NEURAL INFORMATION PROCESSING SYSTEMS
NeurIPS 2026

SAGA: Gaussian Splatting-based Volumetric Video Compression with Sparse 4D Anchors

Ge Gao[1] · Siyue Teng[1] · Changqi Wang[1] · Fan Zhang[1] · Nantheera Anantrasirichai[1] · Jui Chiu Chiang[2] · Wen-Hsiao Peng[3] · David Bull[1]
[1]Visual Information Lab, University of Bristol, UK · [2]National Chung Cheng University, Taiwan · [3]National Yang Ming Chiao Tung University, Taiwan

University of BRISTOL
National Chung Cheng University
National Yang Ming Chiao Tung University
Project page

SAGA A dynamic scene as a **hierarchy of sparse anchors in 4D space-time**, entropy-coded with a **long-context model** over a fixed-size memory.

1.91 MB	−90.88%	−65.14%	85 FPS
for a 300-frame Neu3D scene: 129× smaller than 4D Scaffold-GS	PSNR BD-rate vs. GIFStream on Neu3D	PSNR BD-rate vs. OMG4 on Neu3D	decoding on Neu3D, one RTX 3090

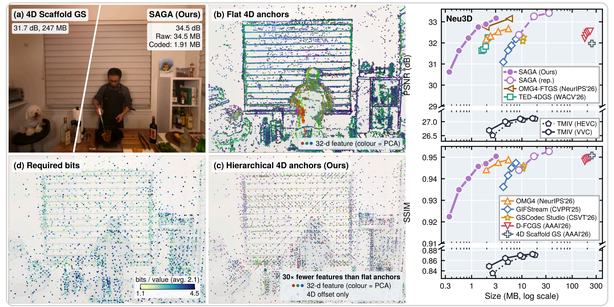
1 Motivation

Dynamic 3DGS is hard to compress: dense primitives, spatiotemporal redundancy. Existing designs have limits:

- **Frame-by-frame coding:** long prediction chains can complicate random access and accumulate drift.
- **Canonical + deformation:** a single reference can be affected by abrupt motion, disocclusion or topology changes.
- **Anchor scaffolds:** each primitive is typically conditioned only on its own anchor, limiting inter-anchor correlations.

SAGA: space-time anchors share features across spatial and temporal neighbours; no canonical scaffold to deform.

2 Flat vs. hierarchical 4D anchors



(a) 4D Scaffold-GS, 31.7 dB at 247 MB, vs. SAGA, 34.5 dB at 1.91 MB. **(b, c)** Anchors that store a feature. **(d)** Bits per value. **Right:** RD on Neu3D.

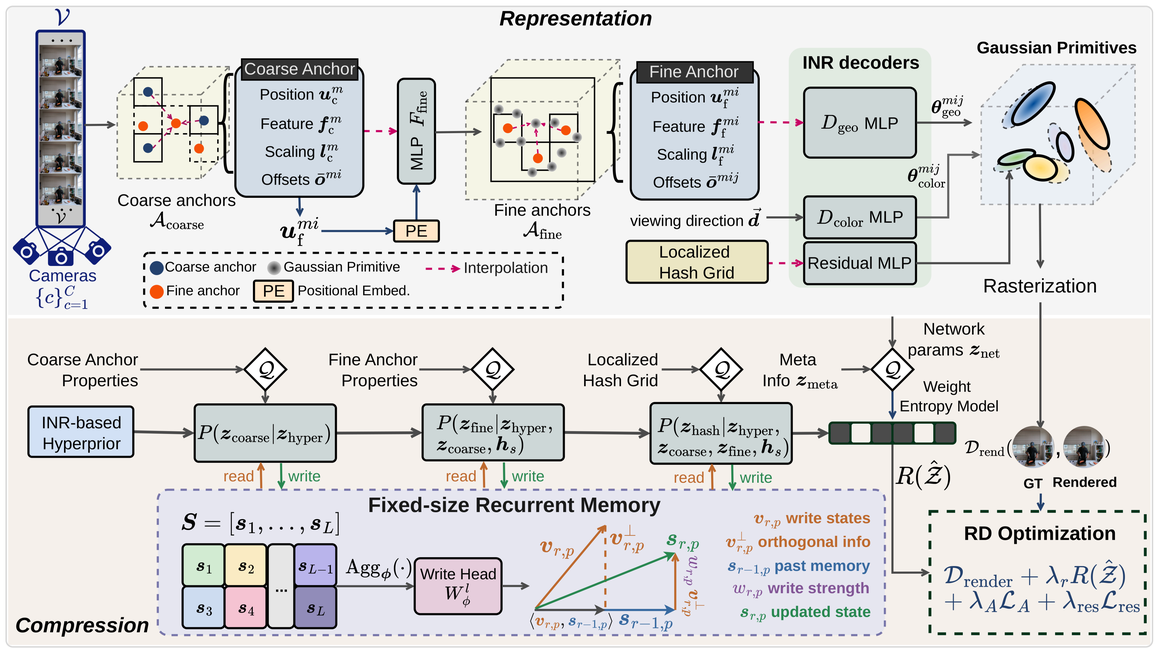
30×	2.1 bits	32-d
fewer feature-bearing anchors	per coded value, on average	features, coarse anchors only

3 Visual comparison



Neu3D (top), **MPEG MIV** (bottom). Labels: size | PSNR | SSIM. SAGA is the smallest with the highest PSNR in both examples.

