

RESEARCH

Open Access

Single view-based 3D face reconstruction robust to self-occlusion

Youn Joo Lee¹, Sung Joo Lee², Kang Ryoung Park³, Jaeik Jo¹ and Jaihie Kim^{1*}

Abstract

State-of-the-art 3D morphable model (3DMM) is used widely for 3D face reconstruction based on a single image. However, this method has a high computational cost, and hence, a simplified 3D morphable model (S3DMM) was proposed as an alternative. Unlike the original 3DMM, S3DMM uses only a sparse 3D facial shape, and therefore, it incurs a lower computational cost. However, this method is vulnerable to self-occlusion due to head rotation. Therefore, we propose a solution to the self-occlusion problem in S3DMM-based 3D face reconstruction. This research is novel compared with previous works, in the following three respects. First, self-occlusion of the input face is detected automatically by estimating the head pose using a cylindrical head model. Second, a 3D model fitting scheme is designed based on selected visible facial feature points, which facilitates 3D face reconstruction without any effect from self-occlusion. Third, the reconstruction performance is enhanced by using the estimated pose as the initial pose parameter during the 3D model fitting process. The experimental results showed that the self-occlusion detection had high accuracy and our proposed method delivered a noticeable improvement in the 3D face reconstruction performance compared with previous methods.

Keywords: 3D face reconstruction, Single image, Self-occlusion, Arbitrary view, 3D model fitting

Introduction

3D face modeling originated with Parke's pioneering studies [1,2], which aimed to generate realistic faces for computer animation. Since Parke's work, 3D face reconstruction has attracted considerable attention from many computer vision researchers because it has many useful applications, such as pose-invariant face recognition [3,4], age-invariant face recognition [5,6], 3D face generation for game and movie characters [7,8], monitoring suspects using surveillance camera systems, video conferencing, and automatic face conversion of a 2D face image into a 3D face for 3D TV.

3D face modeling technologies can be divided into two basic approaches. One approach uses specific sensors, such as stereographic cameras, structured light, or 3D laser scanners [9]. These methods produce accurate 3D face data, but they are expensive and require additional operations, such as calibration. To overcome these limitations, a monocular camera-based approach has been

researched intensively. This approach can be categorized further into single view-based and multi-view-based approaches. The multi-view-based approach uses more 2D facial information than the single view-based approach, but it has some limitations such as a requirement for multiple images and detection of the correspondences among the images. In this study, we focus on the single view-based approach.

Among the single view-based methods, shape-from-shading (SFS) is a traditional method for deriving a 3D facial shape from the brightness variations in a single image. However, SFS-based methods have impractical constraints because the Lambertian reflectance model and a known light source direction need to be assumed to produce accurate results [10-12]. Recently, several new techniques have been proposed to overcome this problem. These methods reconstruct a 3D face by modeling the relationships between the intensities and the depth information of the face using statistical learning techniques, such as principal component analysis (PCA), partial least squares, and canonical correlation analysis [13-16]. However, all of these methods assume that the input face is viewed from the front, but this strict

* Correspondence: jhkim@yonsei.ac.kr

¹School of Electrical and Electronic Engineering, Yonsei University, Seoul, Republic of Korea

Full list of author information is available at the end of the article

constraint cannot always be satisfied in real world applications, such as surveillance camera systems.

The state-of-the-art 3D morphable model (3DMM), which was proposed by Blanz and Vetter [3,17], is a single-image-based 3D face reconstruction method that requires no frontal pose constraint. In this method, 3D face reconstruction is performed by fitting a morphable model to a 2D image. The reconstructed 3D face is represented by model parameters that minimize the texture residual errors between the rendered model image and the input image. To generate a high-quality 3D face, the model parameters contain rendering parameters, such as the camera geometry and illumination direction, as well as facial texture and shape parameters. Therefore, 3DMM can reconstruct a more realistic 3D face in less restrictive conditions. However, it has a high computational complexity because of the large number of model parameters that need to be estimated simultaneously [4]. In addition, 3DMM requires manual initialization and a dense point-to-point correspondence between all of the face images.

Simplified versions [5,6,18-20] of 3DMM have been proposed to reduce these computational costs. Unlike the original 3DMM, most simplified 3DMM (S3DMM)-based methods use a sparse shape model, which is constructed by statistically learning a data set of 3D facial feature points (FFPs). The FFPs indicate the salient features of a face, such as the corners of the eyes, nose tip, and the corners of the mouth.

S3DMM reconstructs a 3D facial shape (that consists of 3D FFPs) by finding the optimal shape parameter and the pose parameter that minimizes the difference between the projected 3D FFPs of the shape model and the input 2D FFPs. Therefore, the S3DMM-based method does not need to find a dense correspondence between the faces and it has a lower computational complexity because of the drastic reduction in the number of parameters. However, some of these methods [18,19] still have the limitation that the input face needs to be a frontal view whereas others [5,6,20] are unaffected by pose variations. Wang et al. [20] proposed an automatic framework for 3D face reconstruction on the basis of an arbitrary view image and estimated the shape and pose parameters using an expectation-maximization (EM) algorithm. However, they did not report any quantitative results and only qualitative results were presented from some test images in their experiments. Park et al. [5,6] used S3DMM to create a 3D aging model for age-invariant face recognition. They derived a 3D facial shape by alternating the pose parameter and the shape parameter until the shape residual error converged. The pose and shape parameters were estimated separately using the least squares method. This alternation methodology is used in most S3DMM-based methods [5,6,18,20] because it reduces the computational time and produces a

linear cost function, compared with the one-step methodology that estimates all of the parameters at the same time using a non-linear optimization process.

Existing methods [5,6,20] are known to perform using an arbitrary view, but they are not robust to pose variations. This is because they are vulnerable to self-occlusion errors in the 2D input FFPs caused by pose changes. If an input face is rotated, some parts of the face are self-occluded by other parts. Therefore, the occluded facial region is not visible in a rotated face image, and the real FFPs placed on the region are also not observable. As a result, the FFPs detected on the rotated face image contain location errors caused by self-occlusion. These errors degrade the performance of S3DMM-based 3D face reconstruction. Unfortunately, existing S3DMM-based methods have not addressed this problem.

Therefore, we propose a solution to the self-occlusion problem of S3DMM-based 3D face reconstruction. Our proposed method consists of the following two steps. In the first step, visible FFPs that are not affected by self-occlusion are automatically discriminated from self-occluded FFPs by estimating the head pose using a cylindrical head model. In the next step, 3D face reconstruction is performed using a model fitting scheme based on selected visible FFPs to reduce the self-occlusion effect. The performance of the 3D face reconstruction is enhanced by using a pose estimated with the cylindrical head model as the initial pose parameter in the proposed model fitting. In experiments, we evaluated the performance of our proposed method qualitatively and quantitatively by using ground-truth 3D face data acquired with a 3D laser scanner. The experimental results showed that the visible FFPs were selected with high accuracy and that the 3D face reconstruction performance with the proposed method was improved greatly compared with previous S3DMM-based methods. Comparisons of previous methods and our proposed method are summarized in Table 1.

The remainder of this paper is organized as follows. In the following section, the S3DMM is explained and the self-occlusion problem is analyzed. In Section "Proposed 3D face reconstruction method", our proposed method is described, including the automatic selection of visible FFPs and the model fitting strategy. The experimental results are presented in Section "Experimental results". Finally, our conclusions are summarized in Section "Conclusions".

S3DMM and self-occlusion problem

S3DMM

In S3DMM, the geometry of a face is defined as a shape vector $\mathbf{S} = (X_1, Y_1, Z_1, X_2, \dots, Y_n, Z_n)^T \in \mathfrak{R}^{3n}$, which contains the X , Y , and Z -coordinates of n vertices. The

Table 1 Comparison of the previous methods and the proposed method

Categories	Methods	Strengths	Weaknesses
Frontal view	Shape-from-Shading [10-12]	- No requirement for training data	- Infeasible constraints (Lambertian reflectance model and known light source direction)
	Modeling the relationship between intensities and depths [13-16]	- Low computational complexities	- Requires pixel-by-pixel alignment between intensities and depth map - Requirement of training data
	Simplified version of 3D morphable model[18,19]	- Low computational complexity	- Requirement of training data
Arbitrary view	3D morphable model [3,17]	- Robust to pose and illumination variation - Less sensitive to self-occlusion	- High computational complexity - Dense correspondence between faces - Requirement of training data
	Simplified version of 3D morphable model - EM algorithm [20] - Least squares [5,6]	- Low computational complexity	- Requirement of training data - Self-occlusion problem
	Simplified version of 3D morphable model - Proposed method	- Low computational complexity - Robust to self-occlusion	- Requirement of training data

original 3DMM generally uses a dense shape with thousands of vertices, whereas S3DMM uses a sparse shape with only dozens of vertices. In order to build a morphable shape model, S3DMM performs PCA on a training set of shape vectors \mathbf{S}_j . The mean shape \mathbf{s}_0 and m shape variations \mathbf{s}_i are then obtained, and a new shape \mathbf{S} can be expressed as a linear combination of the mean shape \mathbf{s}_0 and the shape variations \mathbf{s}_i as follows:

$$\mathbf{S} = \mathbf{s}_0 + \sum_{i=1}^m \beta_i \mathbf{s}_i \quad (1)$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_m)^T$ is the shape parameter and m is the dimension of the shape parameter which was determined to represent 99 % of the shape variations of the training face set [21]. Finally, a new 3D facial shape can be generated by changing the shape parameter $\boldsymbol{\beta}$.

Given the 2D FFPs of an input face image, such as $\mathbf{s}_{2d} = (x_1, y_1, x_2, \dots, y_n)^T \in \mathcal{R}^{2n}$, the shape parameter $\boldsymbol{\beta}$ needs to be determined such that it minimizes the shape residual between the projected 3D facial shape generated by the shape parameter and the input 2D facial shape. The optimal shape and pose parameters $(\boldsymbol{\beta}, \mathbf{R}_\theta, \mathbf{T})$ are obtained from (2):

$$\arg \min_{\boldsymbol{\beta}, \mathbf{R}_\theta, \mathbf{T}} \|\mathbf{P}(\mathbf{R}_\theta \tilde{\mathbf{S}} + \tilde{\mathbf{T}}) - \tilde{\mathbf{s}}_{2d}\|^2 \quad (2)$$

Where $\tilde{\mathbf{S}}$ is a $3 \times n$ matrix that is reshaped from the $3n \times 1$ model shape vector \mathbf{S} obtained using (1), $\tilde{\mathbf{s}}_{2d}$ is a $2 \times n$ matrix that is reshaped from the $2n \times 1$ input shape vector \mathbf{s}_{2d} , \mathbf{P} is a 2×3 orthographic projection matrix, $\tilde{\mathbf{T}}$ is a $3 \times n$ translation matrix consisting of n translation vectors $\mathbf{T} = [t_x \ t_y \ t_z]^T$, and \mathbf{R}_θ is a 3×3 rotation matrix where the yaw angle is θ . Note that in this paper, we consider mainly yaw rotation because the self-occlusion caused by yaw rotation is relatively greater than that caused by pitch rotation, and t_z is set to 0 because an orthographic projection is assumed.

