Neural Implicit Morphing of Face Images

Guilherme Schardong, Tiago Novello, Hallison Paz, Iurii Medvedev, Vinícius da Silva, Luiz Velho, Nuno Gonçalves

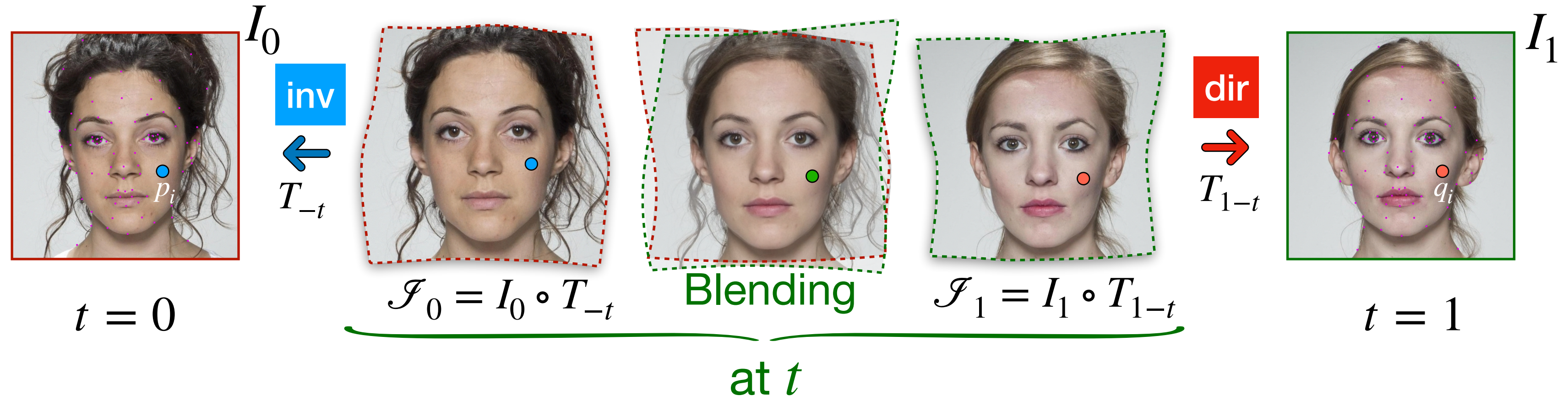
paper&code

CVPR JUNE 17-21, 2024

We investigate the use of **smooth** neural networks for morphing face images regularized by the **thin-plate** energy. For this, we model the time as a parameter and **disentangle** deformation from blending.

Neural Morphing Primer

A morphing between **faces** consists of a **warping** T of their domains for feature alignment and **blending** of the resulting **warped faces**.

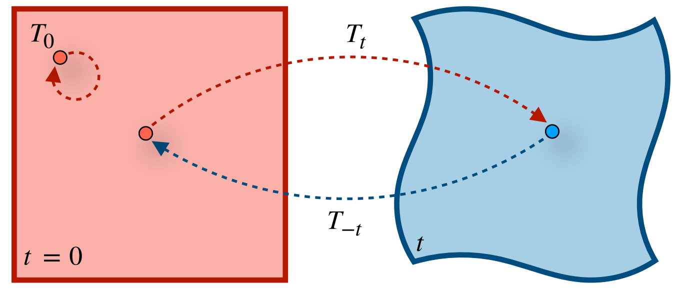


We parametrize the **warping** by a network $T : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ and train it using

$$\mathcal{L} = \lambda_1 \mathcal{W} + \lambda_2 \mathcal{D} + \lambda_3 \mathcal{T}$$

