3D Face Modeling based on 3D Dense Morphable Face Shape Model

Yongsuk Jang Kim, Sun-Tae Chung, Boogyun Kim, and Seongwon Cho

Abstract—Realistic 3D face model is more precise in representing pose, illumination, and expression of face than 2D face model so that it can be utilized usefully in various applications such as face recognition, games, avatars, animations, and etc.

In this paper, we propose a 3D face modeling method based on 3D dense morphable shape model. The proposed 3D modeling method first constructs a 3D dense morphable shape model from 3D face scan data obtained using a 3D scanner. Next, the proposed method extracts and matches facial landmarks from 2D image sequence containing a face to be modeled, and then reconstructs 3D vertices coordinates of the landmarks using a factorization-based SMF technique. Then, the proposed method obtains a 3D dense shape model of the face to be modeled by fitting the constructed 3D dense morphable shape model into the reconstructed 3D vertices. Also, the proposed method makes a cylindrical texture map using 2D face image sequence. Finally, the proposed method generates a 3D face model by rendering the 3D dense face shape model using the cylindrical texture map. Through building processes of 3D face model by the proposed method, it is shown that the proposed method is relatively easy, fast and precise.

Keywords—3D Face Modeling, 3D Morphable Shape Model, 3D Reconstruction, 3D Correspondence.

I. INTRODUCTION

REALISTIC 3D face model is more precise than 2D face model in representing pose, illumination, and expression of face so that it can be utilized successfully in various applications such as face recognition, games, avatars, animations, and etc. During the last decades, tremendous research efforts have been poured to obtain more realistic and easier 3D face modeling [1-13].

3D face can be represented by volume or surface, but usually represented by polygonal surface, which consists of triangular meshes. Mesh is composed of vertices. Therefore, 3D vertices coordinates need to be precisely determined. After face meshes are built, one can obtain 3D face model by rendering the face meshes using vertex colors or a face model texture map.

Among 3D realistic face modeling works reported until now, 3D face modeling obtained by 3D scanner [14], 3D face modeling based on a generic face model [2,3,4,6,7,8,10], 3D face modeling based on 3D morphable model [5,11], and 3D face modeling based on 3D morphable shape model [9,12,13] are representative.

Yongsuk Jang, Sun-Tae Chung and BooGyun Kim are with School of Electronic Engineering, Soongsil University, Seoul, Korea (e-mail: cst@ssu.ac.kr).
Seongwon Cho is with A.I. Lab., Hongik University, Seoul, Korea.

3D scanner captures 3D vertices coordinates and colors for sampling points in 3D face surface and provides a 3D face model based on these scan data. However, 3D scanner is very expensive and 3D modeling using 3D scanner is not appropriate for real situations. 3D face modeling by deforming a 3D face generic model to fit the 3D vertex of some facial landmarks is relatively simple and fast in processing, but cannot obtain an accurate 3D face model since the adopted 3D generic model has a sparse number of vertices and triangles. 3D face modeling based on a 3D morphable model first constructs a 3D morphable model which consists of 3D shape PCA model and texture 3D PCA model. Next, for a new input 2D face image, the 3D face modeling method fits the 3D morphable model into the new input 2D face image. The 3D morphable model based 3D face modeling is very accurate and needs just single 2D face image to produce a 3D face model, but it costs a prohibitive computation time to generate a 3D face model for real-time processing. The 3D face modeling based on 3D morphable shape model first constructs a 3D morphable face shape model. Given a series of 2D face images of a person to be 3D modeled, the corresponding facial landmarks are detected and 3D vertices coordinates of the facial landmarks are reconstructed using a technique of Structure from Motion (SfM), and build a 3D surface mesh consisting of the reconstructed 3D vertices. Next, this 3D face modeling method obtains a 3D face shape model of the person to be modeled by fitting 3D morphable shape model into the 3D surface mesh. Also, a cylindrical blended texture map is obtained using the textures of 2D face images. Finally, the modeling method accomplishes a 3D face modeling by rendering the 3D face shape model using the cylindrical blended texture map.

