Project Page with **Video** results

Animal Avatar: Reconstructing Animatable-Animals from Casual Videos

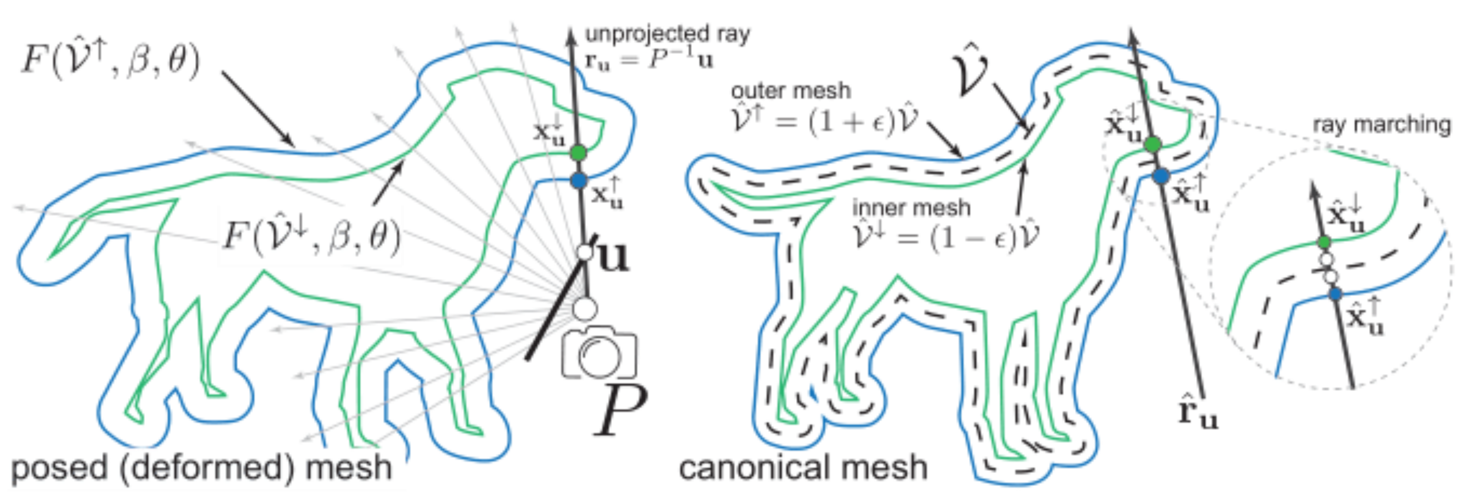
Remy Sabathier [1], [2], Niloy J. Mitra [2], David Novotny [1]

[1] Meta [2] University College London (UCL)

Improving supervision on challenging views

We **align** the **CSE** mesh template with the **SMAL** template to establish **dense** correspondences between surface points of the SMAL template and the corresponding image pixels.

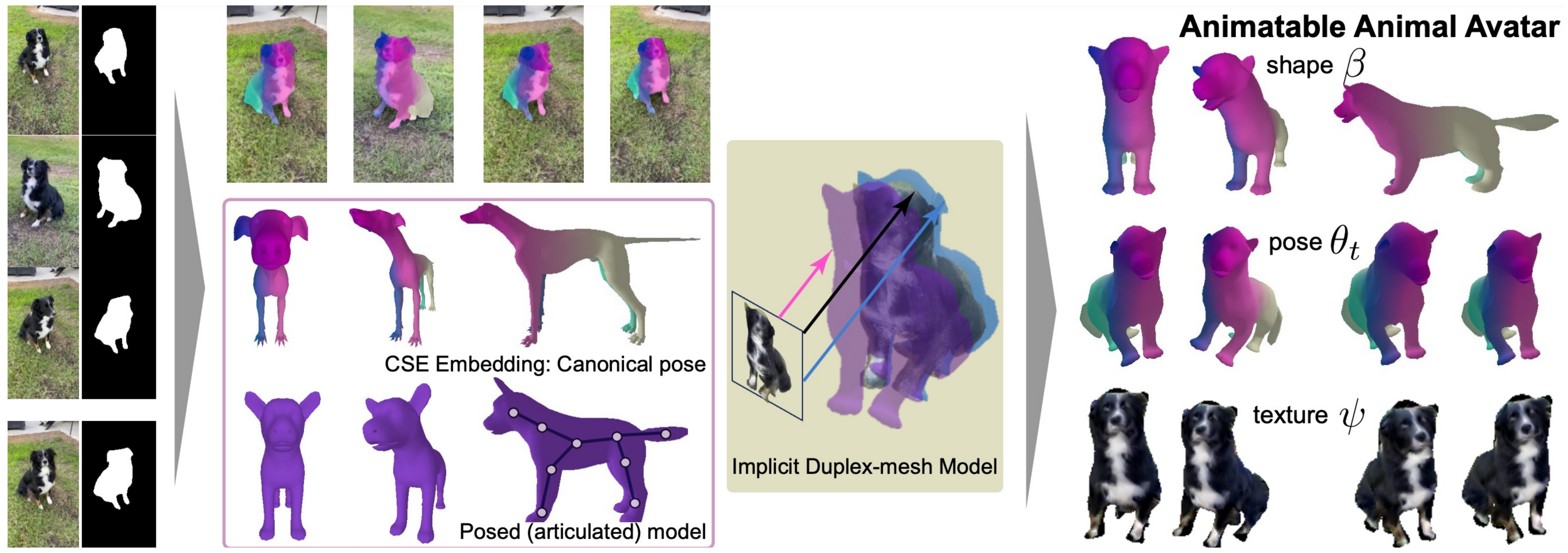
Duplex-Mesh Renderer



We render the color of a view-space ray $\boldsymbol{r_u}$ via EA raymarching over ray $\hat{\boldsymbol{r}}_{\boldsymbol{u}}$, obtained by transforming into the **canonical** space the intersections of $\boldsymbol{r_u}$ with the posed **duplex** (outer/inner) meshes $\{\hat{\mathcal{V}}^{\uparrow}, \hat{\mathcal{V}}^{\downarrow}\}$.