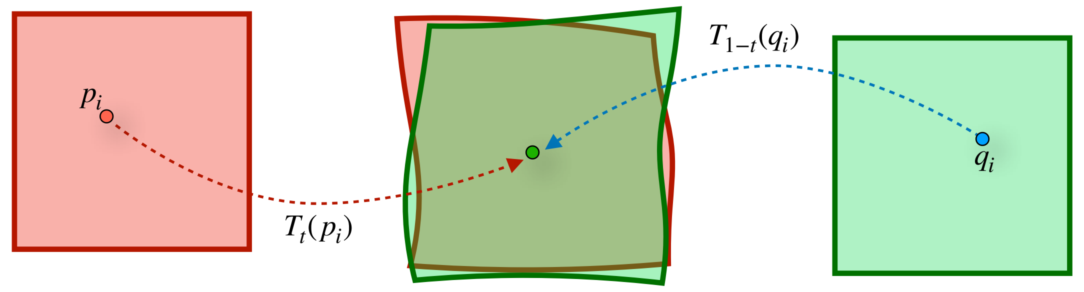
Warping constraint

$$\mathcal{W} = \underbrace{\int_{\mathbb{R}^2} \|T_0 - Id\|^2 dx}_{\text{Identity constraint}} + \underbrace{\int_{\mathbb{R}^2 \times \mathbb{R}} \|T_t \circ T_{-t} - Id\|^2 dx dt}_{\text{Inverse constraint}}$$



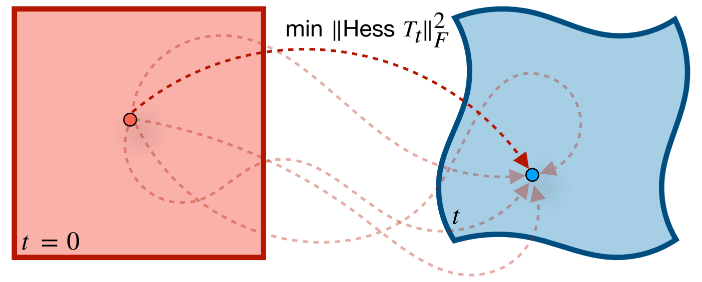
Data constraint

$$\mathcal{D} = \sum_i \int_{[0,1]} \|T_t(p_i) - T_{1-t}(q_i)\|^2 dt$$



Thin-plate constraint

$$\mathcal{T} = \int_{\mathbb{R}^2 \times \mathbb{R}} \|\text{Hess}\,(T)\|_F^2 dx dt$$



Finally, we blend the warped images $\mathcal{I}_i(\cdot, t)$ to define the morphing $\mathcal{I}$.

Poisson and generative blendings

We propose a *Poisson blending* where we align $\text{Jac}(\mathcal{I})$ with a vector field U, defined in terms of $\text{Jac}(\mathcal{I}_i)$. We optimize

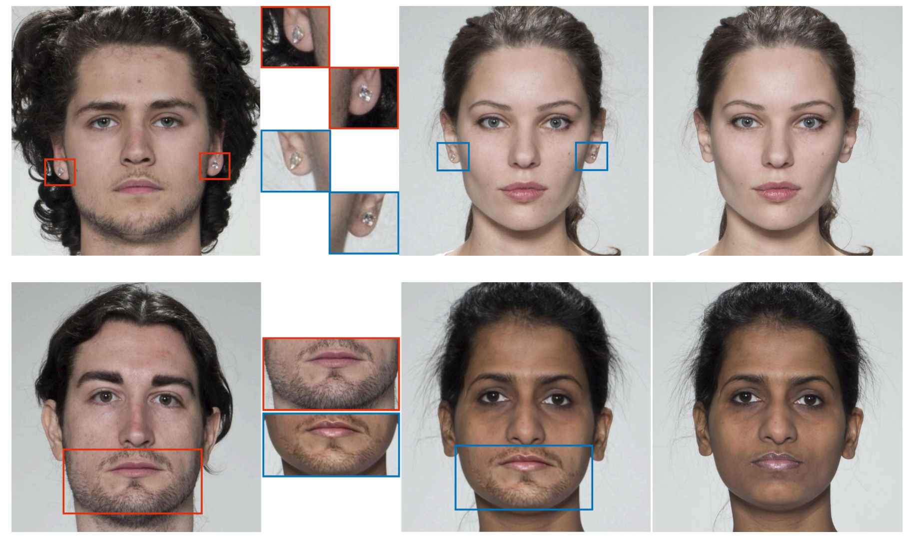
$$\mathcal{M} = \int_{\Omega} \|\text{Jac}(\mathcal{I}) - U\|^2 dx dt + \int_{S-\Omega} (\mathcal{I} - \mathcal{I}^*)^2 dx dt$$

