ACM SIGGRAPH / Eurographics Symposium on Computer Animation 2026
D. Levin and M. Chu
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 45 (2026), Number 8

LORAMI: Low-Rank Adaptation for Multi-Identity Physics-Based Face Rigs

D. Corigliano[1], D. Peter[1,2], B. Solenthaler[1], and B. Thomaszewski[1]

[1]ETH Zurich, Switzerland
[2]Align Technology, Switzerland

1. Anatomical Head Data

1.1. Anatomical Components

Our simulation-ready head models closely follow the definition of [CPH*25]. Each identity is represented using a multi-layered anatomical model compatible with physics-based simulation, inclusive of facial geometry and internal anatomy, such as bones and teeth. The anatomical head models are driven using a user-defined rig, whose controls $\mathbf{q} \in \mathbb{R}^m$ induce transformations of the anatomical components, while respecting biomechanical limits. In our implementation, we account for $m = 14$ controls, which we observe to produce a large number of expressions and facial changes.

Teeth Each identity is equipped with a personalized set of teeth, with unique locations and orientations.

Skull and Mandible Our anatomical head model contains three main personalized rigid components, namely the skull $\mathcal{S}$, the mandible $\mathcal{M}$ and the teeth $(\mathcal{T}_{upper}, \mathcal{T}_{lower})$. Facial expression and articulation require the motion of the jaw. We achieve it by providing $\mathbf{q}_m \in \mathbb{R}^4$ degrees of freedom for the mandible, corresponding to three rotations and one translation in the anatomical limits. The transformations of the mandible are propagated to the lower teeth arch $\mathcal{T}_{lower}$, whereas the cranium does not move.

Orthognathic Controls The mandible and skull meshes are split along pre-defined cut planes, selected according to reference medical procedures. The lower skull (maxilla) and the frontal mandible are then associated with additional degrees of freedom $\mathbf{q}_o \in \mathbb{R}^3$, that represent various bone translations in the horizontal and vertical directions. The range of $\mathbf{q}_o$ is bounded to guarantee anatomical feasibility.

Soft Tissue The layer between the bone and the facial surfaces is tetrahedralized, providing a deformable soft tissue $\mathcal{ST}$ for simulation. The soft tissue nodes are denoted as $x_s \in \mathbb{R}^{n_v}$, where $n_v = 3|\mathcal{V}|$ and $\mathcal{V}$ is the set of tetrahedral mesh vertices. The nodes deform elastically in response to a change of boundary conditions corresponding to $(\mathbf{q}_o, \mathbf{q}_m)$.

Expressions Realistic facial expressions are incorporated through a dual-representation approach that utilizes expression blendshapes $\{\mathcal{B}_i\}$ for surface-level deformation and soft-tissue expression operators $\{\mathbf{F}_{\mathrm{E}_i}\}$ for mechanical simulation. For the six provided expressions, we add $q_e \in \mathbb{R}^6$ degrees of freedom to the model. These operators are derived by aligning the template soft tissue $\mathcal{ST}_i$ to target surface scans and extracting local actuations via polar decomposition of the associated deformation gradients [YZC*24, CPH*25]. Those expression pairs are then adapted per identity, as described in Section 2.6.

Rest Configuration In the neutral expression, lip attachment is guaranteed by volumetric stresses in the mouth area. Our rest configuration $\mathcal{ST}_{Rest}$ accounts for this phenomenon by including a slight overlap between the lower and the upper lips. In our implementation, the rest pose is represented as an additional expression, which is enabled when all the other expressions are deactivated, effectively representing the absence of active muscular stresses, and adding one extra degree of freedom.

Neutral Configuration The neutral configuration $\mathcal{ST}_{Neutral}$ represents the neutral expression, where all the controls are inactive. In such a pose, the lips are in tight contact, but not intersecting. It provides the starting configuration for subsequent physically-based simulation.

