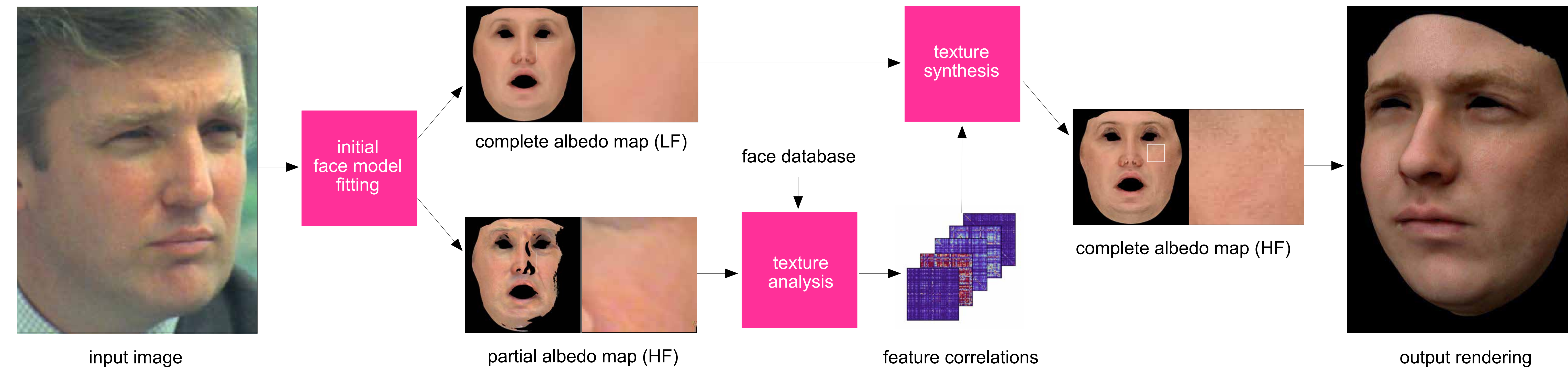


Photorealistic Facial Texture Inference Using Deep Neural Networks

SHUNSUKE SAITO^{1,2} LINGYU WEI^{1,2} LIWEN HU^{1,2} KOKI NAGANO³ HAO LI^{1,2,3}

¹Pinscreen Inc. ²University of Southern California ³USC Institute for Creative Technologies



PIPELINE

Face Model Fitting

PCA coefficients: optimization of photo consistency, landmark distances and regularization.

$$E(\chi) = w_c E_c(\chi) + w_{lan} E_{lan}(\chi) + w_{reg} E_{reg}(\chi)$$

$$E_c(\chi) = \frac{1}{|\mathcal{M}|} \sum_{p \in \mathcal{M}} \|C_{input}(p) - C_{synth}(p)\|_2$$

$$E_{lan}(\chi) = \frac{1}{|\mathcal{F}|} \sum_{f_i \in \mathcal{F}} \|f_i - \Pi_P(RV_i + t)\|_2^2$$

$$E_{reg}(\chi) = \sum_{i=1}^{80} \left[\left(\frac{\alpha_{id,i}}{\sigma_{id,i}} \right)^2 + \left(\frac{\alpha_{al,i}}{\sigma_{al,i}} \right)^2 \right] + \sum_{i=1}^{29} \left(\frac{\alpha_{exp,i}}{\sigma_{exp,i}} \right)^2$$

Texture Analysis

Optimal Gramian matrix: convex combination of Gramian matrices from dataset of high resolution real images

$$\begin{aligned} \min_w \quad & \sum_l \left\| \sum_k w_k G^l(\mathcal{M}(I_k)) - G^l(\mathcal{M}(I_0)) \right\|_F \\ \text{s.t.} \quad & \sum_{k=1}^K w_k = 1 \\ & w_k \geq 0 \quad \forall k \in \{1, \dots, K\} \\ & \mathcal{M}(I)_p = \begin{cases} 0.5, & \text{if } p \text{ is non-visible} \\ I_p, & \text{otherwise} \end{cases} \end{aligned}$$

Texture Synthesis

$$\min_I \sum_{l \in L_F} \|F^l(I) - F^l(I(\alpha_{al}))\|_F^2 + \lambda \sum_{l \in L_G} \left\| G^l(I) - \sum_k w_k G^l(I_k) \right\|_F^2$$

USER STUDY