Several methodologies are used to estimate the model parameters for S3DMM. We used the alternation methodology of [5], which alternately finds the optimal shape parameter and pose parameter until the shape residual converges, because this approach reduces the computational cost by transforming a non-linear cost function into a linear one. As shown in Algorithm 1, the procedure for 3D model fitting is as follows. First, the shape parameter $\boldsymbol{\beta}_0$ and translation parameter \mathbf{T}_0 are initialized to $\mathbf{0}$ and the input 2D FFPs \mathbf{s}_{2d} are aligned with the 2D mean shape obtained by projecting the 3D mean shape (\mathbf{s}_0) with a frontal pose onto the x - y plane. As the alignment method, we use the Procrustes analysis, which includes translation, rotation, and scaling [22]. The optimal model parameters are determined by alternately updating the pose parameter $(\mathbf{R}_\theta, \mathbf{T})$ at the fixed $\boldsymbol{\beta}$ and updating the shape parameter $\boldsymbol{\beta}$ at fixed $(\mathbf{R}_\theta, \mathbf{T})$ until the shape residual error converges. The cost function is solved as a least squares problem and the rotation matrix is calculated by QR decomposition, as in [5]. Finally, a new 3D facial shape \mathbf{S}_{3d} is reconstructed by applying the optimal shape parameter $\boldsymbol{\beta}$ to (1).

Algorithm 1 3D Model Fitting [5]

Input: $\mathbf{s}_{2d} = (x_1, y_1, x_2, \dots, y_n)^T$
 Output: $\mathbf{S}_{3d} = (X_1, Y_1, Z_1, X_2, \dots, Y_n, Z_n)^T$

1. Initialization
Set $\boldsymbol{\beta}_0 = \mathbf{0}$, $\mathbf{T}_0 = \mathbf{0}$ and $k = 1$.
2. Alignment
 \mathbf{s}_{2d} is aligned with the 2D mean shape obtained by projecting the frontal 3D mean shape (\mathbf{s}_0) onto the x - y plane.
3. Update \mathbf{R}_θ and \mathbf{T} with the fixed shape parameter

$$\arg \min_{(\mathbf{R}_\theta, \mathbf{T})_k} \|\mathbf{P}(\mathbf{R}_\theta \tilde{\mathbf{S}} + \tilde{\mathbf{T}}) - \tilde{\mathbf{s}}_{2d}\|_{\boldsymbol{\beta}=\boldsymbol{\beta}_{k-1}}^2$$

Update the shape parameter with the fixed pose parameter

$$\arg \min_{\beta_k} \left\| \mathbf{P}(\mathbf{R}_\theta \tilde{\mathbf{S}} + \tilde{\mathbf{T}}) - \tilde{\mathbf{s}}_{2d} \right\|_{(\mathbf{R}_\theta, \mathbf{T}) = (\mathbf{R}_\theta, \mathbf{T})_k}^2$$

4. Reconstruct $(\mathbf{S}_{3d})_k$ using the shape parameter β_k
5. Verify whether

$$\left\| \mathbf{P}(\mathbf{R}_\theta (\tilde{\mathbf{S}}_{3d})_k + \tilde{\mathbf{T}}) - \tilde{\mathbf{s}}_{2d} \right\|_{(\mathbf{R}_\theta, \mathbf{T}) = (\mathbf{R}_\theta, \mathbf{T})_k} < \epsilon$$

If not, go to Step 3 and $k = k + 1$

6. $\beta = \beta_k$ and $(\mathbf{R}_\theta, \mathbf{T}) = (\mathbf{R}_\theta, \mathbf{T})_k$
7. Reconstruct \mathbf{S}_{3d} using the final shape parameters

Self-occlusion problem

S3DMM-based methods reconstruct a 3D facial shape from 2D FFPs detected in an input 2D image. Therefore, these methods are vulnerable to large location errors that affect the 2D FFPs observed in a given image. Unfortunately, the observed 2D FFPs may have severe location errors caused by self-occlusion when detecting the 2D FFPs in a highly rotated face image, as shown in Figures 1 and 2.

Figure 1 shows self-occlusion errors that occur after comparing the observed 2D FFPs with the ground-truth 2D FFPs of the facial contours. When detecting the facial contour FFPs in a half-profile view image, the visible FFPs laid on the visible facial region can be detected as those of the ground-truth facial contour, but the observed FFPs on the occluded facial region are located in the outline of the face because the occluded real facial contour cannot be observed, as shown in Figure 1a. Therefore, the 2D FFPs observed in a rotated face image have location errors, which are the differences between the observed FFPs and the occluded real FFPs, as shown in Figure 1c. These errors increase as the degree of head rotation increases, as shown in Figure 2. As a result, this self-occlusion problem degrades the S3DMM performance.

Proposed 3D face reconstruction method

Overall procedure of the proposed method

The proposed 3D face reconstruction process starts with the localization of the FFPs in a given 2D face image. To detect self-occlusion in an input face, the head pose is estimated using a cylindrical head model-based method [23]. The estimated pose can then be used to determine which FFPs are self-occluded. Next, a sparse 3D facial shape is reconstructed using the model fitting process based on the selected visible FFPs. Subsequently, a dense 3D facial shape is interpolated from the reconstructed sparse 3D facial shape using the Thin Plate Spline (TPS) method [24,25]. Finally, the facial texture directly extracted from the input image is mapped onto the dense 3D facial shape. The overall procedure of the proposed method is shown in Figure 3.

Head pose estimation

The head orientation of an input face is useful for determining whether the FFPs are self-occluded. There are many appearance-based head pose estimation methods. Manifold embedding methods such as PCA, KPCA, LDA, and kernel discriminant analysis have been used to extract texture features, and these features are used to estimate the discrete head pose [26-29]. Murphy-Chutorian et al.[30] used local gradient orientation and estimated the continuous yaw and pitch using support vector regression. However, the accuracies of these methods can be affected by the detection performance in the face region and these methods require many training samples [31]. Therefore, we used a cylindrical head model [23] in this study, which is a geometry-based head pose estimation method. In general, S3DMM-based methods alone can estimate the head pose because they can rotate the 3D shape model as closely as possible to the pose of the input face to find the best-matched shape model for the input 2D facial

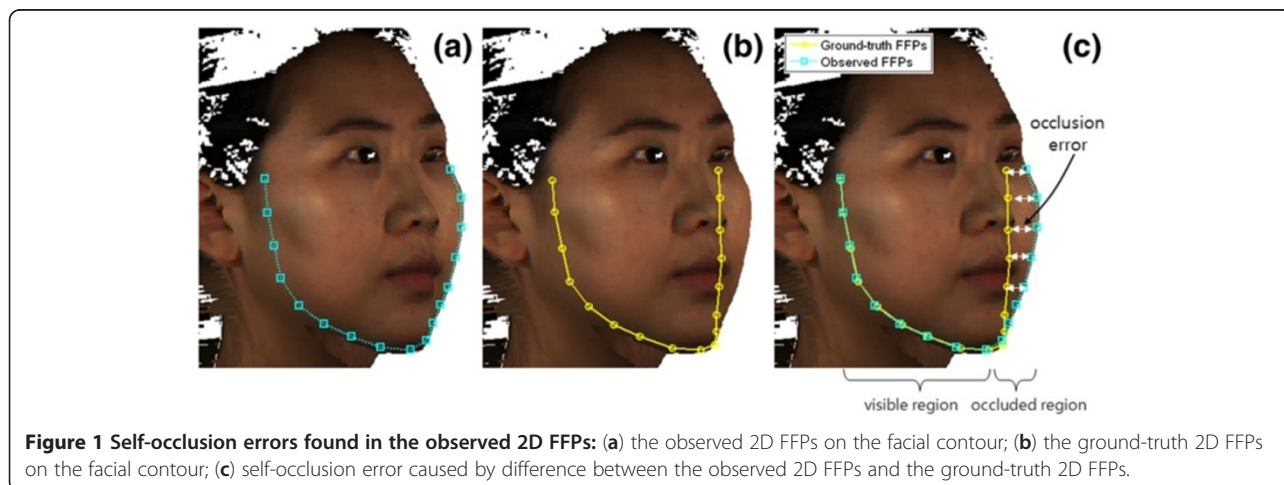


Figure 1 Self-occlusion errors found in the observed 2D FFPs: (a) the observed 2D FFPs on the facial contour; (b) the ground-truth 2D FFPs on the facial contour; (c) self-occlusion error caused by difference between the observed 2D FFPs and the ground-truth 2D FFPs.

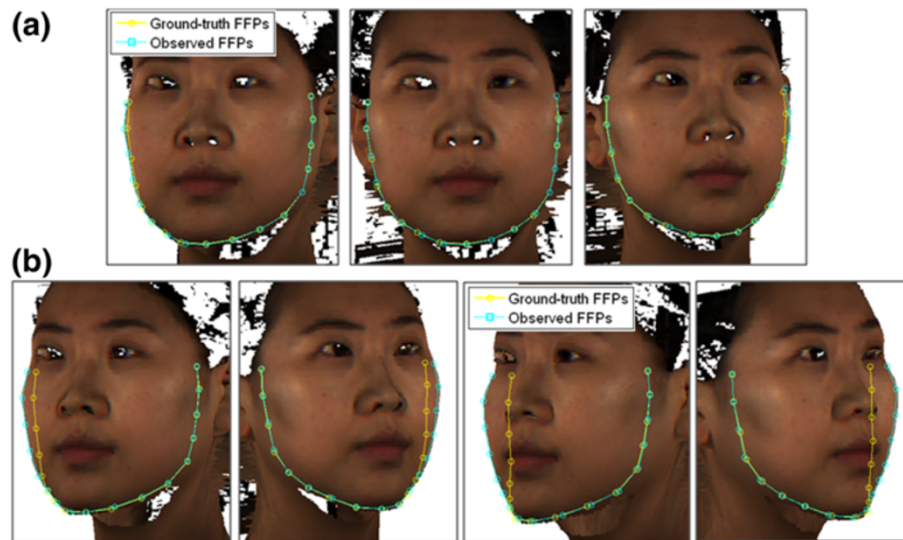


Figure 2 Self-occlusion errors increase as the degree of the head rotation increases: (a) almost frontal views ($0^\circ, \pm 15^\circ$); (b) highly rotated views ($\pm 30^\circ, \pm 45^\circ$).

shape. However, pose estimation by S3DMM is vulnerable to self-occlusion because the head pose is obtained from the relationship between the projected 3D FFPs of the shape model and the 2D FFPs detected in the image. As shown in Figure 1, self-occlusion errors may occur between the projected 3D FFPs and the observed 2D FFPs in a highly rotated face image. Consequently, these errors lead to lower accuracy results during head pose estimation with S3DMM.