Skin A high-resolution textured skin layer is embedded in the barycentric coordinates of the soft tissue. This layer allows for seamless rendering of the head models using available ray tracers.

1.2. Identity Dataset

For each subject, fitting of the anatomical head model begins with the neutral facial surface $\mathcal{F}_{Neutral}$ obtained via photogrammetry, the skull $\mathcal{S}$ and mandible $\mathcal{M}$ obtained through a PCA fit, and the teeth $(\mathcal{T}_{upper}, \mathcal{T}_{lower})$ acquired using either an intra-oral scan or a PCA fit. The rest configuration for the face, $\mathcal{F}_{Rest}$, is derived through an optimization of the inner mouth surface. After scaling all models to a common unit, a shared template topology for the soft tissue $\mathcal{ST}_{Template}$ is registered to every identity in two sequential steps: first, the closest rigid transformation between the template surface and the subject surface $\mathcal{F}_{Rest}$ is estimated and applied to the template tetrahedral mesh; second, a nonrigid fitting is performed by solving a finite-element optimization with point constraints at the surface vertices of $\mathcal{F}_{Rest}$, deforming the aligned template to match the subject's geometry and produce the soft tissue mesh $\mathcal{ST}_{Rest}$. Subsequently, another fitting procedure aligns $\mathcal{ST}_{Rest}$ to $\mathcal{F}_{Neutral}$ to generate $\mathcal{ST}_{Neutral}$. Finally, all subjects are brought into a common reference frame by aligning the skull to the Frankfort plane via a rigid roto-translation. For each identity, we define an identity vector $\phi = [\mathcal{ST}_{Rest}, \mathcal{ST}_{Neutral}, \mathcal{S}, \mathcal{M}, \mathcal{T}_{upper}, \mathcal{T}_{lower}]$, composed of the linearized meshes for the anatomical components.

1.3. Statistical Shape Model

A PCA model is fitted on the collection of identity vectors $\{\phi_i\}$. Each training identity is then encoded as a coefficient vector by projecting onto this subspace. At inference time, a new identity geometry is reconstructed by inverse-transforming a coefficient vector through the PCA model, producing the full set of registered meshes (soft tissue, skull, mandible, teeth) for that identity. Each identity is represented in low-dimensional PCA space by a vector $\mathbf{p} \in \mathbb{R}^{\mathbf{40}}$.

2. Computational Face Model

After selecting the identity corresponding to an encoding $\mathbf{p}$, the facial configuration is determined by a set of rig control values $\mathbf{q}$, which dictate the boundary conditions of the physical system. We define the resulting facial deformation as the minimizer of a total potential energy functional that encompasses soft-tissue elasticity, bone-tissue attachments, and contact interactions, following the formulation in [CPH*25]. In the next sections, we describe the potentials, omitting the identity codes $\mathbf{p}$ for simplicity.

2.1. Soft Tissue Elasticity

Our tetrahedralized soft tissue mesh $\mathcal{ST}$ naturally lends itself to simulation using linear tetrahedral elements. Employing the approach of [SNF05, KE22], the total elastic energy is expressed as the sum of per-element energies $E_{\text{Elastic}}(\theta, \bar{\mathbf{x}}, \mathbf{x}) = \sum_i E^i_{\text{Elastic}}(\theta, \bar{\mathbf{x}}, \mathbf{x})$, with $\bar{\mathbf{x}} \in \mathbb{R}^{n\times 3}$ and $\mathbf{x} \in \mathbb{R}^{n\times 3}$ being the vertex positions in the rest and deformed states, and n being the number of nodes. The elastic energy is parameterized in terms of material properties θ. Following the approach of [CPH*25], we pick the Stable Neo-Hookean material model for our simulations. The energy density is defined as:

$$E_{NH} = \frac{\mu}{2}(I_C - 3) + \frac{\lambda}{2}(J - \alpha)^2 - \frac{\mu}{2}\log(I_C + 1), \tag{1}$$

where $\alpha = 1 + \frac{\mu}{\lambda} - \frac{\mu}{4\lambda}$, $\theta = (\lambda, \mu)$ are the Lamé parameters, $I_C = \text{tr}(\mathbf{F}^T\mathbf{F})$, $J = \det(\mathbf{F})$, and $\mathbf{F}$ is the deformation gradient.

