

## BACKGROUND AND INTRODUCTION

In this work, we revisit 3D character animation by reformulating the classical skinning pipeline in latent space. Traditional methods rely on hand-crafted rigs, skeletal animation, and skinning operators such as Linear Blend Skinning (LBS), which, despite their efficiency, introduce artifacts and require manual setup. More advanced deformations reduce these artifacts but increase complexity and still depend on mesh topology. Recent learning-based techniques show that animation data is highly redundant and compressible, yet typically remain tied to specific rigs or topologies and treat motion, shape, and deformation separately, so compression, synthesis, and transfer are rarely addressed in a unified framework.

We introduce Latent-Space Skinning (LSS), a neural framework that learns a compact, motion-conditioned latent representation of mesh deformation. Instead of predicting skinning weights, LSS encodes **skeletal motion** and **rest-pose geometry** into a shared latent space and decodes directly to per-vertex deformations over time, remaining compatible with skeleton-based pipelines while reducing reliance on explicit skinning. This leverages temporal redundancy for high compression, corrects common skinning artifacts, and naturally supports interpolation, blending, and motion transfer.

## RELATED WORK

Positioned within modern motion modeling (e.g., BiMotion [WYC\*26]), 4D generative models, deep skinning prediction [MTF23], and latent-space transfer such as SMF [MDM25]), our method provides a simple, topology-dependent baseline that unifies deformation, reconstruction, and motion reuse in a single latent formulation.

## OVERVIEW

**Latent-Space Skinning (LSS)** is a neural framework that models mesh animation as a learned mapping from skeletal motion to vertex deformations in a compact latent space. In contrast to classical skinning methods, which explicitly compute vertex positions from bone transformations and skinning weights, **LSS** directly learns the relationship between skeletal motion, rest-pose geometry, and the resulting animated mesh. Consequently, traditional rigging, skinning weights, and deformation modules are replaced by a latent representation and a decoder that reconstructs the animation sequence. This formulation enables the network to capture both skeletal motion and complex non-linear deformation effects within a unified representation.

## RESULTS

## Compression and reconstruction performance (results in table below).

Comparison between LSS and Fusible Skinning (FS) across multiple characters. We report distortion percentage, root-mean-square error, mean maximum vertex error and compression ratio (percentage reduction relative to the original skeletal representation). Positive values in the improvement columns indicate that LSS improves over FS.

LSS consistently achieves significantly higher compression, while maintaining comparable or improved reconstruction quality in most cases.

Character	Vertices	Orig. Bones	FS Bones	FS				LSS				Improvement of LSS over FS (%)			
				DisPer	ERMS	Max	Comp.	DisPer	ERMS	Max	Comp.	DisPer	ERMS	Max	Comp.
Spiderman	5336	78	14	0.0230	1.1400	0.0320	82.1	0.0151	1.0710	0.0416	95.7	34.3	6.1	-30.0	16.7
Dragon	19895	182	41	3.1000	5.2000	0.5900	77.5	0.0555	4.5255	0.8143	98.2	98.2	13.0	-38.0	26.7
Sailfish	2344	16	7	2.1000	0.9100	0.0129	56.3	0.0042	0.2831	0.0148	79.2	99.8	68.9	-14.7	40.7
Alpha_surface	27850	82	22	1.2800	4.7900	1.2800	73.2	0.0958	5.8230	1.3690	95.9	92.5	-21.6	-7.0	31.1