Therefore, instead of using the pose estimator included in S3DMM, we employed a cylindrical head model to estimate the pose. The estimated pose is used for self-occlusion detection. The cylindrical head model, proposed by Ohue et al. [23], was used to detect the direction of a driver's face in a real-time system. This method is based on the assumption that a human head is basically cylindrical in shape, as shown in Figure 4a. The head pose is calculated simply by using three facial

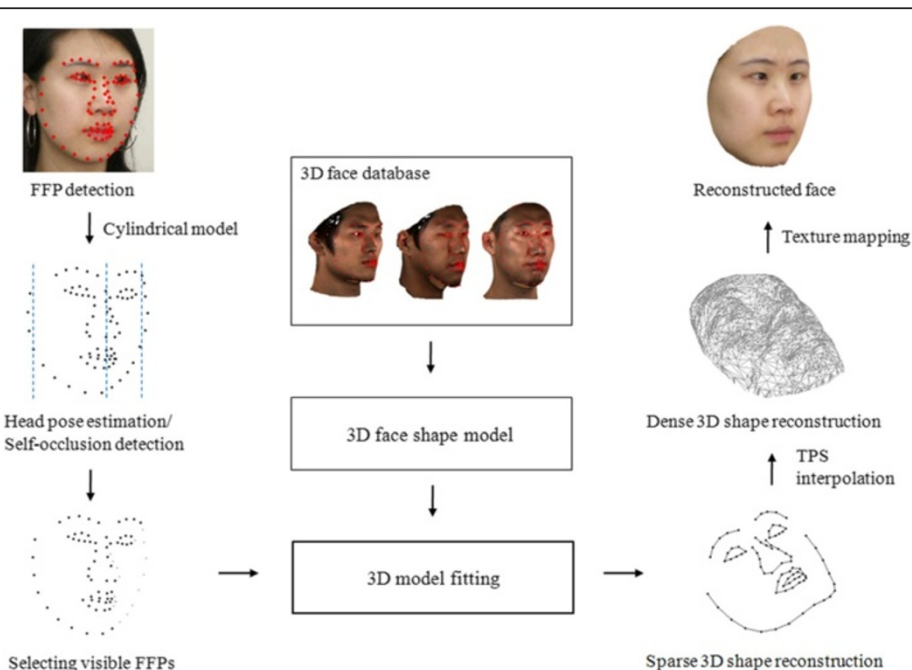
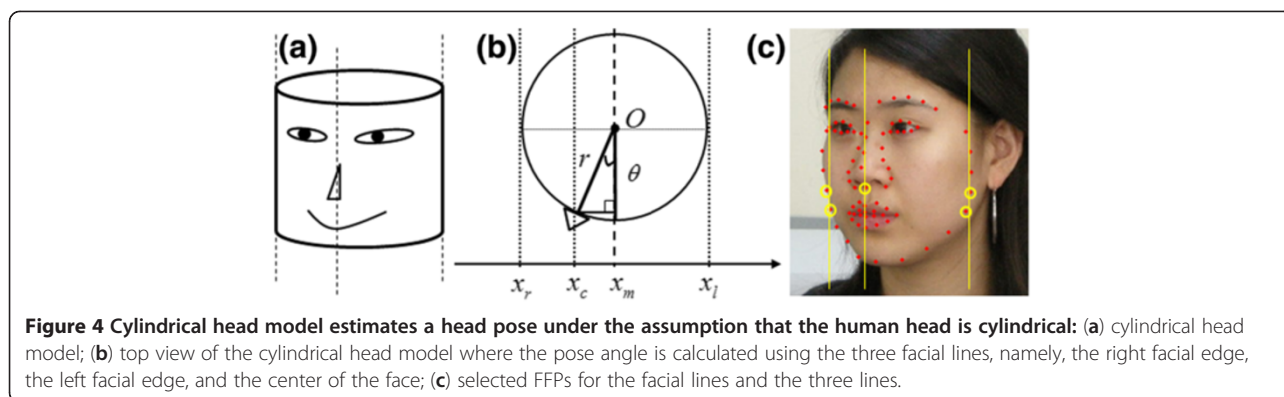


Figure 3 Overall procedure of the proposed method.



lines, namely, the right facial edge, left facial edge, and the center line of the face, as shown in Figure 4b. The yaw angle θ of the face is calculated using the following equation:

$$\theta = \arcsin\left(\frac{x_m - x_c}{r}\right) \quad (3)$$

where r is the radius of the cylinder, x_m is the x -coordinate of the cylinder center line and x_c is the x -coordinate of the facial center line. Here, the radius r and the center line x_m are obtained from the x -coordinates of the right and left facial edges, i.e., $r = (x_l - x_r)/2$ and $x_m = (x_l + x_r)/2$, respectively, as shown in Figure 4b. When detecting the facial edge lines x_l and x_r in a rotated view, the cylindrical head model does not require the occluded facial edge line, and instead it uses the newly observed facial edge line of the rotated view, as shown in Figure 5. Therefore, this method is less affected by self-occlusion.

As shown in (3), the performance of this method depends on how accurately the three lines are detected in the face image. Ohue et al. [23] used image processing methods, such as a Sobel filter and histogram analysis, to detect the lines. However, in our study, the three lines are obtained from the 2D FFPs which are manually annotated or automatically detected using the Active Appearance Models (AAMs) algorithm [32]. The selected FFPs and the three facial lines are shown in

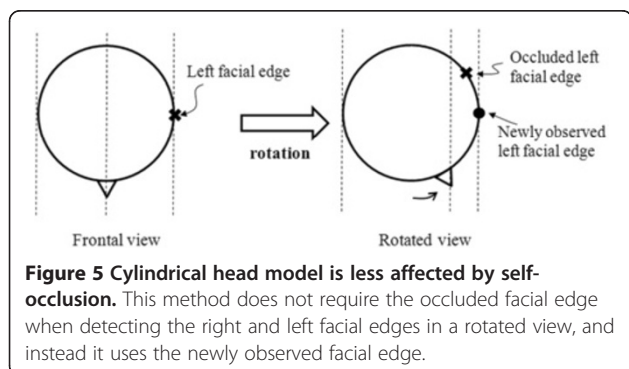


Figure 4c, where the facial center line is on the center point of the bottom of nose while the facial right and left boundaries are obtained from the midpoint between the two points located on each side of the facial contour, respectively.

Determination of visible FFPs

In this section, we present a method for automatically discriminating the visible FFPs and occluded FFPs by using the estimated head pose. The human head has a 3D structure but its shape varies slightly from person to person. Thus, the set of visible FFPs is inconsistent even with the same head pose and it changes depending on the individual. However, the individual differences between the sets of visible FFPs are not very large, and therefore, we use the generic visible FFP set for each yaw angle. This set is defined by manually analyzing the training set of 3D FFPs annotated in the 3D face scans. When defining the generic set of the visible FFPs for each pose, a point is excluded as occluded in the generic set of visible FFPs if it is occluded in the face of a certain individual. In this manner, we can prepare an FFP index table that contains the indices of the visible FFPs for each head pose. Figure 6 shows the defined visible FFPs (as crosses) and the occluded FFPs (as circles) for each yaw angle. In this study, the yaw angles are quantized as seven discrete angles, which range from -45° to $+45^\circ$ at intervals of 15° , because the visible FFPs do not change greatly over a 15° interval.

Finally, given the estimated head pose, the pose angle is quantized into one of the seven discrete angles, and a masking matrix \mathbf{M}_θ is obtained from the index table of the visible FFPs related to the quantized angle θ , as shown in Figure 7. Each component of the masking matrix \mathbf{M}_θ represents a visible FFP as 1 and an occluded FFP as 0:

$$\mathbf{M}_\theta = [\mathbf{m}_1 \mathbf{m}_2 \mathbf{m}_3 \dots \mathbf{m}_n]$$

$$\mathbf{m}_i = \begin{cases} [1 \ 1]^T, & \text{if } i^{\text{th}} \text{ point is visible} \\ [0 \ 0]^T, & \text{if } i^{\text{th}} \text{ point is occluded} \end{cases} \quad (4)$$

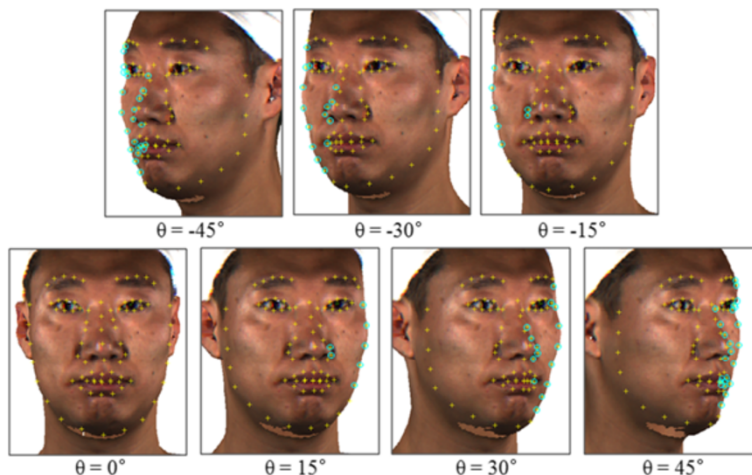


Figure 6 Generic visible FFP sets for seven head poses (cross: visible FFP; circle: occluded FFP).

where $i = 1, 2, \dots, n$, and n is the number of vertices. We used 80 vertices in this study. \mathbf{M}_θ is a 2×80 binary matrix, the pose angle θ has one of the seven discrete values, and \mathbf{m}_i is a 2×1 column-vector.

3D Face model fitting

A detailed description of the proposed model fitting method is shown in Algorithm 2. This is a modified version of the earlier S3DMM-based algorithm mentioned in Section “S3DMM”. The proposed model fitting scheme is based on the selected visible FFPs, which

eliminates the self-occlusion effect. As a result, the cost function of (2) is modified as follows:

$$\arg \min_{\beta, \mathbf{R}_\theta, \mathbf{T}} \|\mathbf{M}_\theta^\circ (\mathbf{P}(\mathbf{R}_\theta \tilde{\mathbf{S}} + \tilde{\mathbf{T}}) - \tilde{\mathbf{s}}_{2d})\|^2 \quad (5)$$

where the symbol “ \circ ” represents the Hadamard product, which is known as entry-wise multiplication [33], while \mathbf{M}_θ is the masking matrix at rotation angle θ . \mathbf{M}_θ is obtained from the index table of the visible FFPs for the estimated pose, as explained in Sections “Head pose estimation” and “Determination of visible FFPs”. We can

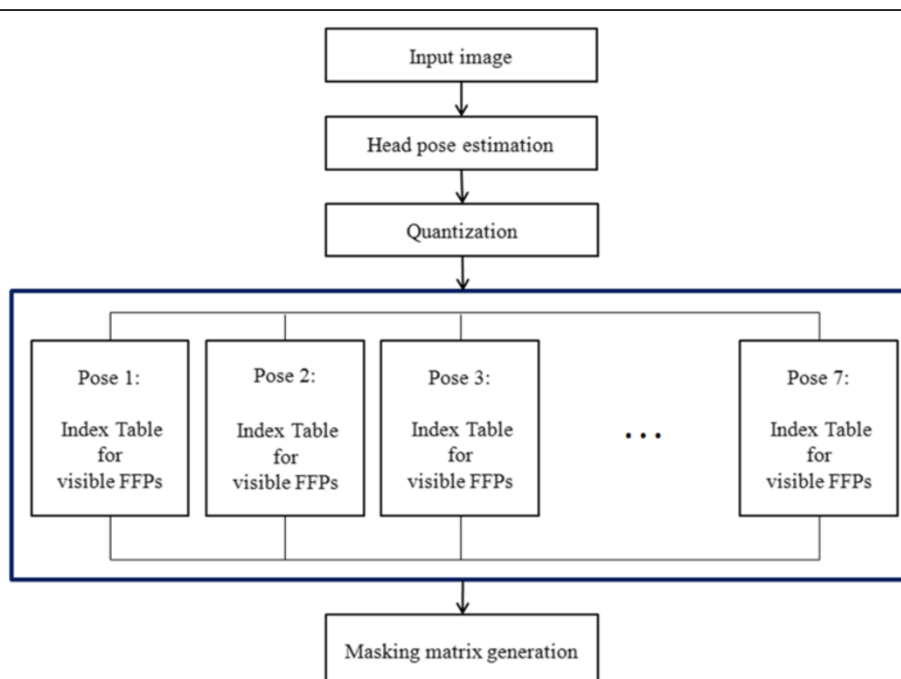


Figure 7 Masking matrix generation.

calculate the shape residual between the visible FFPs of the shape model and the input 2D facial shape using this masking matrix. The shape parameter β and pose parameter $(\mathbf{R}_\theta, \mathbf{T})$ can be obtained without any self-occlusion effect by minimizing this shape residual. The proposed 3D model fitting algorithm has the following two advantages compared with the previous method:

- 1) The pose angle $\hat{\theta}$ estimated by the cylindrical model is used for the pose parameter initialization. Therefore, the parameter estimation starts from a relatively exact initial pose parameter, which enhances the 3D face reconstruction performance. During the alignment step, an accurate alignment result is obtained by aligning the input 2D FFPs with the FFPs of the 2D mean shape, which are obtained by rotating the 3D mean shape (\mathbf{s}_0) from 0° to $\hat{\theta}$ and projecting it onto the x - y plane.
- 2) 3D model fitting is achieved on the basis of the visible FFPs by using the masking matrix. Therefore, the proposed method can reconstruct 3D faces that are less affected by self-occlusion.