The 3D morphable shape model-based 3D face modeling takes much less time than 3D morphable model-based face modeling, and produces a fairly good 3D face model. [12] proposed a 3D morphable shape face model consisting of a neutral face model and a span of 65 metrics. The neutral face contains 194 vertices and 360 triangles. Metrics are introduced for modeling variations of face shapes and designed by an artist. But, since vertices and triangles are not dense enough, the resulting 3D morphable face shape model is not accurate enough for some applications. Moreover, [12] obtains the corresponding corner points by using Plessy corner detector, which is not robust under various illumination environments.

In this paper, we propose a 3D face modeling method based on a 3D dense morphable face shape model. In order to construct a 3D dense morphable face shape model, the 3D dense correspondence problem should be preceently resolved.

International Scholarly and Scientific Research & Innovation 2(1) 2008 44

ISNI:0000000091950263
In this paper, we develop a fast and reliable way to resolve 3D correspondence problem among 3D dense face scan data. As explained before, a reliable extraction of the corresponding facial landmarks and reconstruction of 3D vertices coordinates of the facial landmarks is an important step to generate a realistic 3D face shape model. As supposed to [12], we employ AAM to extract the corresponding facial landmarks, which is known to be relatively stable under various illumination conditions and poses, and a factorization-based SfM technique of orthographic case to reconstruct 3D vertices coordinates of facial landmarks, which is known to be robust to measurement noises. Also, the proposed modeling method constructs a cylindrical texture map using 2D face image sequence. Finally, the proposed modeling method generates a 3D face model by rendering the 3D dense face shape model using the cylindrical texture map.

The proposed 3D face modeling method produces a 3D face model relatively faster than the morphable model-based 3D face modeling, and more accurate than the generic model-based 3D face modeling. Compared to the previous morphable shape model-based 3D face modeling methods, the proposed one generates more accurate 3D face model.

The rest of the paper is organized as follows. Section 2 explains some technical background necessary for this paper, and Section 3 presents a 3D dense correspondence solution developed for building 3D dense morphable face shape model in this paper. Section 4 presents the proposed 3D face modeling method based on a 3D morphable shape model, and the conclusion is presented in Section 5.

II. BACKGROUND

A. Shape and Procrustes Analysis

According to D. G. Kendal [15], ‘shape is all the geometrical information that remains when location, scale and rotational effects are filtered out from an object’. Shape is usually represented as a set of landmarks on an object. A landmark is defined to be a point of correspondence on each object that matches between and within populations [15]. Procrustes analysis is a rigid shape analysis that uses isomorphic scaling, translation, and rotation to find the best fit between two or more landmarked shapes.

B. AAM (Active Appearance Model)

AAM is a computer vision algorithm for matching a statistical model of object shape and appearance to a new image [16]. The statistical model is constructed in training phase from model object images with positions of landmarks. AAM matches a statistical model by an optimization process minimizing the difference the current estimate of appearance and the new image. AAM is known to work relatively well under various illumination conditions and poses.

C. ICP (Iterative Closest Point)

ICP is an algorithm employed to match two clouds of points [17,18]. This matching is used to reconstruct 3D surfaces from different scans, to localize robots, to match bone models with measures in real-time, etc. The algorithm is very simple and is commonly used in real-time. It iteratively estimates the transformation (translation, rotation) between two raw scans. Since ICP algorithm is introduced by Besl and Mckay [17], many variants of ICP algorithm [18] and extensions [19] had been proposed.

D. TPS (Thin Plate Spline)

TPS is a class of non-rigid nonlinear spline mapping function and is an effective tool to connect sparse points in space smoothly [20,11,19]. In this paper, we use TPS to align a sample face into a reference face locally.

E. Structure from Motion

Structure from Motion(SfM) which refers to the process of obtaining scene structure, camera motion(pose) from the corresponding features (feature points, corners, edges, and etc.) in 2D image sequence has been one of the important fundamental problems in computer vision [21,22,23,24,25]. In SfM, locations of 2D features are dependent on 1) their coordinates in 3D space, 2) the relative 3D motion between the camera and the scene and 3) the camera’s internal geometry, and these three causes are usually assumed not to be known beforehand. Epipolar geometry based SfM[23], factorization-based SfM [22,23,24], and SfM using nonlinear kalman filter [25] are some of representative works in SfM researches.