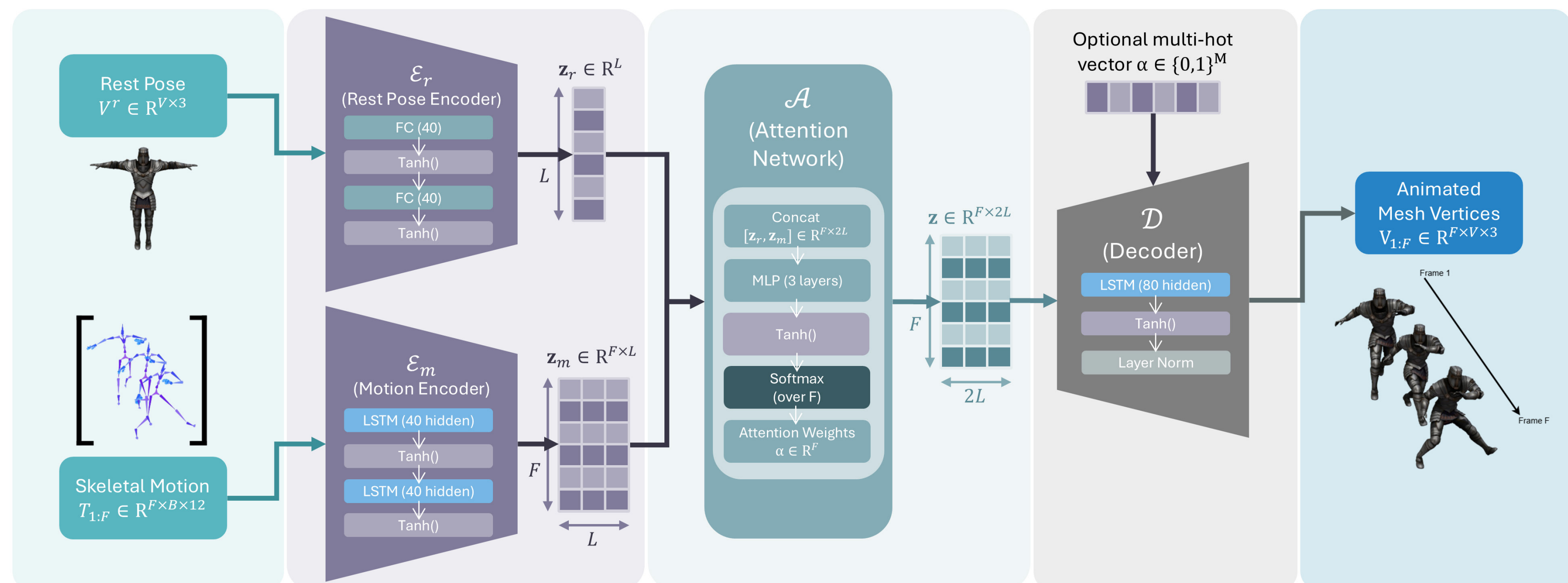
## Animation reconstruction and blending (explaining figure on the right).

Top: Dragon animation reconstructed by our model, preserving motion with high fidelity.

Bottom: Walk and turn motion bending, generating a smooth transition from walking (light blue) to turning (purple).

## METHODOLOGY

Modeling mesh animation requires capturing both temporal motion dynamics and shape-dependent deformations, as character motion evolves continuously over time while remaining strongly influenced by morphology. To address these challenges, we employ an encoder-decoder sequence architecture based on LSTMs with latent-space fusion. An overview of the proposed LSS architecture is shown in the figure below.

