



3D face dense reconstruction based on sparse points using probabilistic principal component analysis

Xiaoxiao Xie^{1,2} · Xingce Wang^{1,2} · Zhongke Wu^{1,2}

Received: 31 May 2021 / Revised: 18 September 2021 / Accepted: 22 October 2021 /
Published online: 9 November 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Reconstructing 3D face from sparse points is an ill-posed problem. While there already exists available solutions addressing this problem, to our knowledge, we propose a better-performed approach which can robustly reconstruct fine 3D face shape. Our method includes two modules: face model establishment based on probabilistic principal component analysis (PPCA) trained in an unsupervised manner to learn transformation between landmarks and point cloud in their low-dimensional representation, and 3D face reconstruction based on learned relation between them to reconstruct fine face shape. Overall, our method considers the probability of face shape and learns more useful information of 3D face shape. We compare our method with 3 typical and state-of-the-art methods on 2 datasets and the effectiveness of our method is demonstrated generally. Further experiments on datasets with noise of different intensities show the stability of our method.

Keywords 3D face reconstruction · Landmarks · Probabilistic principal component analysis · Learning-based method

1 Introduction

3D face reconstruction has become an important research topic in the fields of computer graphics, computer vision and pattern recognition with its wide using in national security, password unlocking, financial payment and cosmetic medicine. 3D face shape contains abundant information about the identity of individuals and face reconstruction methods have been applied in identification [15], face animation [26] and computer games. Subsequently, many methods have been proposed to reconstruct facial expressions, poses, and hair [8, 18]. Though 2D images and videos data are convenient to access, the reconstruction of 3D face shape from them are susceptible to changes of

✉ Zhongke Wu
zwu@bnu.edu.cn

¹ School of Artificial Intelligence, Beijing Normal University, Beijing, China

² Engineering Research Center of Virtual Reality and Applications, Ministry of Education, Beijing Key Laboratory of Digital Preservation and Virtual Reality for Cultural Heritage, Beijing Normal University, 100875 Beijing, China

light and angles. With the expansions of multi-eye vision and 3D model construction methods, the acquisition of 3D data becomes easier. Combining the advantage of containing more depth information, research on 3D face reconstruction from 3D feature points becomes increasingly important.

Since sparse points contain less information regarding the 3D face shape, face reconstruction from sparse points is a highly challenging, ill-posed problem. 3D morphable model (3DMM) [2] is a widely used geometry-based method to solve this problem, which obtains the correspondence between landmarks in 2D images and 3D face shape and minimizes optimized function to estimate 3DMM parameters. Nonlinear 3D morphable model [28, 29] expresses 3D face shape in a nonlinear manner, which has greater representation power than traditional linear 3DMM. These methods cause the disappearance of some facial features and high similarity between reconstructed faces and neutral face model.

Learning-based method is more suitable for 3D face reconstruction from sparse points and it has two major advantages than geometry-based methods. First, the learned prior knowledge can assist in completing missing information of 3D face shape. Second, the reconstructed faces are more detailed as a result of learning face features. Recently, many works applied deep learning networks [4, 19, 35] to facilitate accurate reconstruction. However, these methods require large 3D datasets so as to learn abundant prior knowledge about 3D face shape, and utilize the loss of landmarks to optimize reconstructed face which result in the insufficient use of landmarks information.

To address these issues, we combine two feature-extraction statistical methods in model establishment: Principal component analysis (PCA) and factor analysis (FA). As a widely used statistical method, PCA finds correlations between variables and represents original space using principal axes. During dimension reduction procedure, PCA method eliminates the noise directly and leads to the loss of some detailed information, while FA method considers small deviation of data. Through this combination, the relationship between sparse points and 3D face shape can be learned in greater details using fewer sample faces.

