, Y., Zhou, K.: Interactive rendering of relightable and animatable gaussian avatars. *IEEE Transactions on Visualization and Computer Graphics* **31**(10) (2025)
123. Zhang, J.Y., Pepose, S., Joo, H., Ramanan, D., Malik, J., Kanazawa, A.: Perceiving 3d human-object spatial arrangements from a single image in the wild. In: *European Conference on Computer Vision (ECCV)* (2020)
124. Zhang, J., Luo, H., Yang, H., Xu, X., Wu, Q., Shi, Y., Yu, J., Xu, L., Wang, J.: Neuraldome: A neural modeling pipeline on multi-view human-object interactions. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2023)
125. Zhang, L., Zhang, Q., Wu, M., Yu, J., Xu, L.: Neural video portrait relighting in real-time via consistency modeling. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2021)

126. Zhang, X., Bhatnagar, B.L., Starke, S., Guзов, V., Pons-Moll, G.: Couch: Towards controllable human-chair interactions. In: European Conference on Computer Vision (ECCV) (2022)
127. Zhang, Y., Tang, S.: The wanderings of odysseus in 3d scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022)
128. Zhao, K., Li, G., Tang, S.: DartControl: A diffusion-based autoregressive motion model for real-time text-driven motion control. In: International Conference on Learning Representations (ICLR) (2025)
129. Zhao, K., Zhang, Y., Wang, S., Beeler, T., Tang, S.: DIMOS: Synthesizing diverse human motions in 3d indoor scenes. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2023)
130. Zhao, Y., Wu, C., Huang, B., Zhi, Y., Zhao, C., Wang, J., Gao, S.: Surfel-based gaussian inverse rendering for fast and relightable dynamic human reconstruction from monocular video. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2025)
131. Zheng, I., Tang, G., Doronin, A., Teal, P.D., Zhang, F.L.: SSD-GS: Scattering and shadow decomposition for relightable 3D gaussian splatting. In: International Conference on Learning Representations (ICLR) (2026)
132. Zheng, Z., Zhao, X., Zhang, H., Liu, B., Liu, Y.: AvatarReX: Real-time expressive full-body avatars. *ACM Transactions on Graphics* **42**(4) (2023)
133. Zhu, H., Zhan, F., Theobalt, C., Habermann, M.: Trihuman: A real-time and controllable tri-plane representation for detailed human geometry and appearance synthesis. *ACM Transactions on Graphics* (2024)
134. Zielonka, W., Bagautdinov, T., Saito, S., Zollhöfer, M., Thies, J., Romero, J.: Drivable 3D gaussian avatars. In: International Conference on 3D Vision (3DV) (2025)