

# 3D Face Reconstruction from a Single Non-Frontal Face Image

Naoki NOZAWA\* Daiki KUWAHARA\* Shigeo MORISHIMA †  
Waseda University\* Waseda Research Institute for Science and Engineering †

## 1. Introduction

A reconstruction of a human face shape from a single image is an important theme for criminal investigation such as recognition of suspected people from surveillance cameras with only a few frames. It is, however, still difficult to recover a face shape from a non-frontal face image. Method using shading cues on a face depends on the lighting circumstance and cannot be adapted to images in which shadows occurs, for example [Kemelmacher et al. 2011]. On the other hand, [Blanz et al. 2004] reconstructed a shape by 3D Morphable Model (3DMM) only with facial feature points. This method, however, requires the pose-wise correspondences of vertices in the model to feature points of input image because a face contour cannot be seen when the facial direction is not the front. In this paper, we propose a method which can reconstruct a facial shape from a non-frontal face image only with a single general correspondence table. Our method searches for the correspondences of points on a facial contour in the iterative reconstruction process, and makes the reconstruction simple and stable.

## 2. Deforming 3D face model

We build the 3D face model for representing a shape. This model is the same as [Maejima et al. 2008]. For each person  $j \in (1, \dots, M)$ , vertex position  $\mathbf{x} = (x_1, y_1, z_1, \dots, x_N, y_N, z_N)^T$  is represented as Equation (1) by a Principal Component Analysis with model parameter  $(p_1, \dots, p_{M-1}) \in \mathbf{p}$ :

$$\mathbf{x}_j = \mathbf{m} + \sum_{i=1}^{M-1} p_i \mathbf{b}_i \quad (1)$$

where  $\mathbf{m}$  and  $\mathbf{b}$  are the average shape and eigenvector. To avoid the influence of noise in scanned data and over-fitting of the model, we compress the dimension from  $M - 1$  to  $D$  ( $D < M - 1$ ).

After building the model, we solve a nonlinear optimization problem by minimizing the energy on Equation (2). We introduce *Silhouette* term  $E_s$  in addition to *Inner-Point* term  $E_p$  and *Facial-Likelihood* term  $E_l$  proposed in [Maejima et al. 2008].  $\alpha, \beta$ , and  $\gamma$  decide contribution for each term on  $E$ .

$$E(\mathbf{p}) = \alpha E_p(\mathbf{p}) + \beta E_s(\mathbf{p}) - \gamma E_l(\mathbf{p}) \quad (2)$$

$E_p$  is defined as the distance on the 2D image coordinates between facial feature points and corresponding model vertices except for those lying on a facial contour.

$$E_p(\mathbf{p}) = \frac{1}{N} \sum_{i=1}^D \|\mathbf{W}_i \{\mathbf{A} \mathbf{M}_i(\mathbf{p}) - \mathbf{t}_i\}\|^2 \quad (3)$$

$\mathbf{M}$  and  $\mathbf{t}$  are the positions of vertices and feature points, and  $\mathbf{A}$  indicates the affine matrix which is decided by a facial direction and scale of input image.  $\mathbf{W}$  deletes the influence of depth and vertices which have no relation to feature points.

$E_s$  is defined as the distance on the 2D image coordinates between feature points and vertices lying on a facial contour. Moreover, in order to prevent a shape from becoming awkward, we add the energy when right and left reversed input image is used. The positions of  $\mathbf{M}'$ ,  $\mathbf{t}'$  and those of  $\mathbf{M}$ ,  $\mathbf{t}$  are symmetrical

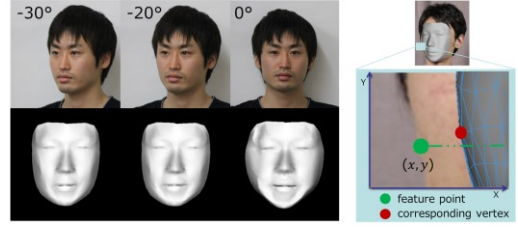


Figure 1: Reconstructed faces in our method (left)

Figure 2: The correspondence on a contour (right)

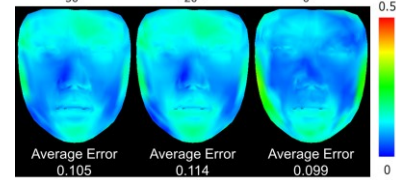


Figure 3: Error from grand truth

about the central axis which is parallel to the Y axis.  $\mathbf{A}'$  indicates affine matrix decided by the revised image.

$$E_s(\mathbf{p}) = \frac{1}{N} \sum_{i=1}^D \{ \|\mathbf{W}_i \{\mathbf{A} \mathbf{M}_i(\mathbf{p}) - \mathbf{t}_i\}\|^2 + \|\mathbf{W}_i \{\mathbf{A}' \mathbf{M}'_i(\mathbf{p}) - \mathbf{t}'_i\}\|^2 \} \quad (4)$$

$E_l$  is represented of Mixture of Gaussian  $\Psi$  which was learned by EM algorithm. We take the log of this distribution and set it as restriction to the deformation of 3D face model.

$$E_l(\mathbf{p}) = \sum_{k=1}^K \ln(P(k) \Psi_k(\mathbf{p} | \mu_k, \Sigma_k)) \quad (5)$$

$\Sigma$  and  $\mu$  are variance and mean of Gaussian, and  $P(k)$  is the mixing coefficient.

While calculating  $E$ , we search for the vertices corresponding to the feature points on a facial contour until the iteration of nonlinear optimization converges. Firstly, the model is rotate and scaled according to the input image. Secondly, the polygons which contain the y value of a feature point are searched for. Finally, the polygon which is on the silhouette line of the model is found, and a vertex in outside the most is selected.

## 3. Results and Future Work

Our results are shown in Figure 1. In our method, reconstruction results from a frontal face become the same as [Maejima et al 2008]. Figure 3 shows error maps which represent the distance on 3D coordinate between our results and grand truth. To match the coordinate of the model and grand truth, we use Axis Aligned Bounding Box (AABB). The scale is adjusted by half the distance of a space diagonal line of AABB to 1, and the position is adjusted by the center. The results show that our method achieved to reconstruct the models from non-frontal face with small difference of error comparing with that of frontal faces. Our method shows significant improvement around cheek.

Applying to arbitrary angles is the prime task in our future work.

## References

- V. BLANZ, A. MEHL, T. VETTER, H. SEIDEL. 2004. A Statistical Method for Robust 3D Surface Reconstruction from Sparse Data. *Int. Symp. on 3DPVT*, pp293-300.
- I. K-SHLIZERMAN. 2010. 3D Face Reconstruction from a Single Image Using a Single Reference Face Shape. *Trans. on PAMI, Vol.33 No.2* pp.394-405.
- A. MAEJIMA, S. MORISHIMA. 2008. Fast Plausible 3D Face Generation From Single Photograph. *SIGGRAPH ASIA, Poster*.

\*e-mail: [s112800563@akane.waseda.jp](mailto:s112800563@akane.waseda.jp)

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).

SIGGRAPH 2015 Posters, August 09 – 13, 2015, Los Angeles, CA.

ACM 978-1-4503-3632-1/15/08.

<http://dx.doi.org/10.1145/2787626.2792634>