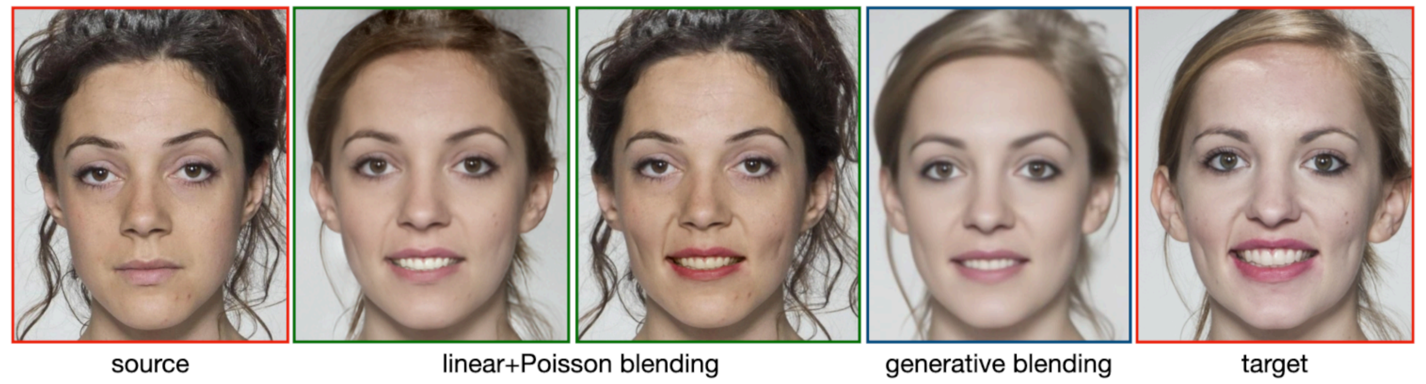
Let $\mathcal{E}$ and $\mathcal{D}$ be encoder and decoder models. We embed the warped images $\mathcal{I}_i$ in the latent space: $\mathcal{C}_i(t) = \mathcal{E}\big(\mathcal{I}_i(\cdot, t)\big)$. Then, we interpolate them using

$$\mathcal{I}_i(\cdot, t) = \mathcal{D}\big((1-t)\mathcal{C}_0(t) + t\mathcal{C}_1(t)\big)$$

Faces with different gender and ethnicity

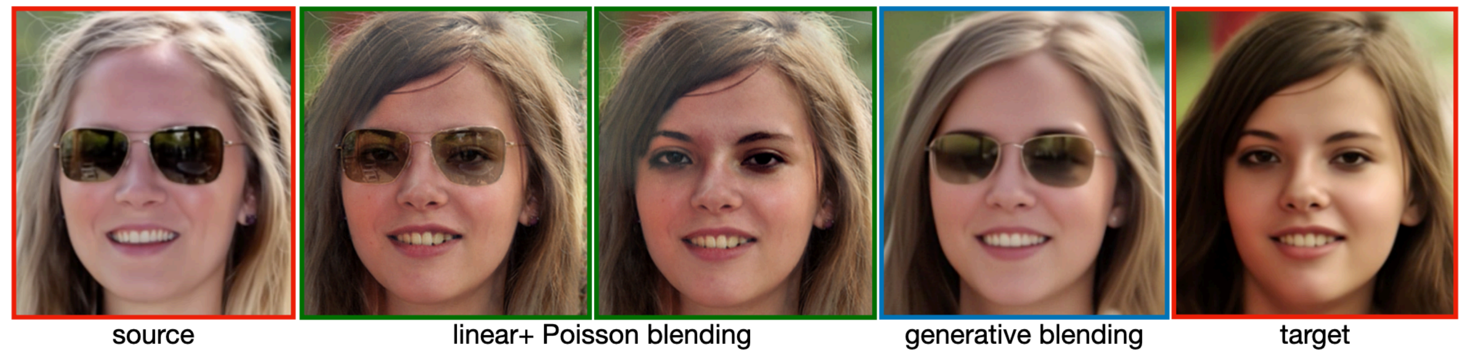
Poisson blending for feature transfer





Faces w/o full correspondence, e.g. the smile

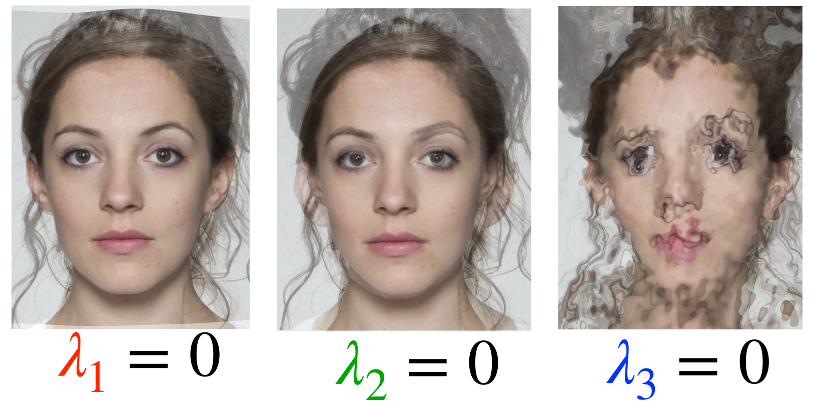
Faces w/ occlusion, e.g. the eyes.