Algorithm 2 Proposed 3D Model Fitting

Input: $\mathbf{s}_{2d} = (x_1, y_1, x_2, \dots, y_n)^T$

Output: $\mathbf{S}_{3d} = (X_1, Y_1, Z_1, X_2, \dots, Y_n, Z_n)^T$

1. Initialization
 Set $\theta_0 = \hat{\theta}$, $\mathbf{T}_0 = \mathbf{0}$, $\mathbf{M} = \mathbf{M}_{\hat{\theta}}$, and $k = 1$.
 ($\hat{\theta}$ is the angle estimated by the cylindrical model and $\bar{\theta}$ is a quantized angle of θ_0 .)
2. Alignment
 \mathbf{s}_{2d} is aligned with the 2D mean shape produced by rotating the 3D mean shape (\mathbf{s}_0) from 0° to $\hat{\theta}$ and projecting it onto the x - y plane.
3. Update the shape parameter with the fixed pose parameter
 $\arg \min_{\beta_k} \|\mathbf{M}^\circ(\mathbf{P}(\mathbf{R}_\theta \tilde{\mathbf{S}} + \tilde{\mathbf{T}}) - \tilde{\mathbf{s}}_{2d})\|_{(\mathbf{R}_\theta, \mathbf{T}) = (\mathbf{R}_\theta, \mathbf{T})_{k-1}}^2$
 Update \mathbf{R}_θ and \mathbf{T} with the fixed shape parameter
 $\arg \min_{(\mathbf{R}_\theta, \mathbf{T})_k} \|\mathbf{M}^\circ(\mathbf{P}(\mathbf{R}_\theta \tilde{\mathbf{S}} + \tilde{\mathbf{T}}) - \tilde{\mathbf{s}}_{2d})\|_{\beta = \beta_k}^2$
4. Reconstruct $(\mathbf{S}_{3d})^k$ using the shape parameter β_k
5. Verify whether
 $\|\mathbf{M}^\circ(\mathbf{P}(\mathbf{R}_\theta (\tilde{\mathbf{S}}_{3d})_k + \tilde{\mathbf{T}}) - \tilde{\mathbf{s}}_{2d})\|_{(\mathbf{R}_\theta, \mathbf{T}) = (\mathbf{R}_\theta, \mathbf{T})_k} < \epsilon$
 If not, go to Step 3 and $k = k + 1$
6. $\beta = \beta_k$ and $(\mathbf{R}_\theta, \mathbf{T}) = (\mathbf{R}_\theta, \mathbf{T})_k$
7. Reconstruct \mathbf{S}_{3d} using the final shape parameter

Dense 3D facial shape and texture mapping

A sparse 3D facial shape is produced as the reconstruction result after model fitting. Therefore, we have to

perform interpolation to produce a dense 3D facial shape. Interpolation is achieved by mapping a generic dense mean shape onto the reconstructed sparse 3D facial shape using the TPS algorithm [24,25]. The TPS mapping function is designed by learning the relationship between the FFPs of the generic 3D mean shape and the reconstructed 3D facial shape, which is similar to [24]. Let \mathbf{u} be the FFPs of the generic mean shape and $F(\mathbf{u})$ be the FFPs of the reconstructed facial shape. The mapping function is then:

$$F(\mathbf{u}) = c + \mathbf{A}\mathbf{u} + \mathbf{W}^T s(\mathbf{u}) \quad (6)$$

where c represents a translation, \mathbf{A} is a rotation, \mathbf{W} is the non-linear deformation, and $s(\mathbf{u})$ is a spline function. The mapping function $F(\mathbf{u})$ is then used to transform all of the other vertices in the mean shape, which produces adapted dense 3D face.

Finally, an available texture in the input 2D image is mapped onto the dense 3D facial shape to complete the 3D face reconstruction. However, some facial texture regions can be self-occluded in rotated view images. Thus, the bilateral symmetry of a face is used for texture mapping to recover the self-occluded texture. Specifically, we produce a mirrored image of the half-face contained in the visible region, and use the mirrored texture of the visible half-face for texture mapping in the occluded facial region.

Experimental results

Face database

For the experiments, we acquired 86 male and 64 female 3D face scans using a 3D laser scanner (*Cyberware 3030* RGB model) [34]. To construct 3D shape models, each 3D face scan was manually annotated with 80 FFPs, and the annotated 3D FFPs were aligned with others by Procrustes analysis [22]. An iterative alignment method was applied to our data to produce accurate aligned data. A reference face scan was randomly selected and 3D FFPs in the remaining face scans were aligned with 3D FFPs in the reference face scan. The mean FFPs of the aligned data became the new reference FFPs in the next iteration, and this process was repeated until the mean FFPs stopped changing. The final aligned 3D FFPs were used to build a 3D shape model in the training stage. The training and test data-sets were divided using the leave-one-out methodology, which uses a single data as the test data and the remaining data as the training data.

2D face images of 150 subjects were obtained by projecting the textured 3D face scans as test images. In other words, we obtained face images of seven different head poses by rotating the face scans at 15° intervals in the range -45° to $+45^\circ$ and projecting them onto the x - y plane. Therefore, the total number of test images was 1050 (150 subjects \times 7 poses). The resolution of the

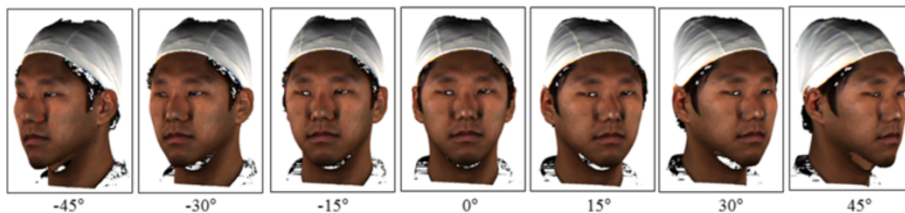


Figure 8 Test image samples for the seven head poses.

obtained 2D images is 1200×900 pixels and the size of the facial region was approximately 350×350 pixels. Figure 8 shows examples of the 2D face test images obtained. Given a test image, 2D FFPs were detected for 3D model fitting in S3DMM. In the experiments, we used two methods to detect 2D FFPs in the test images for ideal and practical cases. In the ideal case, Test data 1 was created by combining the ground-truth 2D FFPs in the visible facial region and the manually marked 2D FFPs (observed 2D FFPs) for the self-occluded facial regions, as shown in Figure 9a. The ground-truth 2D FFPs were acquired by rotating the ground-truth 3D FFPs (manually annotated points on a face scan) and projecting them onto the x - y plane. In the case of a rotated facial image, it is difficult to detect the ground-truth 2D FFPs on the self-occluded region because they are not visible. Therefore, the manually marked 2D FFPs (observed 2D FFPs) were used in the self-occluded facial regions. Test data 1 could be regarded as ideal data because all of the FFPs were obtained manually. In the

practical case, Test data 2 was derived by automatically detecting the FFPs using the AAMs fitting algorithm, as described in Figure 9b. The AAMs used in this work is based on the simultaneous inverse compositional algorithm [32]. These data may have AAMs fitting errors as well as errors caused by self-occlusion.

We used the root mean squared error (RMSE) to measure the similarity between the reconstructed facial shape and the corresponding ground-truth facial shape. Let the reconstructed facial shape vector be $\mathbf{S}_{re} = [X'_1, Y'_1, Z'_1, X'_2, \dots, Y'_m, Z'_m]^T$ and the corresponding ground-truth facial shape vector be $\mathbf{S}_{gt} = [X_1, Y_1, Z_1, X_2, \dots, Y_m, Z_m]^T$. Given these two shape vectors, the RMSE is calculated as follows:

$$e = \frac{1}{n} \sum_{i=1}^n \sqrt{(X_i - X'_i)^2 + (Y_i - Y'_i)^2 + (Z_i - Z'_i)^2} \quad (7)$$

where n is the number of FFPs.

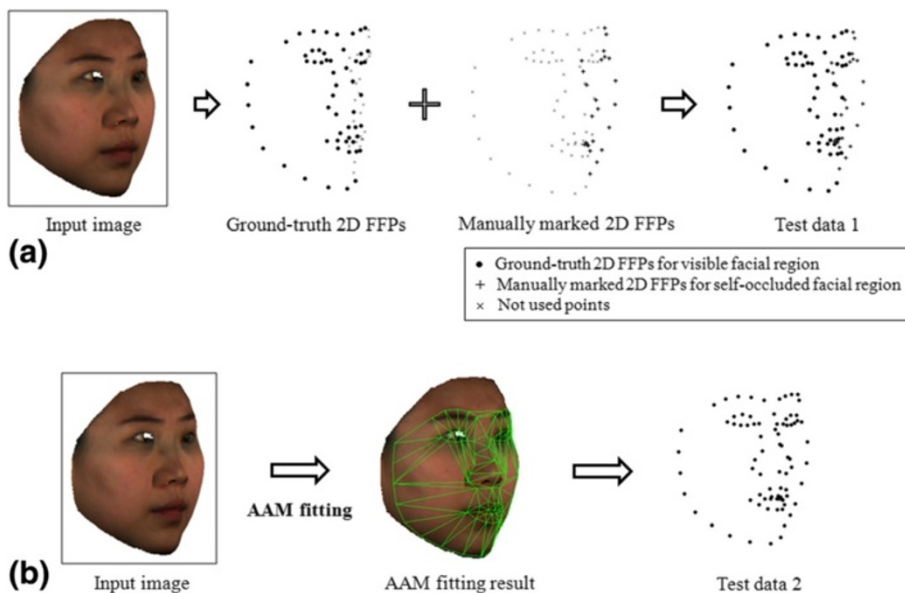


Figure 9 Procedure for generating the two types of test data: (a) Test data 1 is manually obtained by combining the ground-truth 2D FFPs for the visible facial regions and the manually marked 2D FFPs for the self-occluded facial region; (b) Test data 2 is derived by automatically detecting the FFPs using the AAMs fitting algorithm.

