

# Adaptive Contour Fitting for Pose-Invariant 3D Face Shape Reconstruction

Chengchao Qu<sup>1,2</sup>  
 chengchao.qu@iosb.fraunhofer.de  
 Eduardo Monari<sup>2</sup>  
 eduardo.monari@iosb.fraunhofer.de  
 Tobias Schuchert<sup>2</sup>  
 tobias.schuchert@iosb.fraunhofer.de  
 Jürgen Beyerer<sup>2,1</sup>  
 juergen.beyerer@iosb.fraunhofer.de

<sup>1</sup> Vision and Fusion Laboratory  
 Karlsruhe Institute of Technology  
 Karlsruhe, Germany  
<sup>2</sup> Fraunhofer IOSB  
 Karlsruhe, Germany

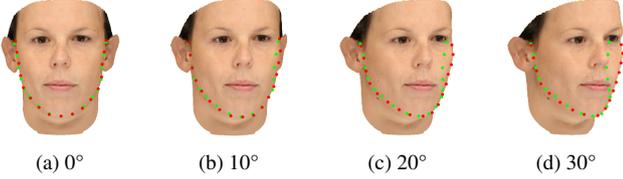


Figure 1: Correspondence errors of 2D (red) and 3D (green) facial contour landmarks w.r.t. yaw angles of (a) 0°, (b) 10°, (c) 20° and (d) 30°.

**Motivation** Direct reconstruction of 3D face shape—solely based on a sparse set of 2D feature points localized by a facial landmark detector—offers an automatic, efficient and illumination-invariant alternative to the widely known analysis-by-synthesis framework, which is extremely time-consuming considering the enormous parameter space for both shape and photometric properties. Given 2D landmarks  $\mathbf{y}$  and their correspondence on the 3D Morphable Model (3DMM), the 3D shape can be recovered by minimizing the distance between 2D and the projected 3D landmarks

$$\hat{\alpha} = \arg \min_{\alpha} E(\alpha) = \arg \min_{\alpha} \|\mathbf{Q}\alpha - \mathbf{y}\|_2^2 + \eta \|\alpha\|_2^2, \quad (1)$$

where  $\alpha$  is the target 3DMM shape parameter.  $\mathbf{Q}$  is a simplification of the projected 3D landmarks  $\mathbf{Q} = \Pi\Phi\mathbf{S}\text{diag}(\sigma)$  containing the affine camera projection matrix  $\Pi$ , sparse landmark selection  $\Phi$ , as well as 3DMM eigenvectors  $\mathbf{S}$  and eigenvalues  $\sigma$ . The term  $\eta \|\alpha\|_2^2$  prevents overfitting and regulates the plausibility of the reconstructed faces.

The empirical assumption of a fixed mapping between 2D and 3D landmarks gives rise to a major flaw of the landmark-based methods. Fig. 1 demonstrates that with increasing yaw angle, remarkable deviation in the self-occluded face half can be observed. Moreover, we argue that not only the invisible contour landmarks, but also the visible ones are susceptible to 2D–3D correspondence discrepancy. Even in frontal view, a tight correspondence cannot be necessarily guaranteed (see Fig. 1a). In this paper, we propose a novel algorithm to address these two problems.

**Fast Detection of Silhouette Vertices** Due to the fact that 2D contour landmarks are always located on the face silhouette, the tangent planes of such 3D vertices are perpendicular to the image plane, meaning that their normal vectors projected onto the z-axis are close to zero. It is then an intuitive idea to treat those with the absolute z-projection values of the normals  $|\mathbf{n}_z| < t$  as silhouette points (see Fig. 2d). However, a universally valid threshold  $t$  is hard to find and the spatial distribution is uncontrollable, too. Instead, we impose geometric constraints by labeling extended vertices starting from the original 3D contour landmarks on the 3DMM towards the center of the face (see Fig. 2a). During fitting, the ones with the smallest  $|\mathbf{n}_z|$  on each line are chosen (see Figs. 2b and 2c). The number of the evaluated normals is thereby reduced by two orders of magnitude, allowing for fast and robust online update during the iterations.

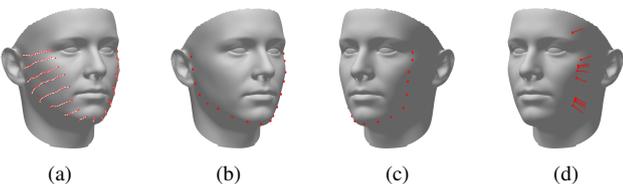


Figure 2: Fast detection of silhouette vertices using (a) annotated candidates. (b) and (c) show our result compared to (d) the direct approach.

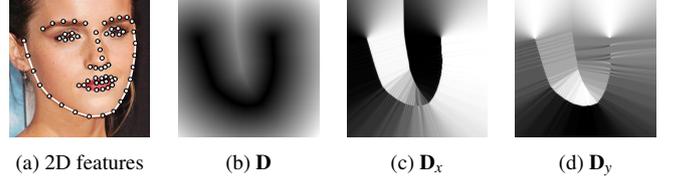


Figure 3: (a) Improved 2D features with connected contour lines, (b) their Distance Transform and its derivatives in (c) x and (d) y directions.