F. 3D Dense Morphable Face Shape Model

The 3D dense morphable face shape model adopted in this paper is constructed from 3D scan data. The 3D face shape scan data consisting of n vertices can be represented as a shape vector

\[ S = (x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_n, y_n, z_n)^T \]

where \( x_i, y_i, z_i \) are \( x, y, z \) coordinates of vertex \( i \). Now, if one finds the mean 3D shape vector \( \bar{S} \) and shape principal modes \( S_i \) (i = 1, ..., m - 1) by applying PCA analysis for m 3D face shape scan data, a face shape vector \( S \) can be represented as

\[ S = \bar{S} + \sum_{i=1}^{m-1} \alpha_i S_i \quad (1) \]

III. 3D FACE DENSE CORRESPONDENCE

In order to construct a 3D dense morphable face shape model from 3D dense face scan data like (1), 3D dense correspondence among 3D face scan data has to be preceددently resolved. Many approaches for 3D dense correspondence have been suggested [5,11,13,26,27]. Optical flow-based dense correspondence [5] is affected by illumination conditions, and normal vector search [13] and TPS-based dense correspondence [11] had been suggested to alleviate defects of optical-flow based one, but are
not sufficiently satisfactory. Optimization-based dense correspondence such as SPHARM [26] and MDL[27] are known to show the excellent performance, but SPHARM and MDL can be applied to genus 0 closed surface which 3D scan face scan data do not belong to.

In this paper, we develop a simple and computationally fast 3D dense correspondence solution. The procedures for 3D dense correspondence developed in this paper can be summarized as follows.

1) Select control points and a reference 3D face shape scan data
2) Align globally a sample 3D face shape scan data into a reference face shape scan data by using procrustes analysis
3) Align locally a sample 3D face shape into a reference face shape scan data by using TPS(Thin Plate Spline)
4) Search the corresponding nearest point on a sample face scan data to each vertex on the reference face scan data. If the distance is within the threshold, then determine the closest point as the corresponding point, and if not, discard the closest point
5) Find the corresponding point on the original sample face scan data by taking inverse TPS transformation and inverse Procrustes transformation

Fig. 1 shows the result of global alignment of a sample face scan data into a reference face scan data using procrustes analysis and local alignment using TPS.

![Fig. 1](image)

Fig. 1 (a) two original 3D face scan data with one overlaid on the other, (b) two 3D face scan data after Procrustes, (c) two 3D face scan data after TPS

From Fig. 1 (c), a sample face aligns well into a reference face after global alignment using Procrustes analysis and local alignment using TPS. Through the developed 3D correspondence procedures, we could construct 40 3D scan data consisting of the corresponding approximate 40,000 points from 40 face scan data with 81,134 points through 598,121 points. Fig. 2 compares the result of the 3D dense correspondence procedures developed in this paper with original face scan data. One can see very close similarity between the original face scan data and face scan data consisting of the corresponding points only.

![Fig. 2](image)

Fig. 2 Comparison between original face scan data (left) and face scan data consisting of the corresponding points only (right)

IV. 3D FACE MODELING BASED ON 3D DENSE MORPHABLE FACE SHAPE MODEL

A. Outline

The proposed 3D face modeling method based on 3D dense morphable face shape model consists of two stages: 1) construction of 3D morphable face shape model, 2) 3D face modeling.

B. Construction of 3D Dense Morphable Face Shape Model

In this stage, we construct a 3D morphable face shape model like (1) by applying PCA to several 3D scan face data. The most important process in this stage is to arrange all 3D face scan to consist of corresponding points by using 3D dense correspondence, which we deal with by utilizing the 3D dense correspondence procedures in Section III.

In this paper, for the construction of 3D dense morphable face shape model, we use 40 persons’ 3D face scan data captured by Cyberware Model 3030 Color Scanhead [14]. Fig. 3 shows some examples of 3D scan face data.

![Fig. 3](image)

Fig. 3 Some face scan data

Hair and ears are cut from the face scan data, and each resulting face scan data contains more than 70,000 points points (81,134 points through 598,121 points) sampled from each face.