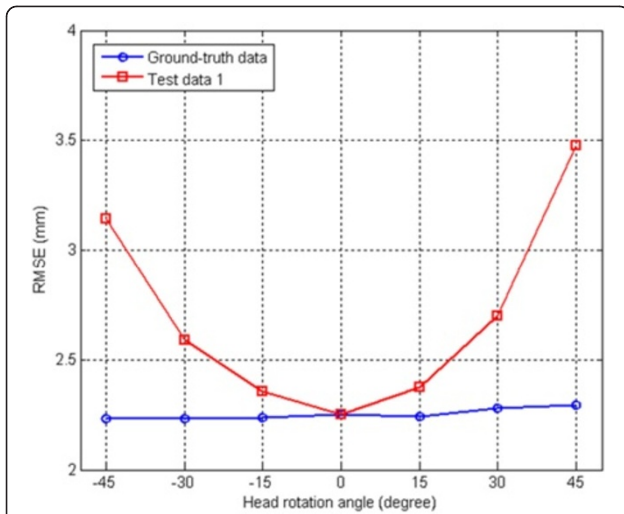


Figure 10 Performance comparisons of the previous S3DMM with the ground-truth data and Test data 1 according to the pose of the input face.

Effect of self-occlusion on 3D face reconstruction

To demonstrate the effect of self-occlusion on the performance of the previous S3DMM method, we performed 3D face reconstruction using the S3DMM algorithm found in Section “S3DMM” with the ground-truth data and Test data 1. The ground-truth data was obtained by projecting the ground-truth 3D FFPs onto the x - y plane, and hence no errors were caused by self-occlusion, whereas Test data 1 contained errors, as shown in Figure 9a. Figure 10 shows the RMSEs of the reconstructed facial shape for the two types of test data. As shown in Figure 10, the RMSE increased sharply when using Test data 1

because the degree of head rotation increased, and thus the location errors in the 2D FFPs in the occluded facial region also increased. In the highly rotated views ($\pm 45^\circ$), the RMSE differences between the two data sets were greater than approximately 1 mm. Both data-sets had the same RMSE in the frontal view (0°), because the frontal view was not affected by self-occlusion. The results showed that the 2D FFP location errors caused by self-occlusion severely deteriorated the reconstruction performance of the earlier S3DMM algorithm.

Head pose estimation performance

This section details the performance of the cylindrical model-based pose estimation method compared with QR decomposition-based head pose estimation, which was used in [5], and the effect of using the estimated pose as the initial pose parameter on the proposed 3D reconstruction method. Figure 11 and Table 2 show the mean absolute errors (MAEs) with the two different head pose estimation methods using Test data 1 and 2. The results clearly show the following:

- 1) The cylindrical model-based method showed better performance with the highly rotated face images ($\pm 30^\circ$, $\pm 45^\circ$) because this method used facial features that were less sensitive to self-occlusion.
- 2) The QR decomposition-based method showed slightly better performance with the almost frontal face images ($\pm 15^\circ$, 0°) because the FFP location error in these images caused by self-occlusion was small and this method used a relatively larger number of facial features. However, these facial features were more sensitive to self-occlusion because the

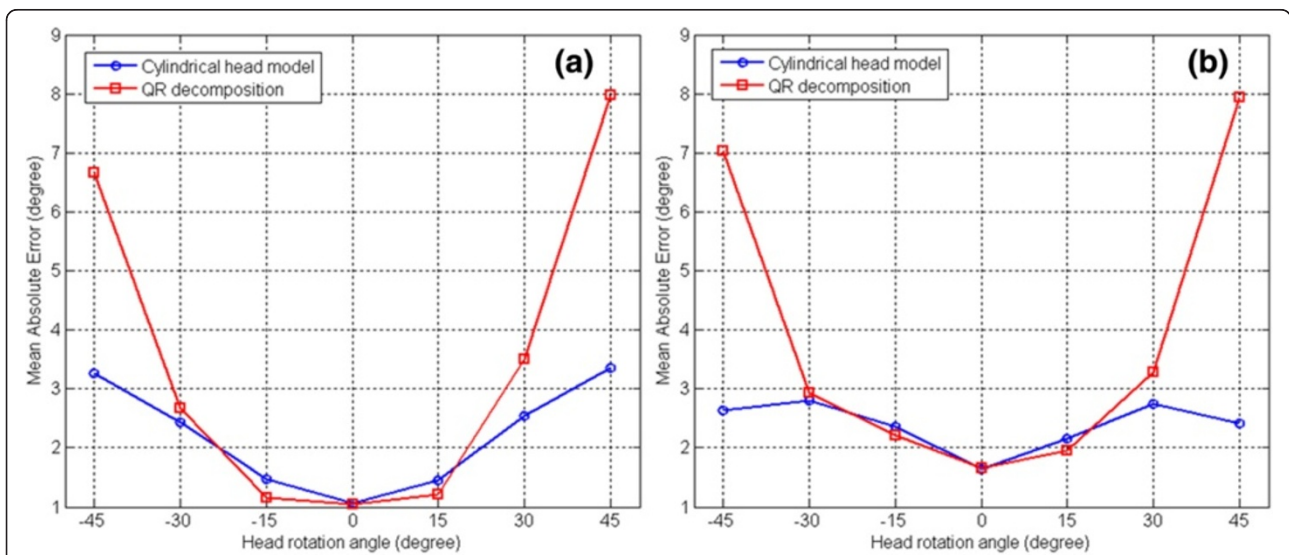


Figure 11 Performance comparisons of the different head pose estimation methods: (a) Test data 1; (b) Test data 2.

Table 2 Performance comparisons of the two different methods for head pose estimation

Data set	Test data 1		Test data 2	
Method	Cylindrical model (proposed method)	QR decomposition [5]	Cylindrical model (proposed method)	QR decomposition [5]
MAE	2.23°	3.46°	2.39°	3.86°

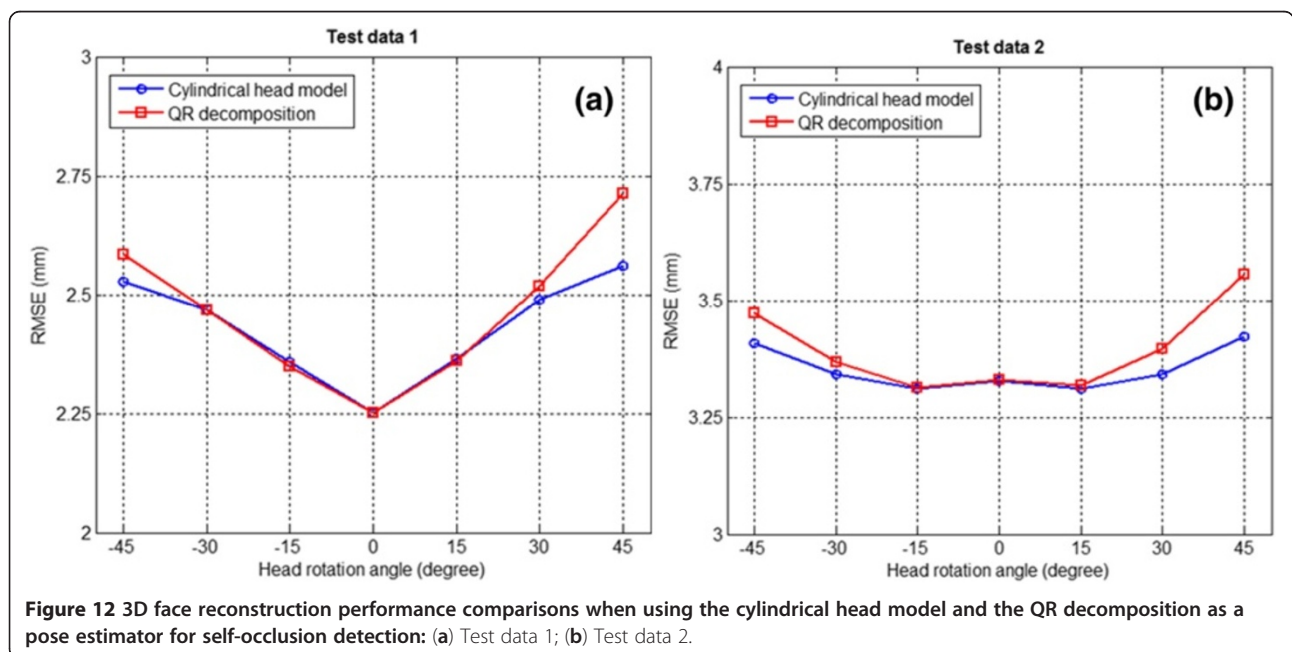
performance of this method was severely degraded in the highly rotated images.

- The results were similar irrespective of the data set that was used, but the pose estimation error with Test data 2 was slightly larger than that with Test data 1 because the automatically obtained test data had more severe FFP location errors than the manually obtained test data. However, even in such case, the proposed 3D reconstruction method showed better performance than the previous method. This result will be discussed in Section “Quantitative results with yaw variations”. In order to remove noisy FFPs, we need a different approach such as texture analysis around each FFP.

To evaluate the effects of different head pose estimation methods on 3D face reconstruction, we applied QR decomposition and the cylindrical head model to our proposed S3DMM algorithm seen in Section “3D face model fitting”, respectively. As shown in Figures 12a,b, the two methods showed similar performances in the -15° to $+15^\circ$ pose range, but the performance of the cylindrical model is better than that of the QR decomposition method at $\pm 30^\circ$ and $\pm 45^\circ$, regardless of the test data-set. The method based on QR decomposition showed better pose estimation performance with the

almost frontal face images, but this method did not show a better 3D face reconstruction performance, as shown in Figure 12, because the low pose estimation error was compensated for by our proposed model fitting algorithm from Section “3D face model fitting”. Note that the cylindrical model-based method showed considerably better pose estimation performance with highly rotated face images, which led to an improvement in 3D face reconstruction performance.

In this work, the yaw angle estimated by the cylindrical model was used as the initial pose parameter ($\theta_0 = \hat{\theta}$), as shown in the initialization step of the proposed 3D model fitting algorithm in Section “3D face model fitting”. To demonstrate the effect of using the estimated pose as the initial pose parameter, we further obtained results to compare the following two cases with Test data 1 and 2: the first case used the estimated pose whereas the second case did not. In the first case, 3D model fitting began with shape parameter estimation because the approximate initial pose was known, as shown in Algorithm 2 in Section “3D face model fitting”. In the second case, the initial pose parameter was not provided, and hence the pose parameter had to be estimated prior to shape parameter estimation, as shown in Step 3 of Algorithm 1. As shown in Figures 13a,c, both cases produced similar RMSEs with the almost frontal views