2.2. Attachment Constraints

Rigid attachments between the soft tissue and bone are implemented as hard constraints, which effectively reduce the total degrees of freedom of the soft tissue mesh $\mathcal{ST}$ during simulation. The interfaces between the rigid skeletal structures and the soft tissue are defined via a barycentric embedding of the soft tissue into the bone. Then, attachment is strictly enforced by projecting the tissue nodes located on the interface directly onto their corresponding coordinates on the bone geometry, ensuring they transform rigidly with the underlying cranium and mandible.

2.3. Contact Constraints

The lower and upper lips produce both self-collisions and contacts with the teeth. To handle such interactions, we follow the approach of [CPH*25]. We represent the collider geometries $\mathbf{y}(\mathbf{q})$ as Implicit Moving Least Squares (IMLS) surfaces [DLCT24]. The IMLS approximation consists in representing the colliders as signed distance functions:

$$d_{\text{IMLS}}(\mathbf{x}) = \frac{\sum_i \mathbf{n}i^T(\mathbf{x} - \mathbf{x}i) \cdot \gamma_i(\mathbf{x})}{\sum_i \gamma_i(\mathbf{x})}, \tag{2}$$

where $\mathbf{x}_i$ and $\mathbf{n}_i$ denote the surface vertices and their corresponding normals. We use the same radial basis function as the reference implementation: $\gamma_i(\mathbf{x}) = (1 - |\mathbf{x} - \mathbf{x_i}|^2/h^2)^4$ if $|\mathbf{x} - \mathbf{x_i}| \leq 0$ and 0 otherwise. For the contact barrier, we opt for a smoothly clamped cubic penalty $\Phi(d, d_0)$. The total contact potential is the sum over all the contact nodes:

$$E_{contact}(\mathbf{x}, \mathbf{q}) = w_c \sum_i \sum_j \Phi\big(d_{IMLS}(\mathbf{x}_i, \mathbf{y}^j(\mathbf{q})), d_0\big), \tag{3}$$

where w_c is the contact penalty weight and d_0 represents the activation distance at which the barrier potential engages.

2.4. Total Mechanical energy

The potentials are all combined in the mechanical energy of the system. Each pair of controls-identity $\mathbf{h} = [\mathbf{q}, \mathbf{p}]$ directly maps to an equilibrium configuration for the system, as:

$$\mathbf{x}(\mathbf{h}) = \arg\min_{\hat{\mathbf{x}}} \left(E_{Elastic}(\theta, \bar{\mathbf{x}}(\mathbf{h}), \tilde{\mathbf{x}}(\mathbf{h})) + E_{Contact}(\tilde{\mathbf{x}}(\mathbf{h}), \mathbf{h})\right). \tag{4}$$

The forward simulation maps the the soft tissue of the identity $\mathbf{p}$ to the equilibrium configuration corresponding to the controls $\mathbf{q}$. The statistical model and the forward simulation together form an high-dimensional map that produces equilibrium configurations for any representable identity $\mathbf{x} : \mathbf{h} \rightarrow \mathbf{x}(\mathbf{h})$

2.5. Expression

Expressions are central to our framework. During simulation, when a set of expressions $\mathcal{E}$ is activated by the controls $\mathbf{q}$, we modify the soft tissue's rest state using the formulation of [CPH*25]:

$$\mathbf{F}_{Elastic}(\mathbf{q}) = \mathbf{F}_{Geometry} \cdot \mathbf{F}_{\mathcal{E}}^{-1}(\mathbf{q}) \tag{5}$$