After we apply 3D dense correspondence procedures in Section III, we obtain aligned 40 face shape vector, each of
which has approximately 40,000 points. For these 40 aligned shape vectors $F_1, F_2, \ldots, F_m$ ($m = 40$), we first seek a mean face shape, $\bar{S} = (F_1 + F_2 + \cdots + F_m) / m$. Now, for scan data matrix $A = (F_1, F_2, \ldots, F_m)$ where $A_i = F_i - \bar{S}$ ($i = 1, \ldots, m$), let’s seek principal modes $S_1, S_2, \ldots, S_{m-1}$ ($m = 40$) from the covariance matrix $C$ by applying PCA analysis to covariance matrix $C = \frac{1}{m} AA^T$. Then, we can represent a 3D sample face scan vector $S$ as follows.

$$S = \bar{S} + \sum_{i=1}^{m-1} \alpha_i S_i$$

(2)

C. 3D Face Modeling

3D face modeling stage is to generate 3D face model of the corresponding person from 2D face image sequence of a person to be modeled, and proceeds in the following steps:

1. Extract and match the corresponding face landmarks and their 2D coordinates from 2D face image sequence
2. Reconstruct 3D vertices coordinates of the face landmarks from the 2D coordinates of the corresponding face landmarks using a factorization-based SfM technique
3. Obtain a 3D dense shape model of the face to be modeled by
4. Fitting the constructed 3D dense morphable face shape model into the 3D vertices reconstructed from 2D facial landmarks
5. Make the facial texture map from 2D face image sequence
6. Generate 3D face model by rendering the obtained 3D dense face shape model using the texture map.

1) Extraction and Matching of the Corresponding Facial Landmarks

The previous approaches to 2D correspondence includes optical-flow based method [5], and corner matching based method [12], but these methods do not work properly if each 2D image is under different illumination conditions. AAM is an algorithm which can be used for extracting facial landmarks and is known to work relatively well under various illumination variations and various poses. Thus, in this paper, we extract facial landmarks using AAM. Fig. 4 shows positions of 63 facial landmarks for AAM in this paper.

2) Reconstruction of 3D Vertices Coordinates of Facial Landmarks

Since the face is located sufficiently distant from camera, one can assume orthographic projection is a fairly good approximation to perspective projection in this case. In this paper, we apply factorization-based SfM method of orthographic projection case [22] to obtain 3D vertices coordinates of facial landmarks.

Fig. 5 shows 3D facial images obtained by rendering the 3D face surface mesh consisting of the reconstructed 3D vertices of facial landmarks with 2D face image texture, and it implies that the reconstructed 3D vertices and meshes are fairly accurate.

![Fig. 5 Reconstructed 3D vertices and meshes](image)

3) Obtainment of 3D Dense Face Shape Model of the Face

Now, we need to obtain a 3D dense shape model of the face to be modeled by fitting the 3D dense morphable shape face model into the reconstructed 3D vertices. Among the reconstructed vertices of the facial landmarks, vertices on the chin are not reliable since the landmarks on the chin of a 2D image may be different from the landmarks on the chin of the other 2D images. Thus, the landmarks on the chin are excluded from fitting process. We take the same approach to [12] in fitting 3D dense morphable face shape model into the reconstructed 3D vertices of facial landmarks.

Given a set of reconstructed 3D points from matched facial landmarks, the fitting process searches for both the pose of the face and the PCA coefficients $\alpha_i$ of 3D dense morphable face shape model. to minimize the distance from the reconstructed 3D points to the face mesh. The pose of the face is the transformation $T = \begin{pmatrix} sR & t \\ 0^T & 1 \end{pmatrix}$ from the coordinate frame of the mean face mesh to the camera frame, where $R$ is a 3x3 rotation matrix, $t$ is a translation, and $s$ is a global scale. For any 3D vector $P$, we use notation $T(P) = sRP + t$. The vertex coordinates of the face mesh in the camera frame is a function of both the PCA coefficients $\alpha_i$ of 3D dense morphable shape model and the pose of the face. Given PCA coefficients $(\alpha_1, \alpha_2, \ldots, \alpha_{m-1})$ and pose $T$, the face geometry in the camera frame is given by

$$S = T(\bar{S} + \sum_{i=1}^{m-1} \alpha_i S_i) .$$

Let’s denote the reconstructed vertices representing eyes, node, and mouth as $Q_1, Q_2, \ldots, Q_k$ and the rest of the reconstructed vertices as $P_1, P_2, \ldots, P_n$ where vertices on the chin are excluded. Also, we denote the vertices on the face mesh corresponding to reconstructed vertices $Q_1, Q_2, \ldots, Q_k$ as $V_j$ ($j = 1, 2, \ldots, k$). Now, the fitting process is processed as an
optimization problem by finding pose $T$ and the PCA coefficients $(\alpha_1, \alpha_2, \ldots, \alpha_{m-1})$ to minimize