- (1) We propose a statistics-based method for face reconstruction by combining PCA and FA methods, where the principal axes intuitively correspond to the subspace of face shapes. In face shape analysis, we express the 3D face shape in a probabilistic manner. The probabilistic presentation preserves more detailed information of a face when establishing our face model. Using few sample faces, our method can still reconstruct a 3D face with fine details.
- (2) We apply probabilistic principal component analysis theory and obtain the distribution of latent variables by the maximum likelihood function and expectation maximization (EM) method. The PPCA method is efficient for dimension reduction based on the faces with large dimensions and the EM algorithm generates a subspace full of stability. As a probabilistic formulation of PCA applied for face features extraction, the reduced representation of a face contains more local details.
- (3) Our method is effective and adaptable on different 3D face databases based on our experiments on the Chinese craniofacial database and 3D FaceWarehouse. Compared to the traditional reconstruction methods, our method reduces the average reconstruction error by 70.35%, 19.23% and 48.78%. Further experiments on datasets with noise demonstrate the stability of 3D face reconstruction; in particular, for faces from FaceWarehouse with added 10% Gaussian noise, our method reduces the face reconstruction loss by 18.88% compared with the traditional non-probabilistic method.

2 Related work

In this section, related works on the reconstruction of 3D face from sparse points and the applications of PPCA method are reviewed.

2.1 Face reconstruction based on landmarks

Previous works chose landmarks as their input data and proposed several methods for the reconstruction of the 3D face shape. Among these methods, the relationship between landmarks and 3D face shape can be derived from the geometric deforming, supervised learning and unsupervised learning methods.

Geometric Deforming method Geometric deforming methods reconstruct 3D face shape through the deformation of existing 3D face models, Basel Face Model [21] for instance. Recently, Xiao [34] reconstructed 3D faces from estimated 3D landmarks by shape deformation. They deformed the 3D faces through a deformation function based on the radial basis function (RBF). Since 3D landmarks were estimated from 2D landmarks, the inaccuracy may cause unrealistic details in the mouth and nose area. Aji [1] added facial soft tissue thickness (FSTT) on Craniofacial data to obtain the corresponding 3D landmarks and generated 3D face shape using the Laplacian deformation. This method estimated 3D landmarks from skull landmarks, therefore, the changes of age and gender influenced the structure of the face and gave rise to inaccuracies for the reconstructed faces. Hu [13] reconstructed the 3D face shape from a set of feature points in an analysis-by-synthesis loop. Reconstructed faces using the geometry-based method are usually similar to the neutral face model. Geometry-based methods reconstruct 3D face shape through the deformation of the established face model, and some details regarding the faces cannot be recovered.

Supervised Learning method Supervised learning methods obtain the relationship between input data and 3D faces based on training feedback. Liu [18] learned 3D shape regressors in an interactive process. Fitting such a regressor is an optimization process that does not require additional face alignment methods. To enhance the performance of face reconstruction, many state-of-the-art methods utilized neural networks for face reconstruction. Tran [30] applied a state-of-the-art CNN to estimate 3DMM parameters and demonstrated its excellent performance on face recognition from in-the-wild images compared to the traditional 3DMM and 3DDFA methods [2, 38]. Liu [19] reconstructed 3D face shapes from in-the-wild unlabeled images with a pose guidance network(PGN) and used 3D landmarks to obtain parameters to adjust face poses. These works can adjust the reconstructed face in real time according to the feedback of the reconstruction error. However, online face reconstruction is not suitable for all environments and these methods are susceptible to the quality of the training datasets.

Unsupervised Learning method Unsupervised learning obtains a relationship between the input and 3D face shape by extracting their features. For an input single image with noise, Zhong [37] proposed a coupled facial denoising and reconstruction network (FDR) method that utilized a three-tier shape consistency including the feature, depth image and surface levels. Compared to the state-of-the-art reconstruction methods, the method of [37] was effective in denoising and face reconstruction. The reconstruction from 2D images tended to obtain a large, estimated error in depth estimation, and thus affected the reconstruction performance. Different from 2D images, the 3D representation of a face not only contains accurate depth information but also performs more stereoscopic. Therefore, we

learn the mapping between the sparse points and the dense point cloud and reconstruct 3D face shape from 3D sparse points.