Here, $\mathbf{F}_{\mathcal{E}}(\mathbf{q})$ represents the expression operator that maps the rest configuration to the target shape defined by controls $\mathbf{q}$, with each individual expression E_i activated proportionally to its scalar control q_i. The terms $\mathbf{F}_{\text{Elastic}}$ and $\mathbf{F}_{\text{Geometry}}$ denote the elastic and geometric deformation gradients, respectively. $\mathbf{F}_{\text{Elastic}}$ ensures that the simulation yields deformations that are physically plausible and consistent with the target defined by $\mathbf{q}$.

2.6. Expression Personalization

Given the template identity $\mathbf{p}_t$ and a target identity $\mathbf{p}_k$ with soft-tissue vertices $\mathbf{x}^t$ and $\mathbf{x}^k$, we define the per-element local frames $\mathbf{E}_t$ and $\mathbf{E}_k$.

Per-Identity expression actuations To adapt actuations, we follow the approach of [YZC*24]. We extract the identity-specific rotation $\mathbf{R}_{t\to k}$ via polar decomposition on the deformation gradient $\mathbf{F}_{t\to k} = \mathbf{E}_k \cdot \mathbf{E}_t^{-1}$. The target actuation becomes $\tilde{\mathbf{F}}^k_{\mathcal{E}} = \mathbf{R}_{t\to k} \cdot \mathbf{F}^t_{\mathcal{E}} \cdot \mathbf{R}^T_{t\to k}$. This ensures the actuation remains consistent with the target geometry.

Per-Identity blendshapes Blendshapes are adapted via a linear deformation-gradient transfer procedure. For each tetrahedron, the nodes are deformed to the target identity frame as $\mathcal{B}^k = \mathbf{F}_{t\to k} \cdot \mathcal{B}^t$. The final per-vertex displacement at each vertex of the target identity is the average of all contributions from incident tetrahedra.

Expression Synthesis. At runtime, the control vector $\mathbf{q}$ drives the linear combination of both mechanical and geometric quantities. In terms of expressions, the blendshapes and deformation gradients are:

$$\mathbf{F}_{\mathrm{E}}(\mathbf{q}) = \sum_{i\in I_{\mathcal{E}}} q_i \mathbf{F}^i_{\mathrm{E}}, \qquad \mathcal{B}(\mathbf{q}) = \sum_{i\in I_{\mathcal{E}}} q_i \mathcal{B}^i \tag{6}$$

where $I_{\mathcal{E}}$ are the indices for the expression controls. To personalize the rest configuration, an identity-specific rest operator $\mathbf{F}_{\mathrm{rest}}$ is blended based on the maximum activation. This ensures the rest correction vanishes at full activation ($q_{\max} = 1$) and applies fully at neutral ($q_{\max} = 0$). The expressions that do not interact with the mouth area are modeled exclusively as blendshapes, as they do not affect the simulation.

3. Reconstructions with PCA

We show the accuracy of our statistical model in the reconstruction of two ground-truth alignments that are not part of the training dataset. As shown in the first row of Figure 1, our PCA model achieves a mean reconstruction error of 1 mm and a maximum reconstruction error of 6.6 mm on the first identity. The second row illustrates that our model achieves 2 mm and 15 mm on the other evaluation identity. We observe that the model produces a satisfactory fit on the face and the lips, while error mostly concentrates under the neck area, of scarce interest for our facial simulation. We remark that the accuracy of the statistical model is independent of the accuracy of the neural backing of physics-principled deformations. The PCA model is uniquely responsible for the quality of the anatomical head models. Instead, the quality of the physics is entirely dependent on the architecture of LORAMI.

