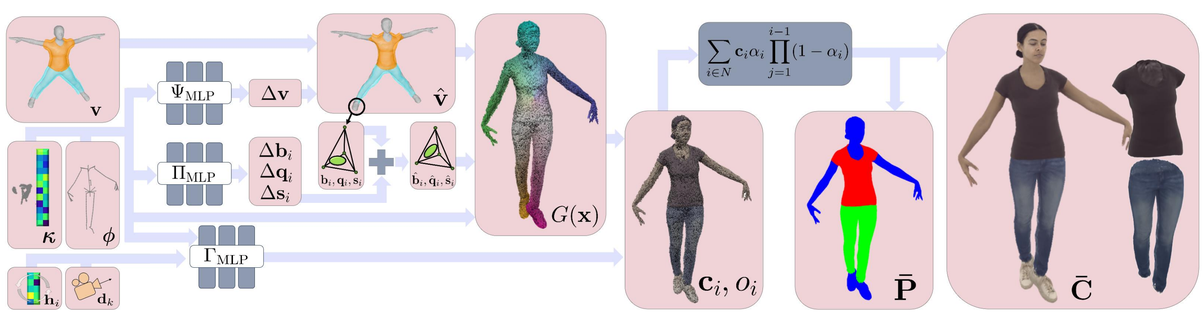
D3GA: Drivable Gaussian Avatars

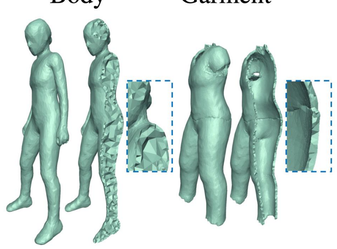
Wojciech Zielonka, Timur Bagautdinov, Shunsuke Saito,
Michael Zollhöfer, Justus Thies, Javier Romero

3DV 2025
International Conference on 3D Vision
Singapore
March 25-28, 2025

Abstract

- A multi-layered, controllable 3D human model using 3D Gaussian primitives embedded in tetrahedral cages.
- Unlike Linear Blend Skinning (LBS), cages naturally reorient and stretch 3D Gaussians via deformation gradients
- Efficient Deformation: Tetrahedron vertex offsets decouple driving poses from rendered primitives, improving optimization.
- Compositional Pipeline: Avatars are decomposed into layers (e.g., garments, hands, face)
- Flexible Conditioning: Different parts can be driven by keypoints (facial expressions) or joint-angle vectors (body & garments).





Given a multiview dataset of a single person, D3GA reconstructs a controllable avatar that can be driven using only pose parameters.

D3GA embeds 3D Gaussian primitives inside tetrahedrons, where the position of each primitive is represented as a barycentric interpolation of tetrahedron vertices. In this way, we perform simultaneous mesh tracking and reconstruction.

Qualitative comparisons show that D3GA models facial expressions and garments better than other SOTA approaches. Especially regions with loose garments like skirts or sweatpants.

D3GA enables motion transfer showing good generalizability while preserving each avatar's high-quality details.

Code

[1] Timur M. Bagautdinov, Chenglei Wu, Tomas Simon, Fabián Prada, Takaaki Shiratori, Shih-En Wei, Weipeng Xu, Yaser Sheikh, and Jason M. Saragih. **"Driving-Signal Aware Full-Body Avatars."**
[2] Stephen Lombardi, Tomas Simon, Gabriel Schwartz, Michael Zollhoefer, Yaser Sheikh, and Jason M. Saragih. **"Mixture of Volumetric Primitives for Efficient Neural Rendering."**