Since the texture is expressed in the canonical space, once optimized it can be rendered on any pose.

Problem Statement



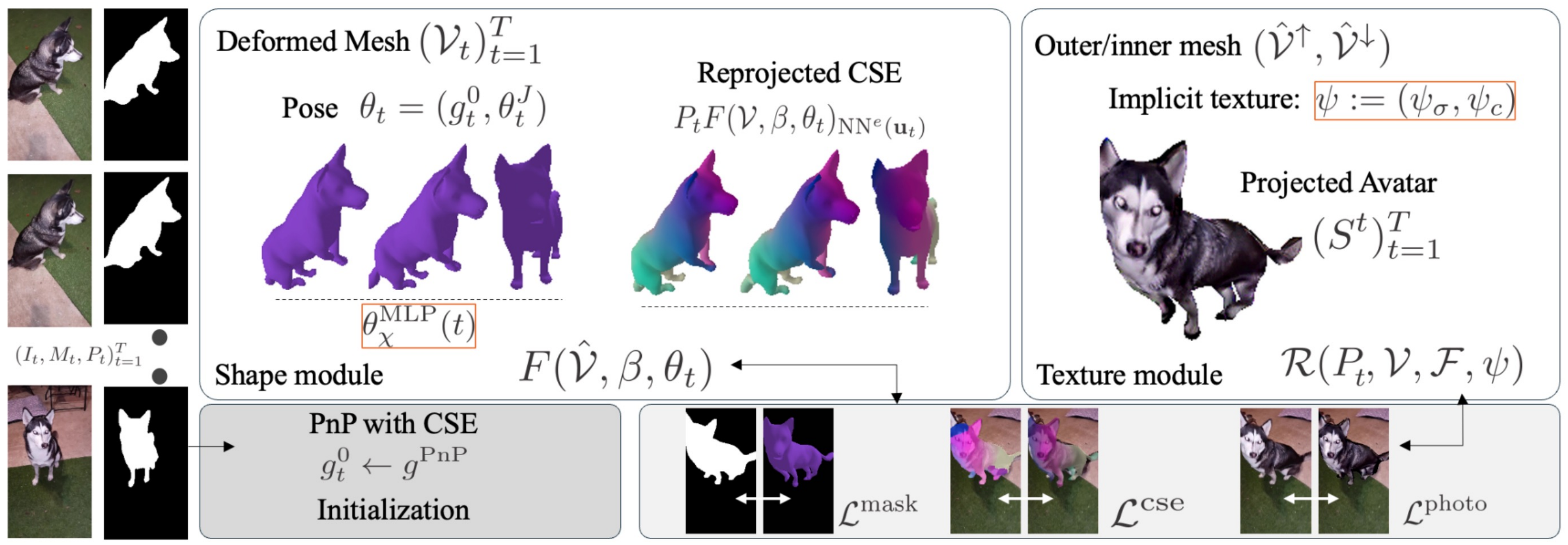
Given a **monocular video of a dog**, we propose a template-based method to reconstruct the **shape** β **motion** θ_t and **texture** ψ.

Traditional 3D-signals for animal reconstruction are **sparse** and fail when animals are captured from non-frontal views.

Our main contributions:

- Integrate **continuous surface embeddings** (**CSE**) with an **articulated mesh** (**SMAL**).
- Introduce an **implicit duplex neural radiance field** to enable texturing of the articulated mesh.

Method



We jointly optimize shape, pose and implicit texture through an **analysis-by-synthesis** approach, leveraging **mask**, **dense correspondence** and **photometric** signals.

Main Results

Dataset	CoP3D					APTv2
	IoU ↑	IoUw5 ↑	PSNR ↑	PSNRw5 ↑	LPIPS ↓	err_{track} ↓
BARC	0.75	0.47	×	×	×	0.047
BITE	0.81	0.59	×	×	×	0.047
RAC	0.76	0.52	21.86	17.51	0.164	0.093
Ours	**0.84**	**0.79**	**22.12**	**19.40**	**0.041**	**0.035**

References

- Rueegg, N., Zuffi, S., Schindler, K., & Black, M. J. (2022). **BARC**: Learning to Regress 3D Dog Shape from Images by Exploiting Breed Information
- Ruegg, N., Tripathi, S., Schindler, K., Black, M. J., & Zuffi, S. (2023). **BITE**: Beyond Priors for Improved Three-D Dog Pose Estimation.
- Yang, G., Wang, C., Reddy, N. D., & Ramanan, D. (2023). **RAC**: Reconstructing Animatable Categories from Videos.