4. Illustration of Architectures

Figure 2 details the architectures evaluated in the main paper. We first illustrate the MLP backbone of NeuRiPhy in (a), followed by a breakdown of the hidden layer implementations for each method (b–e). The figures describe the possible design choices for identity conditioning discussed in the main paper. For clarity of exposition, the hidden layer structure of the proposed LORAMI network is delineated in subpanel (d), allowing for a direct comparison with alternative implementations.

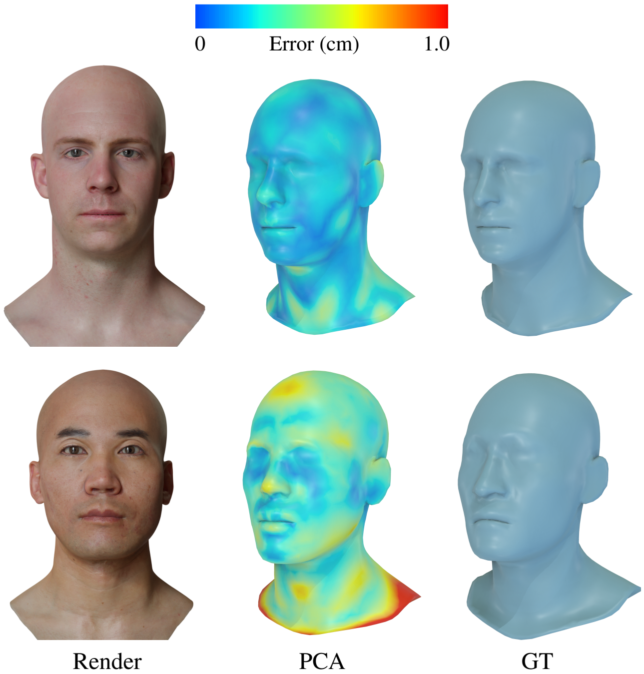

Figure 1: *Reconstruction accuracy of our statistical model, evaluated in terms of distance from the ground-truth alignment for two validation identities.*

References

[CPH*25] Corigliano D., Peter D., Huber N. B., Thomaszewski B., Solenthaler B.: Neuriphy: Neural baking of physics-based deformations for facial rigs. *Proc. ACM Comput. Graph. Interact. Tech. 8*, 4 (Aug. 2025). URL: https://doi.org/10.1145/3747866, doi:10.1145/3747866. 1, 2

[DLCT24] Du Y., Li Y., Coros S., Thomaszewski B.: Robust and artefact-free deformable contact with smooth surface representations. *Computer Graphics Forum 43*, 8 (2024), e15187. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.15187, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.15187, doi:https://doi.org/10.1111/cgf.15187. 2

[KE22] Kim T., Eberle D.: Dynamic deformables: Implementation and production practicalities (now with code!). In *ACM SIGGRAPH 2022 Courses* (New York, NY, USA, 2022), SIGGRAPH '22, Association for Computing Machinery. URL: https://doi.org/10.1145/3532720.3535628, doi:10.1145/3532720.3535628. 2

[SNF05] Sifakis E., Neverov I., Fedkiw R.: Automatic determination of facial muscle activations from sparse motion capture marker data. In *ACM SIGGRAPH 2005 Papers* (New York, NY, USA, 2005), SIGGRAPH '05, Association for Computing Machinery, p. 417–425. URL: https://doi.org/10.1145/1186822.1073208, doi:10.1145/1186822.1073208. 2

[YZC*24] Yang L., Zoss G., Chandran P., Gross M., Solenthaler B., Sifakis E., Bradley D.: Learning a generalized physical face model from data. *ACM Trans. Graph. 43*, 4 (July 2024). URL: https://doi.org/10.1145/3658189, doi:10.1145/3658189. 1, 3