Loss ablations and final remarks

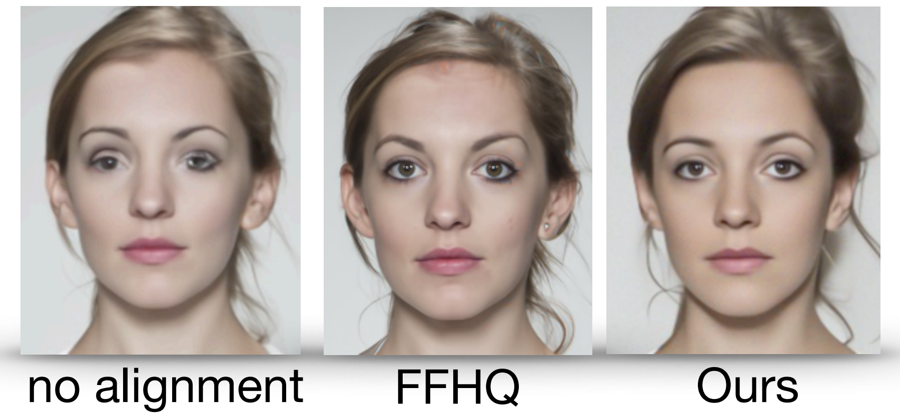
Each loss term plays a part in the warping results. Eliminating constraints by setting $\lambda_1, \lambda_1, \lambda_1 = 0$, lead to interesting effects.

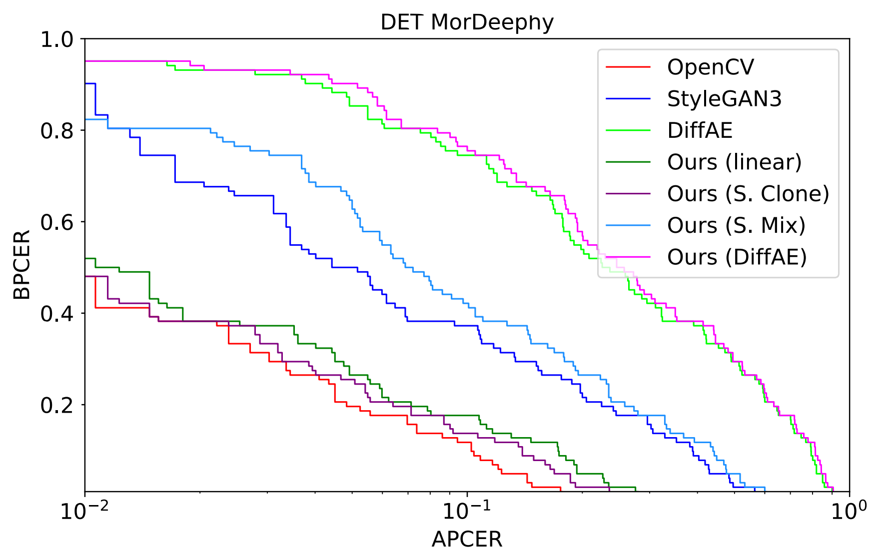
- $\mathcal{W}$ is responsible for morphing continuity.
- $\mathcal{D}$ is crucial for landmark matching. Without it, there is no warping.
- $\mathcal{T}$ minimizes spacial distortions and regularizes point trajectories, ensuring a not-too-strong non-linearity.



Our method handles pose variations gracefully. We may use generative methods for blending.

Generative methods demand aligned faces. Minor misalignments lead to mismatching.





Blending impacts morphing detection. Our warping+diffAE blending is comparable to pure diffAE, followed by StyleGAN3 and ours+S.Mix. Lastly classical warping and ours+(S.Clone/linear)

Non-linear warping leads to smoother alignment. Especially noticeable on videos