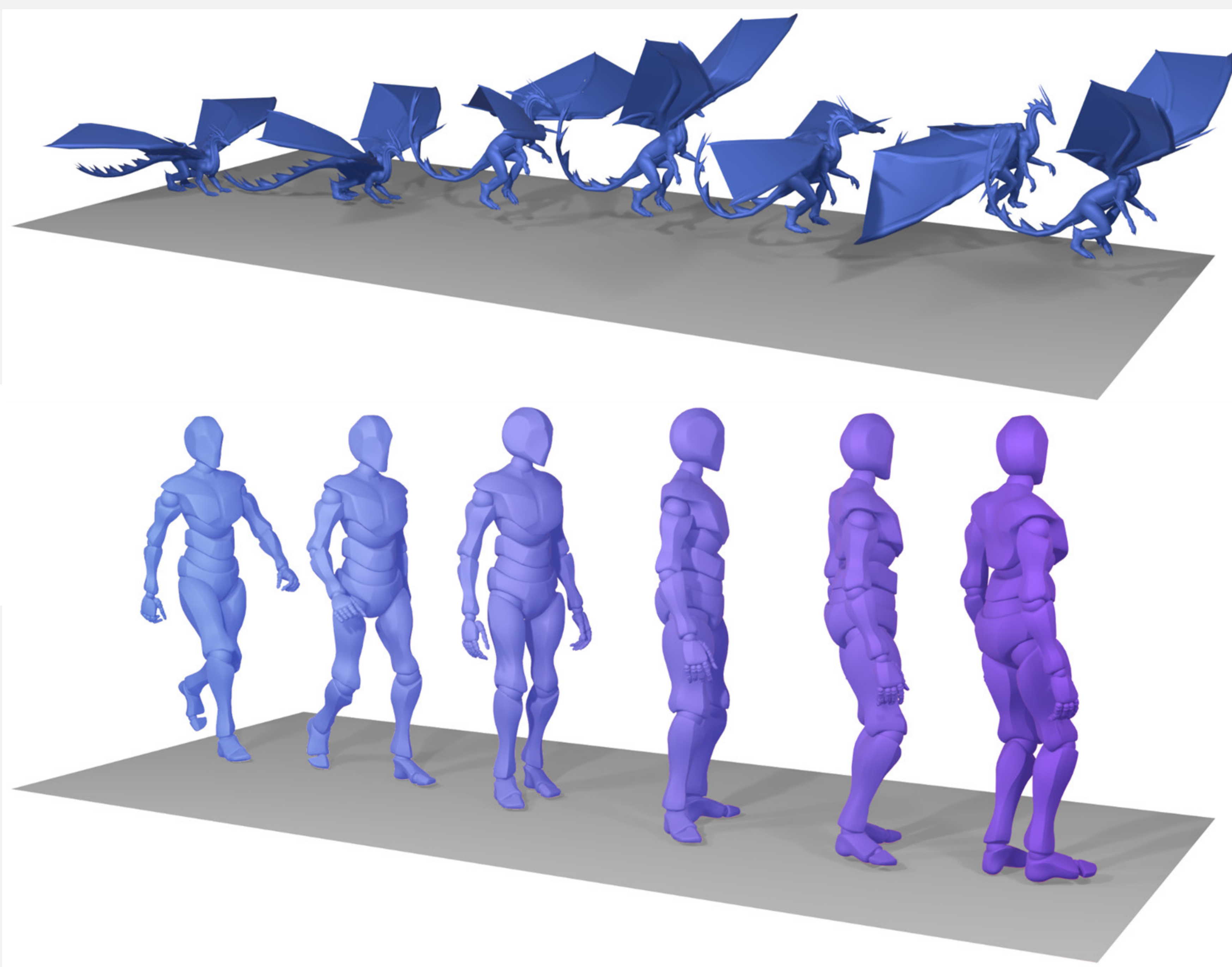


This formulation captures both skeletal motion and non-linear deformation effects within a unified latent representation. The resulting latent space enables compact motion encoding for animation compression, latent-space interpolation for animation synthesis, and the transfer of motion across characters through the recombination of motion and shape representations.

The network is trained to reconstruct the ground-truth animated mesh sequence  $V_{1:F}^{gt}$  from the rest-pose mesh  $V^r$  and the skeletal motion sequence  $T_1$ . Optimization is performed using AdamW with a learning rate of  $10^{-3}$  and weight decay of  $10^{-4}$  for 6,000 epochs on a single GPU. Models are trained independently for each character using from 7 to 20 animation sequences.

The final training objective combines reconstruction accuracy with geometric, volumetric, and temporal regularization terms. Specifically,  $\mathcal{L}_G$  preserves local surface curvature,  $\mathcal{L}_{vol}$  enforces segment volume preservation, and  $\mathcal{L}_t$  promotes temporal smoothness by minimizing per-vertex acceleration:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{MSE} + \lambda_2 \mathcal{L}_C + \lambda_3 \mathcal{L}_G + \lambda_4 \mathcal{L}_{vol}$$



## AFFILIATIONS



## REFERENCES

- [MDM25] MURALIKRISHNAN S., DUTT N. S., MITRA N. J.: Smf: Template-free and rig-free animation transfer using kinetic codes. ACM Trans. Graph. 44, 6 (Dec. 2025).  
 [MTF23] MOUTAFIDOU A., TOULATZIS V., FUDOS I.: Deep fusible skinning of animation sequences. Vis. Comput. 40, 8 (Nov. 2023), 5695–5715.  
 [WYC\*26] WANG M., YAN Q., CAO Z., LI Y., MAC AODHA O., CORSO J. J., VAXMAN A.: Bimotion: B-spline motion for text-guided dynamic 3d character generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2026), CVPR'26

## LINKS

## Source code:

- <https://github.com/gdrongoulas/LatentSpaceSkinning>

## Computer Graphics Group:

- <https://cgrg.eu/>

## Graphics &amp; Extended Reality Lab:

- <https://graphics.cs.ucy.ac.cy/>