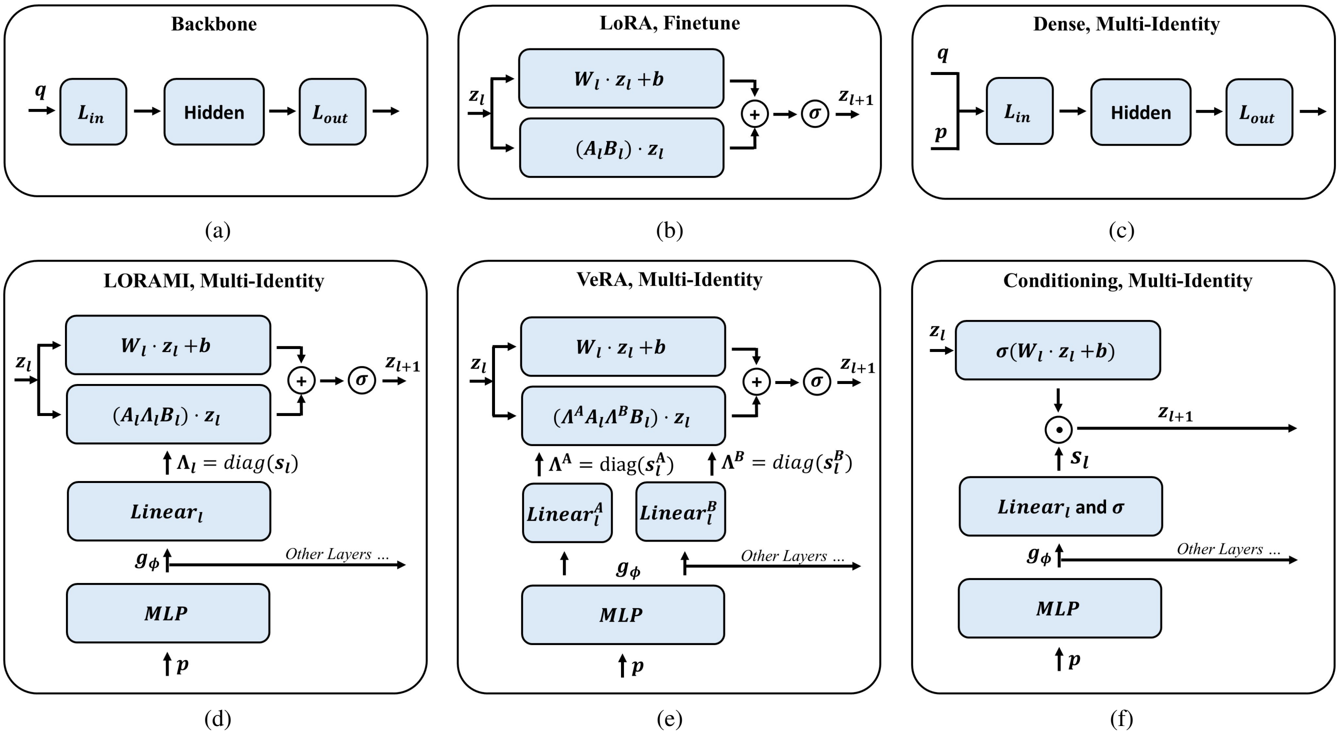

Figure 2: *Overview of multi-identity and single-identity architectures. Subpanel (a) illustrates the NeuRiPhy backbone, implemented as a simple MLP. Subpanels (b–e) detail the hidden layer structures for each method: (b) LoRA (for single-identity finetuning), where low-rank adaptations are added to frozen network parameters* $\mathbf{W}_l$ *and* $\mathbf{b}_l$ *post-training; (c) Dense multi-identity, which concatenates identity and control vectors at the input; (d) LORAMI and (e) VeRA, which learn low-rank vectors jointly with the layer parameters* $\mathbf{W}_l$ *and* $\mathbf{b}_l$*. While VeRA employs two diagonal matrices for modulation, LORAMI utilizes a single diagonal matrix. The conditioning architecture modulates activations at each layer's output. Experimental results demonstrate that LORAMI achieves the optimal tradeoff between accuracy and parameter efficiency.*