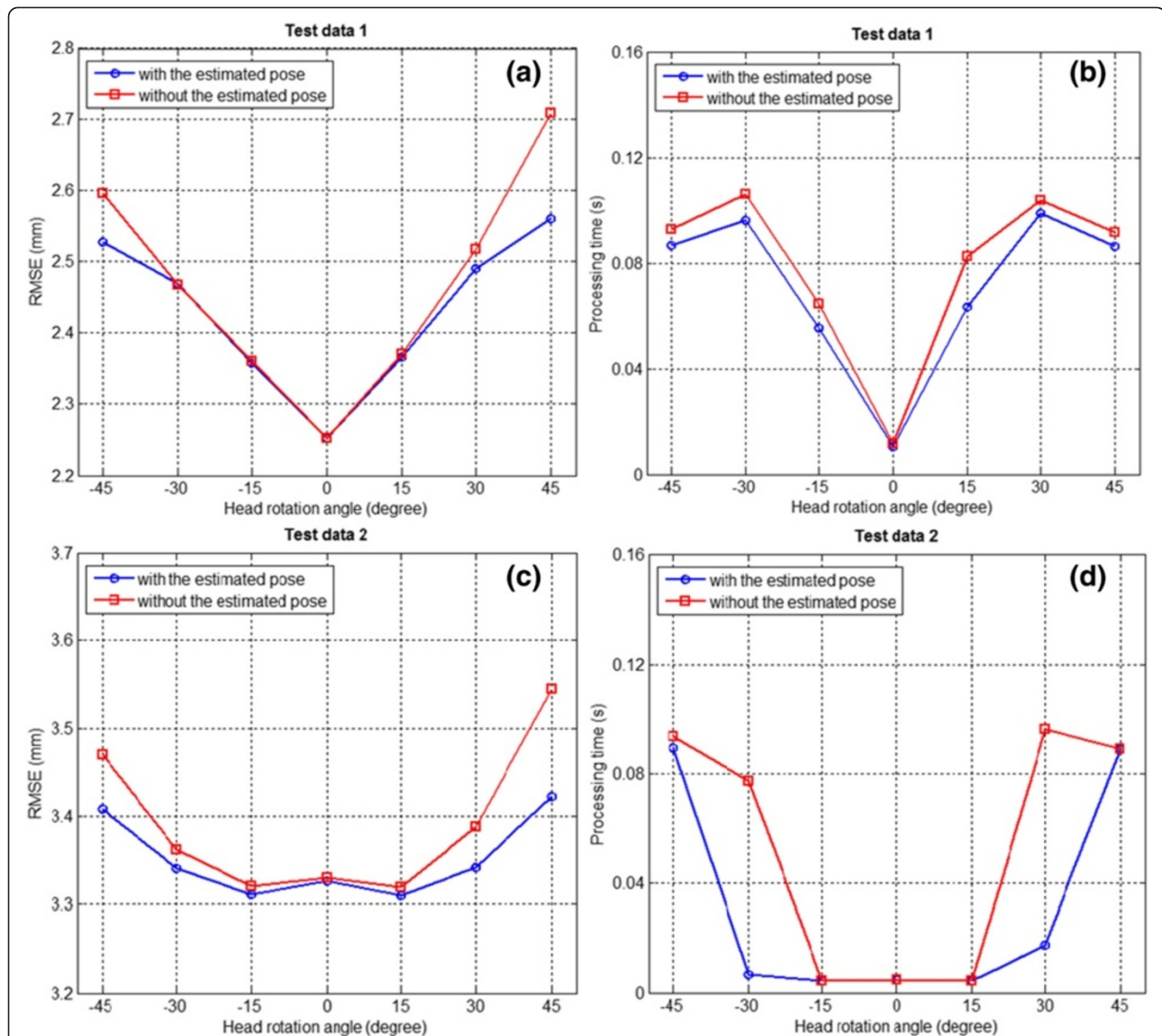


Figure 13 Performance comparisons with the estimated pose and without the estimated pose: (a) 3D face reconstruction performance with Test data 1; (b) processing time with Test data 1; (c) 3D face reconstruction performance with Test data 2; (d) processing time with Test data 2.

(-15° to $+15^\circ$). However, the first case showed the better performance with the highly rotated views ($\pm 30^\circ$, $\pm 45^\circ$), irrespective of the test data-set used. As shown in Figures 13b,d, the use of the estimated pose was also more efficient than not using the estimated pose in terms of the processing time. Based on these observations, we found that an accurate initial pose parameter led to a performance improvement during 3D face reconstruction based on our proposed algorithm and reduced the number of iterations required for 3D model fitting.

In summary, the cylindrical model-based method is suitable for selecting a reliable set of visible FFPs while the use of the estimated yaw angle as the initial pose

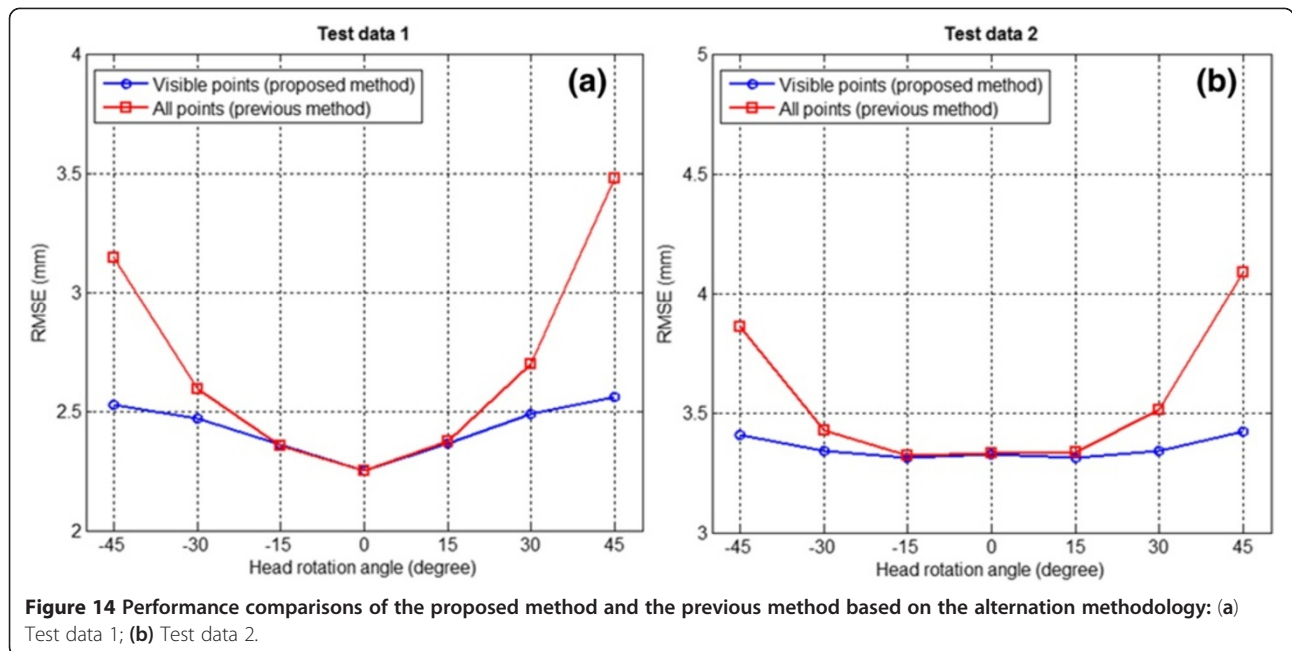
parameter indicates that the proposed 3D model fitting algorithm shows better reconstruction performance with higher efficiency.

Quantitative performance of the proposed method

Quantitative results with yaw variations

To evaluate the performance of our proposed method, we obtained results from the two test data sets with yaw variations using the proposed algorithm and we compared its performance with that of the previous S3DMM algorithm found in Section "S3DMM".

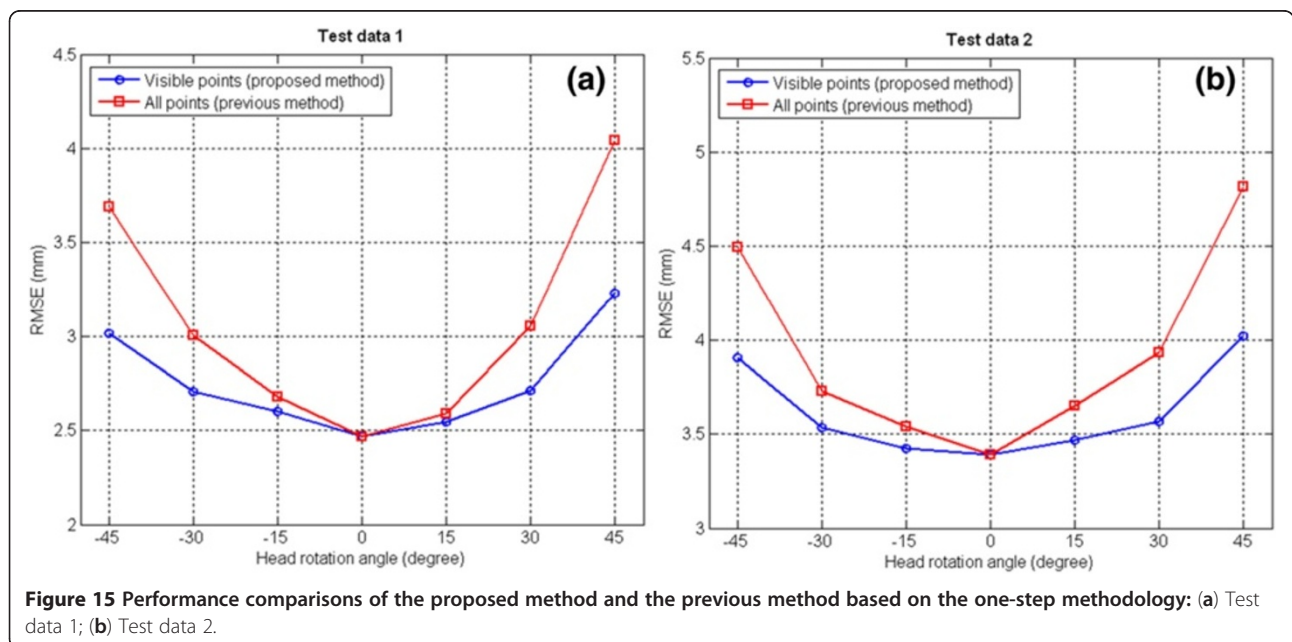
Figure 14 shows the RMSEs of the proposed method and the previous method with Test data 1 and 2. As shown in Figure 14, the proposed algorithm produced a



significantly lower RMSE than the previous method as the head rotation increased. This was because the FFP self-occlusion errors increased with the head rotation. Figure 14b showed that although Test data 2 contained self-occlusion errors and detection errors, because they were obtained automatically using the AAMs algorithm, the results showed that the proposed method provided superior performance compared to the previous method. However, if detection errors of the FFPs were very large, the reconstruction performance of the proposed method could be significantly degraded because these FFPs still

contained large detection errors, although self-occlusion errors were excluded by the proposed method. This problem of FFP detection is a common limitation of S3DMM-based methods.

We also obtained results from Test data 1 and 2 using the one-step parameter estimation method. The one-step methodology has been used for model parameter estimation by the S3DMM-based method, which simultaneously estimates the shape and pose parameters using a non-linear optimization algorithm [3,17,19]. In this study, one-step parameter estimation was achieved using