4 Framework



Top: coarse anchors feed fine anchors (kernel interpolation), which generate Gaussians (INR decoders). **Bottom:** level-by-level entropy coding with a fixed-size recurrent memory.

5 Kernel interpolation

FEATURE AT A QUERY COORDINATE

$$\tilde{\boldsymbol{f}}_\ell(\boldsymbol{u}_q) = \text{MLP}_\ell\Big[\sum_{b\in\mathcal{N}(q)} \text{softmax}_b\big(-d_{qb}/T\big)\, \boldsymbol{f}_b \,;\, \text{PE}(\boldsymbol{u}_q)\Big]$$

$$d_{qb} = \big\|(\boldsymbol{u}_q - \boldsymbol{u}_b) \oslash (\boldsymbol{l}_q + \boldsymbol{l}_b)\big\|_\Lambda^2$$

- **Neighbours** $\mathcal{N}(q)$: the K closest of the M nearest 4D neighbours of the owner anchor a_q with $d_{qb} < \rho^2$, owner included; the decoder recomputes them.
- **Distance:** normalised by both supports; $\Lambda = \text{diag}(1, 1, 1, \lambda_t)$ weights time.
- **Two levels:** $\ell = c$: coarse → fine anchors, $\boldsymbol{l}_q = \boldsymbol{l}_{a_q}/4$; $\ell = f$: fine anchors → Gaussians, $\boldsymbol{l}_q = \boldsymbol{l}_{a_q}$.

6 Orthogonal memory update

L slots summarise what has been decoded and are read to predict the entropy parameters. After each step r, every slot p, with previous value $\boldsymbol{s}_{r-1,p}$, receives a candidate $\boldsymbol{v}_{r,p}$ and a gate $w_{r,p}$.

SLOT UPDATE

$$\boldsymbol{v}_{r,p}^{\perp} = \boldsymbol{v}_{r,p} - \frac{\langle \boldsymbol{v}_{r,p},\, \boldsymbol{s}_{r-1,p}\rangle}{\|\boldsymbol{s}_{r-1,p}\|_2^2 + \epsilon}\, \boldsymbol{s}_{r-1,p}$$

$$\boldsymbol{s}_{r,p} = \text{Norm}\big(\boldsymbol{s}_{r-1,p} + w_{r,p}\, \boldsymbol{v}_{r,p}^{\perp}\big)$$

The component parallel to the slot is removed before writing, limiting repeated writes.

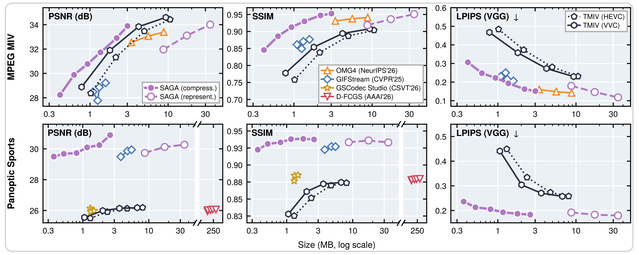
7 Compression efficiency

PSNR BD-rate vs. each codec (negative: SAGA saves bits); baselines need sufficient RD overlap. –: not reported.

Anchor codec	Neu3D	Panoptic Sports	MPEG MIV
TMIV-24.0 (HEVC, main)	–	–	−45.89%
TMIV-24.0 (VVEnc, main)	–	–	−18.41%
4DGC (CVPR'25)	−62.39%	–	−32.38%
GIFStream (CVPR'25)	−90.88%	−84.87%	−67.09%
TED-4DGS (WACV'26)	−64.30%	–	–
IC-4DGS (ICIP'26)	−22.62%	–	–
D-FCGS (AAAI'26)	−99.31%	–	–
OMG4 (NeurIPS'26)	−65.14%	–	–
SAGA (ours)	**0.00%**	**0.00%**	**0.00%**

Real-time decoding: **85 FPS** on Neu3D, **97 FPS** on MPEG MIV.

8 Rate–distortion curves



MPEG MIV (top), **Panoptic Sports** (bottom). Filled markers: SAGA after entropy coding; open: before coding.

9 Ablation on Neu3D

PSNR BD-rate increase vs. full SAGA; one component changed.

13.04%	11.24%	8.17%	7.22%
flat anchors, no hierarchy	no inter-anchor interpolation	non-orthogonal memory update	no MLP compression

No residual hash grid: 1.03%. With 256 memory slots (default 64): −15.66% BD-rate, but 50.3 FPS instead of 85.3.

10 Limitations

Per-scene optimisation may incur notable encoding time; fast motion, topology changes and sparse camera coverage may challenge reconstruction.

MyWorld Funded by UKRI MyWorld Strength in Places (SIPF00006/1) and Royal Society International Exchanges (IEC\R3\253264). Compute: University of Bristol ACRC, Isambard-AI.

Code: github.com/ge1gao/SAGA-NeurIPS26 · arXiv: 2609.33969 · ge1.gao@bristol.ac.uk