**Fitting** To facilitate flexible fitting of 3D contour vertices, the continuous curve formed by the corresponding discrete 2D landmarks is exploited when reconstructing the shape (see Fig. 3a). Separating all landmarks into two disjoint subsets of fixed and contour ones in Eq. (1) leads to

$$E(\alpha) = \left\| \begin{bmatrix} \mathbf{Q}_{\text{contour}} \\ \mathbf{Q}_{\text{fixed}} \end{bmatrix} \alpha - \begin{bmatrix} \mathbf{y}_{\text{contour}} \\ \mathbf{y}_{\text{fixed}} \end{bmatrix} \right\|_2^2 + \eta \|\alpha\|_2^2. \quad (2)$$

An unknown mapping denoted  $\phi(i) = j$  which selects, for each 3D contour vertex  $i$ , the corresponding 2D contour pixel  $j$ , is now also a part of the minimization process

$$E(\alpha, \phi) = \sum_i \left\| \mathbf{Q}_i \alpha - \mathbf{y}_{\phi(i)} \right\|_2^2 + \left\| \mathbf{Q}_{\text{fixed}} \alpha - \mathbf{y}_{\text{fixed}} \right\|_2^2 + \eta \|\alpha\|_2^2. \quad (3)$$

Estimation of  $\alpha$  is then a “minimization of minimization” problem

$$\hat{\alpha} = \arg \min_{\alpha} \sum_i \min_j \left\| \mathbf{Q}_i \alpha - \mathbf{y}_j \right\|_2^2 + \left\| \mathbf{Q}_{\text{fixed}} \alpha - \mathbf{y}_{\text{fixed}} \right\|_2^2 + \eta \|\alpha\|_2^2. \quad (4)$$

The trick to circumvent this difficulty is to apply Distance Transform (DT) to the discrete 2D features with the Levenberg-Marquardt Iterative Closest Point (LM-ICP) algorithm. On the 2D image raster  $\mathbf{x}$ , DT assigns each image pixel with the distance to its closest point on the contour lines  $D(\mathbf{x}) = \min_j \|\mathbf{x} - \mathbf{y}_j\|_2$ , which is reusable for the entire reconstruction procedure by virtue of its independence of the model parameter  $\alpha$  (see Fig. 3b). The merit of DT lies in that the mapping function  $\phi(i)$ , or the minimization over  $j$  in the contour cost of Eq. (3), then vanishes and is thereby simply replaced with  $D(\mathbf{Q}_i \alpha) = \min_j \|\mathbf{Q}_i \alpha - \mathbf{y}_j\|_2$ . Integrating DT into Eq. (3) and vectorizing over all contour vertices  $i$  yields

$$E(\alpha) = \|\mathbf{D}(\mathbf{Q}_{\text{contour}} \alpha)\|_2^2 + \left\| \mathbf{Q}_{\text{fixed}} \alpha - \mathbf{y}_{\text{fixed}} \right\|_2^2 + \eta \|\alpha\|_2^2. \quad (5)$$

Rather than the sum of squares, LM-ICP demands the vector of residuals

$$\mathbf{e}(\alpha) = \begin{bmatrix} \mathbf{D}(\mathbf{Q}_{\text{contour}} \alpha) \\ \mathbf{Q}_{\text{fixed}} \alpha - \mathbf{y}_{\text{fixed}} \\ \sqrt{\eta} \alpha \end{bmatrix}. \quad (6)$$

Differentiating Eq. (6) analytically subject to the shape parameter  $\alpha$  is possible when the chain rule  $\frac{\partial D_i}{\partial \alpha_j} = \frac{\partial D}{\partial \mathbf{x}} \mathbf{f}_i \cdot \frac{\partial \mathbf{f}_i}{\partial \alpha_j} + \frac{\partial D}{\partial \mathbf{y}} \mathbf{f}_i \cdot \frac{\partial \mathbf{f}_i}{\partial \alpha_j}$  and the pre-computed gradient images in x and y directions (Figs. 3c and 3d) are applied to calculate the derivatives of the contour cost  $\nabla_{\alpha} \mathbf{D}(\mathbf{Q}_{\text{contour}} \alpha)$ . The target Jacobian matrix  $\mathbf{J}_{ij} = \frac{\partial e_i}{\partial \alpha_j}$  is then

$$\mathbf{J} = \begin{bmatrix} \mathbf{D}_x(\mathbf{Q}_{\text{contour}} \alpha) \cdot \mathbf{Q}_{\text{contour}}^x + \mathbf{D}_y(\mathbf{Q}_{\text{contour}} \alpha) \cdot \mathbf{Q}_{\text{contour}}^y \\ \mathbf{Q}_{\text{fixed}} \\ \sqrt{\eta} \mathbf{I} \end{bmatrix}. \quad (7)$$

**Results** On the Basel Face Model (BFM), promising results outperforming state-of-the-art methods demonstrate the robustness of the proposed 3D shape reconstruction framework against pose variation. The effectiveness and efficiency justify our theoretical and practical advances.