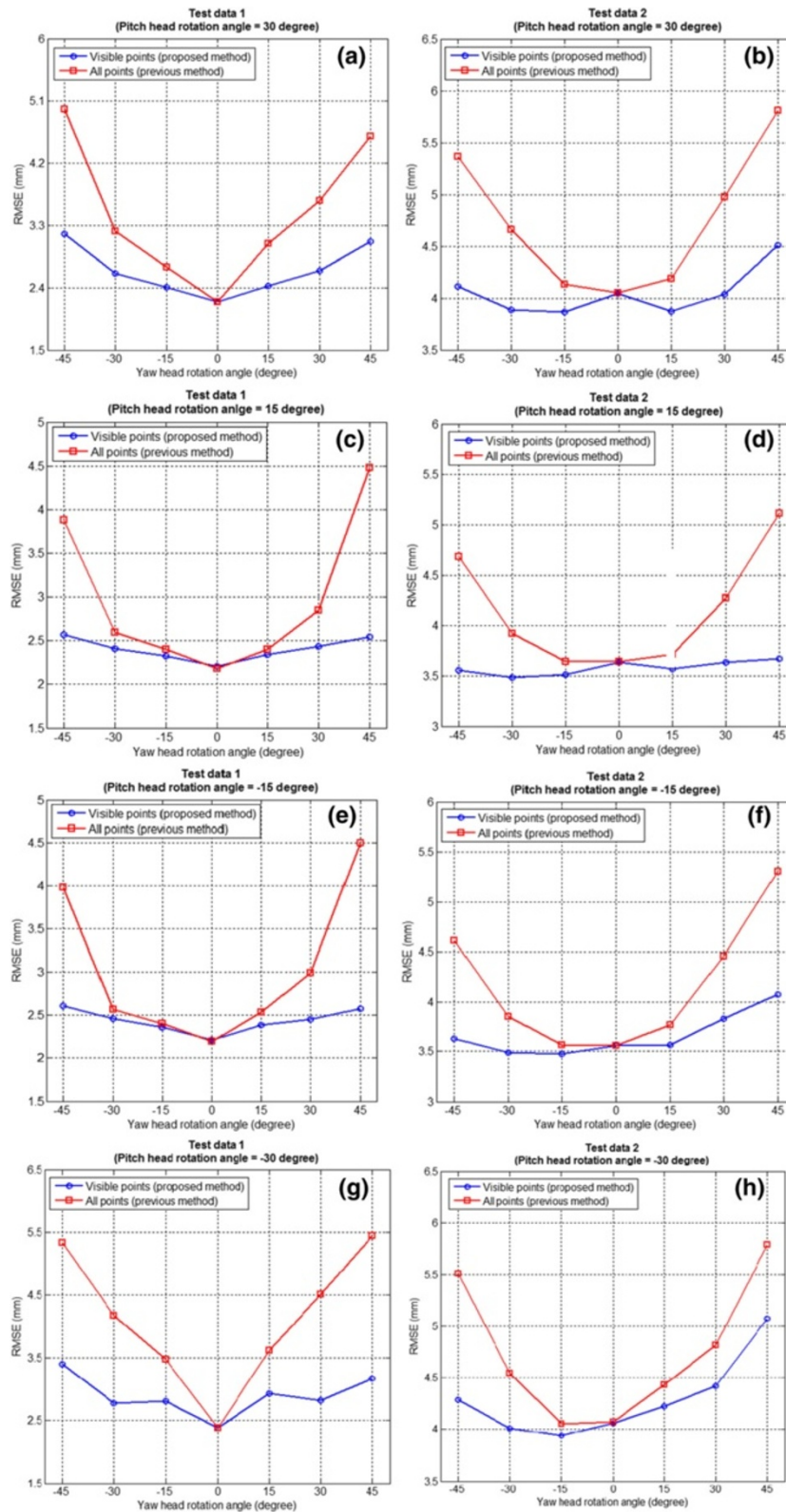


Figure 16 Performance comparisons of the proposed method and the previous method with yaw and pitch variations: (a, b) Test data 1 and 2 at 30° pitch angle; (c, d) Test data 1 and 2 at 15° pitch angle; (e, f) Test data 1 and 2 at -15° pitch angle; (g, h) Test data 1 and 2 at -30° pitch angle.



Figure 17 Examples of faces reconstructed using Test data 1 (a, b: previous method and proposed method at 15°; c, d: previous method and proposed method at 30°; e, f: previous method and proposed method at 45°; red solid lines represent the accurate frontal and profile shapes of the ground-truth face)



Figure 18 Examples of faces reconstructed using Test data 2 (a, b: previous method and proposed method at 15°; c, d: previous method and proposed method at 30°; e, f: previous method and proposed method at 45°; red solid lines represent the accurate frontal and profile shapes of the ground-truth face).

code available in MATLAB [35]. We compared the one-step methodology using all FFPs (previous method) with one using only visible FFPs (proposed method). Figure 15 shows the RMSEs for both methods. In the highly rotated views ($\pm 30^\circ$, $\pm 45^\circ$), the proposed method using selected visible FFPs showed considerably better performance than the previous method using all of the FFPs with both test data-sets. With the almost frontal views (0° , $\pm 15^\circ$), both methods delivered similar reconstruction performances. From the observations in Figures 14 and 15, the proposed strategy of using the visible FFPs was an adequate solution for the self-occlusion problem during S3DMM-based 3D face reconstruction.

Quantitative results with yaw and pitch variations

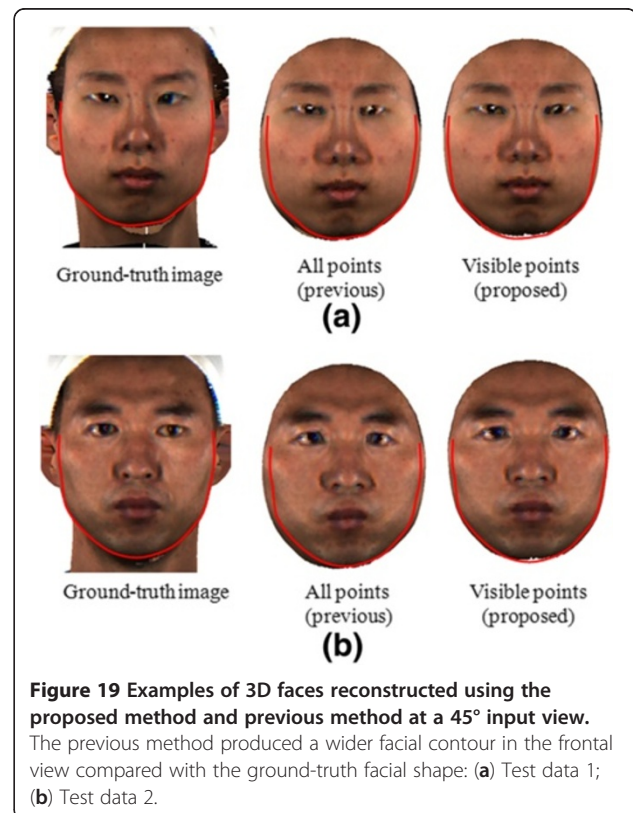
In this experiment, we tested whether yaw and pitch variations affected the performance of the proposed method. For this experiment, face images were generated with yaw and pitch variations by rotating and projecting the textured face scans of 150 subjects, as explained in Section “Face database”. A head pose set consisted of 28 different poses, which were a combination of seven yaw angles (0° , $\pm 15^\circ$, $\pm 30^\circ$, $\pm 45^\circ$) and four pitch angles ($\pm 15^\circ$, $\pm 30^\circ$). The total number of the test images is 4200 (150 subjects \times 28 poses). To evaluate the performance of the proposed method using these test images, Test data 1 and 2 were obtained from the test images, as explained in Section “Face database”, and the performance of the proposed algorithm from Section “3D face model fitting” was compared with that of the previous algorithm seen in Section “S3DMM”. Figure 16 shows the effects of yaw and pitch variations on the performance of the proposed method. In this figure, each graph shows the RMSEs of the proposed method and the previous method with yaw variations at a fixed pitch angle. From the results, it is found that the performance of the proposed method was remarkably improved compared to that of the previous method even with yaw and pitch variations. This was because the self-occlusion error caused by yaw variations was more dominant than that caused by pitch variations. Consequently, it is found that the occlusion error occurred by a combination of yaw and pitch could be compensated for using the proposed algorithm by considering only yaw variations, as shown in Figure 16. However, the performance of the proposed method degraded slightly when both head yaw and tilt occurred because the cylindrical model could only estimate head yaw. In order to improve the performance of the proposed method when both head yaw and tilt occur, we need another tilt estimator.

Qualitative performance of the proposed method

We obtained qualitative results for the proposed method and the previous method using Test data 1 and 2.

Figure 17 shows the reconstructed results for four subjects from our database when using Test data 1. The first column shows the ground-truth images while the a-, c-, and e- columns show the reconstructed faces using the previous method when the yaw angle of the input face was 15° , 30° , and 45° , respectively. The b-, d-, and f- columns show the reconstructed faces using the proposed method when the yaw angle of the input face was 15° , 30° , and 45° , respectively. In the same way, Figure 18 shows the faces reconstructed using Test data 2. As shown in Figures 17 and 18, the performance of the proposed method was almost the same as that of the previous method with a 15° head pose because the almost frontal view is negligibly affected by self-occlusion. However, we observed the following with the highly rotated views (30° , 45°):

- 1) The proposed method showed that the frontal and profile shapes were similar to the ground-truth facial shapes.
- 2) The previous method produced a wider facial contour in the frontal view compared with the ground-truth facial shape because self-occlusion error occurred, as shown in Figure 19. Figure 18 showed that the proposed method remarkably improved the reconstruction performance compared



with the previous method, even with the practical data (Test data 2).

Finally, we present the reconstruction results for real-world images from the FacePix database [36] and the CAS-PEAL-R1 database [37]. Figure 20 shows individual results for five subjects in the Face Pix database, where the rows indicate different subjects and the seventh column shows the test images. The first and fourth columns show different views of the original image. The a- and c- columns show the results with the previous method, whereas the b- and d- columns show the results with the proposed method. Figure 21 shows the results for five subjects in the CAS-PEAL-R1 database. From these figures, we can find that:

- 1) With highly rotated face images, the proposed method provided reconstructed facial shapes in the frontal and profile views that were closer to the ground-truth shape than the previous method were. In particular, the previous method had a wider facial contour in the frontal view compared with the ground-truth facial shape.

- 2) With almost frontal face images, the performance of the proposed method was similar to that of the previous method because self-occlusion error was very small, as shown in the fourth and fifth rows of Figure 20 and in the third, fourth, and fifth rows of Figure 21.

The proposed 3D face reconstruction required about 0.1 s per test image using our proposed algorithm based on an alternation methodology and about 3.6 s per test image using the algorithm based on a one-step methodology. The computation times were measured on an Intel Core i5 CPU 750, 2.7 GHz, 3 GB RAM machine.

Conclusions

We analyzed the self-occlusion problem that occurs in S3DMM-based 3D face reconstruction and proposed a method for solving this problem. Our main contributions are summarized as follows

- The 3D model fitting scheme of S3DMM was modified to make it suitable for 3D face reconstruction based on visible FFPs. The reconstruction accuracy of the proposed method

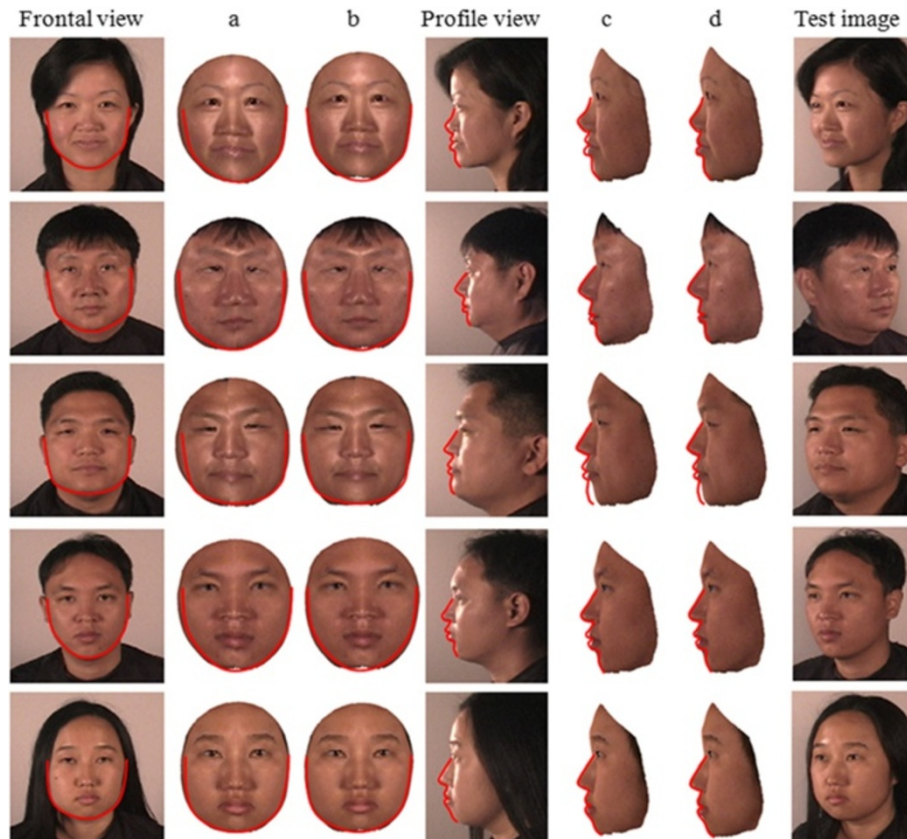


Figure 20 Reconstruction tests using subjects from the FacePix database.