$$
\sum_{i=1}^{n} \omega_i d^2(P_i, S) + \sum_{j=1}^{k} d^2(P_i, V_j)
$$

(3)

where $\omega_i$ is a weighting factor. To solve this problem, we use an ICP approach. At each iteration, we first fix pose $T$. For each $P_i$, we find the closest point $G_i$ on the current face mesh, and then minimize (3). We set $\omega_i$ to be 1 at the first iteration and $\omega_i = \frac{1}{1 + d^2(P_i, G_i)}$ in the subsequent. Such a weighting scheme is an effective way to avoid overfitting to the noisy data [12]. Since both $G_i$ and $V_j$ are linear functions of PCA coefficients, (3) becomes a linear square problem so that one can determine PCA coefficients $\alpha_j$. After PCA coefficients are determined at each iteration, we recalculate $G_i$ and $V_j$, and calculate pose $T$ again for the recalculated $G_i$ and $V_j$ by taking the PCA coefficients as fixed. We do the iteration process until the cost of (3) drops until the threshold.

Fig. 6 shows a 3D face shape model of the face of Fig. 2 obtained after fitting.

![Fig. 6 Reconstructed 3D face shape](image)

4) Making the Face Texture Map

We make the view independent cylindrical texture map in the following way. We first compute the blending weight of each triangle on the face mesh for each image based on the angle between surface normal and the camera direction. Fig. 7 shows the 2D face image where the reconstructed 3D face shape meshes are projected back to the 2D face images. The projected face shape mesh is the simplified mesh with 800 vertices while the reconstructed face shape mesh has approximately 40,000 vertices.

![Fig. 7 Reconstructed 3D face meshes projected back to a 2D face images](image)

If the triangle is invisible, its weight is set to 0.0 and weights are then normalized so that the sum of the weights over all the images is equal to 1.0. For each image, we generate a cylindrical texture map by rendering the cylindrical mapped mesh with the current image as texture map. Let $C_i$ and $W_i (i = 1, 2, \ldots, k)$ be the cylindrical texture maps and the weight maps. The final blended texture map, $C$ is

$$
C = \sum_{i=1}^{k} W_i(u, v)C_i(u, v)
$$

(4)

Fig. 8 shows final cylindrical blended texture for 2D image sequence of the person of Fig. 7.

![Fig. 8 The blended texture image](image)

5) Generation of 3D Face Model

We finally generate a 3D face model by rendering the obtained 3D face shape model with the blended texture map.

Fig. 9 compares the 2D face images with the reconstructed 3D face model images with poses corresponding to the poses of 2D face images. One can see in Fig. 9 that the reconstructed 3D face model from the proposed 3D face modeling method looks similar to 2D images. Due to the view independent texture map of 3D face model, the noses in the reconstructed 3D face model image looks blurred and the skin colors of the reconstructed 3D face model images look different from those of the corresponding 2D face images.

![Fig. 9 Comparison between reconstructed 3D face model images and original 2D face images](image)
V. CONCLUSION

In this paper, we proposed a 3D face modeling method based on 3D dense morphable face shape model. The proposed 3D face modeling method first constructs a 3D dense morphable face shape model from 3D face scan data. The most important issue in constructing a 3D dense morphable face shape model is to resolve the 3D dense correspondence problem. In this paper, we developed an easy and fast way to solve 3D face dense correspondence. Next, given a 2D face image sequences of the person to be modeled, the proposed 3D face modeling method extracts the corresponding facial landmarks by applying AAM into 2D face images, and reconstructs 3D vertices coordinates of the corresponding facial landmarks. Then, the proposed 3D face modeling method fits the constructed 3D dense morphable face shape model into the reconstructed 3D vertices and obtains 3D face shape model of the person to be modeled. Also, the proposed modeling method provides a way to construct a view independent blended texture map. Finally, the proposed 3D face modeling method reconstructs 3D face model of the person by rendering the reconstructed 3D face shape with the blended texture map. The 3D face model produced by the proposed modeling method looks fairly good, and the proposed 3D modeling method turns out to be easy and fast to generate.

ACKNOWLEDGMENT

This work was supported by the Soongsil University Research Fund and BK21.

REFERENCES