High Quality Facial Surface and Texture Synthesis via Generative Adversarial Networks - *Supplementary Material*

Anonymous ECCV submission

Paper ID ***

Maximum likelihood approach

In order to correlate between geometry and texture, we construct a joint 3DMM by concatenating the texture and geometry vectors to each other. Denote by $M = \begin{pmatrix} G \\ T \end{pmatrix}$ the $6m \times n$ matrix which consists of the geometries G and textures T vertically concatenated, and define $\Delta M = M - \mu_M \mathbb{1}^T$. Denote by U the matrix that contains the basis vectors of ΔM in its columns. U can be computed either by the eigenvalue decomposition of $\Delta M \Delta M^T$ as the eigenvectors, or more efficiently by the singular value decomposition of ΔM as the left singular vectors. We assume that the columns of U are ordered by the size of their corresponding singular values in a descending manner. Denote by U_g and U_t the upper and lower halves of U, such that $U = \begin{pmatrix} U_g \\ U_t \end{pmatrix}$. Similarily, denote $\mu_M = \begin{pmatrix} \mu_{M_g} \\ \mu_{M_t}. \end{pmatrix}$. Note that U_g and U_t are not equivalent to the previously defined V_g and V_t and are not orthogonal. Still, any geometry g and texture t of a given face in M can be represented as a linear combination

$$\begin{pmatrix} g \\ t \end{pmatrix} = \begin{pmatrix} \mu_{M_g} \\ \mu_{M_t} \end{pmatrix} + \begin{pmatrix} U_g \\ U_t \end{pmatrix} \beta, \tag{1}$$

where the coefficient vector β is mutual to the geometry and texture.

Given a coefficient vector β of a new face that was not used to construct the model, one can formulate the texture of the face as

$$t = U_t \beta + \mu_{M_t} + noise \tag{2}$$

Our goal is to estimate β given a texture t. According to the maximum likelihood approach, β is estimated as

$$\begin{aligned} \beta^* &= {}_{\beta} P(\beta|t) \\ &= {}_{\beta} P(t|\beta) P(\beta). \end{aligned} \tag{3}$$

Assuming, as before, that $P(t|\beta)$ and $P(\beta)$ follow a multivariate normal distribution with diagonal covariance matrices $\Sigma_{t|\beta}, \Sigma_{\beta}$ and mean $\mu_{t|\beta} = U_t \beta, \quad \mu_\beta = \mathbf{0}$. Then, one can write

$$\beta^* = {}_{\beta} \exp \left\{ -\frac{1}{2} (t - \mu_{t|\beta}) \Sigma_{t|\beta}^{-1} (t - \mu_{t|\beta})^T \right\}.$$

$$
\begin{aligned}
&\exp\left\{-\frac{1}{2}\beta\Sigma_{\beta}^{-1}\beta^{T}\right\} \\
&= {}_{\beta}(t - U_t\beta)\Sigma_{t|\beta}^{-1}(t - V_t\beta)^{T} + \beta\Sigma_{\beta}^{-1}\beta^{T} \quad . \qquad (4)
\end{aligned}
$$

The solution β^* is obtained by taking the gradient to zero, which yields

$$\beta^* = (U_t^T \Sigma_{t|\beta}^{-1} U_t + \Sigma_\beta)^{-1} (t^T \Sigma_{t|\beta}^{-1} U_t). \qquad (5)$$

We estimate the covariance matrix Σ_β empirically from the data. The covariance matrix $\Sigma_{t|\beta}$ can be estimated as well, however an arbitrary scale parameter which determines strength of the prior $P(\beta)$ relative to the data must be calibrated. In other words, this scale takes into account the noise level described in Equation 2. Finally, once the coefficient vector β is known, the geometry is obtained by $g = U_g\beta$.

Blend Shapes for different expressions

Fig. 1: Three identities generated by the proposed method, deformed with different Blend Shapes expressions.

Different poses and lighting conditions

Fig. 2: Identities generated by the proposed method with different pose and lighting.

Fig. 3: Identities generated by the proposed method with different pose and lighting.