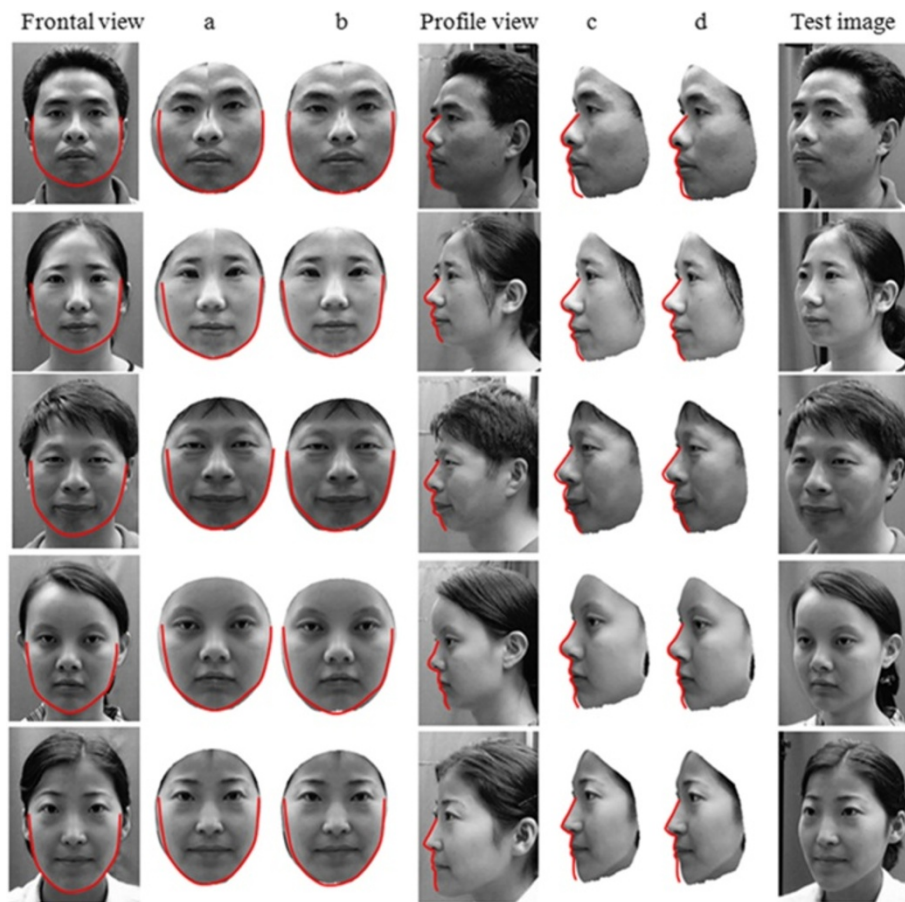


Figure 21 Reconstruction tests using subjects from the CAS-PEAL-R1 database.

was improved greatly compared with the original S3DMM-based method by using the visible FFPs having no self-occlusion errors.

- To exclude self-occluded FFPs in the 3D model fitting process, self-occluded FFPs were separated automatically from visible FFPs using a pose estimation method based on a cylindrical head model and an index table of visible FFPs.
- The reconstruction performance was enhanced by using the estimated pose as the initial pose parameter during the 3D model fitting process.

Since the proposed method can automatically reconstruct a 3D face from an arbitrary-view image, it can be applied for to a variety of useful applications, such as 3D game and animation character generation, and 2D frontal face generation from a side-view face image which can be used for monitoring suspects in surveillance cameras.

In future works, we would research about the 3D face reconstruction method robust to the detection error of

FFP by combining the information of FFP and facial texture. In addition, we would develop a more accurate 3D face reconstruction method for a wide range of pose variations and study a new method to solve the self-occlusion problem without eliminating occluded FFPs.

Abbreviations

SFS: shape-from-shading; PCA: principal component analysis; 3DMM: 3D morphable model; S3DMM: simplified 3D morphable model; FFPs: facial feature points; TPS: Thin Plate Spline; AAMs: Active Appearance Models; RMSE: root mean squared error; MAE: mean absolute error.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012-0005223).

Author details

¹School of Electrical and Electronic Engineering, Yonsei University, Seoul, Republic of Korea. ²Technical Research Institute, Hyundai Mobis, Yongin, Republic of Korea. ³Division of Electronics and Electrical Engineering, Dongguk University, Seoul, Republic of Korea.

Received: 14 January 2012 Accepted: 17 July 2012
Published: 17 August 2012

References

1. FI Parke, Proceedings of ACM National Conference, in *Computer generated animation of faces*, 1972
2. FI Parke, *A parametric model of human faces* (University of Utah, Salt Lake City, 1974). PhD thesis
3. V Blanz, T Vetter, Face recognition based on fitting a 3D morphable model. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(9), 1063–1074 (2003)
4. X Zhang, Y Gao, Face recognition across pose: a review. *Pattern Recognit.* **42**(11), 2876–2896 (2009)
5. U Park, Y Tong, AK Jain, *Face recognition with temporal invariance: a 3D aging model*. Proceedings of International Conference on Automatic Face and Gesture Recognition, 2008, pp. 1–7
6. U Park, Y Tong, AK Jain, Age invariant face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(5), 947–954 (2010)
7. *Example of game character generation*, http://fightnight.easports.com/featureFrame.action?id=Feature-TBA1&fType=video. Accessed 2 January 2012
8. A Maejima, S Wemler, T Machida, M Takebayashi, S Morishima, Instant casting movie theater: the future cast system. *IEICE Transactions on Information and Systems* **E91-D**(4), 1135–1148 (2008)
9. B Gökberk, AA Salah, N Alyüz, L Akarun, Handbook of Remote Biometrics for Surveillance and Security, in *3D face recognition: technology and applications* (Springer, 2009), pp. 217–246
10. R Zhang, P Tsai, JE Cryer, M Shah, Shape from shading: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(8), 690–706 (1999)
11. R Dobgard, R Basri, Statistical symmetric shape from shading for 3D structure recovery of faces. *LNCS on European Conference on Computer Vision* **3022**, 99–113 (2004)
12. A Ahmed, A Farag, T Starr, *A new symmetric shape from shading algorithm with an application to 3-D face reconstruction*. Proceedings of International Conference on Image Processing, 2008, pp. 201–204
13. M Castelan, WAP Smith, ER Hancock, A coupled statistical model for face shape recovery from brightness images. *IEEE Trans. Image Process.* **16**(4), 1139–1151 (2007)
14. M Castelan, JV Horebeek, Relating intensities with three-dimensional facial shape using partial least squares. *IET Computer vision* **3**(2), 60–73 (2009)
15. M Reiter, R Donner, G Langs, H Bischof, *3D and infrared face reconstruction from RGB data using canonical correlation analysis*. Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), 2006, pp. 425–428
16. Z Lei, Q Bai, R He, SZ Li, *Face shape recovery from a single image using CCA mapping between tensor spaces*. Proceedings of IEEE Conference on Computer vision and Pattern Recognition (CVPR 2008), 2008, pp. 1–7. June
17. V Blanz, T Vetter, *A morphable model for the synthesis of 3D faces*. Proceedings of SIGGRAPH, 1999, pp. 187–194
18. D Jiang, Y Hu, S Yan, L Zhang, H Zhang, W Gao, Efficient 3D reconstruction for face recognition. *Pattern Recognit.* **38**(6), 787–798 (2005)
19. S Wang, S Lai, Efficient 3D face reconstruction from a single 2D image by combining statistical and geometrical information. *LNCS on Asian Conference on Computer Vision* **3852**, 427–436 (2006)
20. C Wang, S Yan, H Li, H Zhang, M Li, Automatic, effective, and efficient 3D face reconstruction from arbitrary view image. *LNCS on Advances in Multimedia Information Processing – PCM* **3332**, 553–560 (2004). October
21. TF Cootes, GJ Edwards, CJ Taylor, Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6), 681–685 (2001)
22. I Matthews, S Baker, Active appearance models revisited. *Int. J. Comput. Vision* **60**(2), 135–164 (2004)
23. K Ohue, Y Yamada, S Uozumi, S Tokoro, A Hattori, T Hayashi, Development of a new pre-crash safety system. *SAE World Congress* (2006). April
24. U Park, AK Jain, *3D face reconstruction from stereo video*. Proceedings of the 3rd Canadian Conference on Computer and Robot Vision, 2006, p. 41. June
25. FL Bookstein, Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(6), 567–585 (1989)
26. P Watta, N Gandhi, S Lakshmanan, *An eigenface approach for estimating driver pose*. Proceedings of IEEE Conference Intelligent Transportation Systems, 2000, pp. 376–381
27. S Lakshmanan, P Watta, YL Hou, N. Gandhi, *Comparison between eigenfaces and fisherfaces for estimating driver pose*. Proceedings of IEEE Conference Intelligent Transportation Systems, 2001, pp. 889–894
28. P Watta, S Lakshmanan, YL Hou, Nonparametric approaches for estimating driver pose. *IEEE Trans. Veh. Technol.* **56**(4), 2028–2041 (2007)
29. J Wu, MM Trivedi, A two-stage head pose estimation framework and evaluation. *Pattern Recognit.* **41**(3), 1138–1158 (2008)
30. E Murphy-Chutorian, A Doshi, MM Trivedi, *Head pose estimation for driver assistance systems: A robust algorithm and experimental evaluation*. Proceedings of IEEE Conference Intelligent Transportation Systems, 2007, pp. 709–714
31. E Murphy-Chutorian, MM Trivedi, Head pose estimation in computer vision: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(4), 607–626 (2009)
32. S Baker, R Gross, I Matthews, *Lucas-Kanade 20 years on: a unifying framework: part 3, (Technical Report CMU-RI-TR-03-35)* (Carnegie Mellon University, Robotics Institute, 2003)
33. R.A. Horn, C.R. Johnson, (Cambridge University Press, Matrix Analysis, 1985)
34. *3D Scanner specification*, http://www.cyberware.com/products/pdf/headFace.pdf. Accessed 2 January 2012
35. T Boggs, JW Tolle, Sequential quadratic programming. *Acta Numerica* **1**, 1–51 (1996)
36. J Black, M Garghesha, K Kahol, P Kuchi, S Panchanathan, *A framework for performance evaluation of face recognition algorithms* (ITCOM, Internet Multimedia Systems II, Boston, 2002)
37. W Gao, B Cao, S Shan, X Chen, D Zhou, X Zhang, D Zhao, The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations. *IEEE Trans. SMC-Part A: Systems and Humans* **38**(1), 149–161 (2008)

doi:10.1186/1687-6180-2012-176

Cite this article as: Lee et al.: Single view-based 3D face reconstruction robust to self-occlusion. *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:176.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com