Many researchers prefer to choose 2D images as the input data. Compared to 3D information, 2D images have a drawback of lacking depth information, therefore, 3D landmarks are selected as the input data in our method. 3D information about face shape can reflect depth information and details even they contain only few 3D vertices.

2.2 PPCA method

PCA [32] is an unsupervised dimension reduction method used in statistics and its observed values are described by several related variables. Several works further developed PCA methods such as sparse principal component analysis (SPCA) [39], robust principal component analysis (ROBPCA) [14], kernel principal component analysis (KPCA) [23], probabilistic principal component analysis [27] and generalized probabilistic principal component analysis (GPPCA) [11]. Tipping and Bishop [27] obtained the probabilistic formulation of PCA and established the probabilistic principal component model that regarded latent variables as missing data with latent variables corresponding to principal axes. Then, PPCA method was applied to many fields, for example, image processing [20, 31], disease diagnosis [24] and behaviour analysis [9]. Mredhula and Dorairangaswamy [20] employed the pixel surge model (PSM) and PPCA method to eliminate the salt-and-pepper noise in image processing and utilized filters to improve the quality of image denoising. Vaddi and Manoharan [31] applied the PPCA method to reduce the dimensions of the hyperspectral image (HSI) into a latent subspace prior to classification, obtaining an excellent result compared to the traditional PCA method. The PPCA method has been applied in disease diagnosis, and Shah [24] proposed an automatic heart disease detection method by extracting principal features from datasets using parallel PPCA. Geraci and Farcomeni [9] analyzed the physical activity levels of the children in the UK using this method. PPCA method was used to analyze the relations among multi-variables and obtained the understanding of physical activity and inactivity through non-ignorable missing data. These studies proved the effectiveness of the principal features extraction and reduced the influence of noise through the use of PPCA. The applications of the PPCA method in various fields verified its effectiveness on dimension reduction; therefore, our proposed method applies PPCA to reconstruct 3D face shape and establishes our 3D face model.

3 Formulation and our pipeline

In this section, we briefly describe the problem of reconstructing 3D face shape based on sparse points, and introduce the pipeline of our reconstruction method.

3.1 Problem formulation

Mathematically, the problem of the 3D face reconstruction [13] can be described as:

$$S = \phi(I) \quad (1)$$

where the mapping $\phi : \mathbb{L} \rightarrow \mathbb{S}$ maps sparse representation $l \in \mathbb{L}$ to the corresponding 3D face shape \mathcal{S} . Given a face sparse representation l , our objective is obtaining its corresponding 3D face shape.

Generally, ϕ is a linear mapping, [34] proposed a method to obtain the mapping ϕ by using radial basis function (RBF). The representation of the 3D face shape is described by $\mathcal{S} = \bar{\mathcal{S}} + W_{\mathbb{S}}Z_{\mathbb{S}}$, where $\bar{\mathcal{S}}$ is the average 3D face shape, $Z_{\mathbb{S}}$ is the low-dimensional representation of the 3D face shape in subspace and $W_{\mathbb{S}}$ is the coefficient matrix for subspace construction.

\mathcal{S}_g denotes the groundtruth of a 3D face, and we define the objective function of reconstructing 3D face as (2).

$$\arg \min_{\phi} \{dist(\mathcal{S}_r, \mathcal{S}_g)\} \tag{2}$$

where $dist(\mathcal{S}_r, \mathcal{S}_g)$ denotes the distance between the groundtruth \mathcal{S}_g and the reconstructed face \mathcal{S}_r . The objective function must be optimized to obtain the mapping ϕ and make the reconstructed faces as close to their corresponding groundtruths as possible.

Sparse points of a 3D face contain partial information about distinctive face features. The process of generating detailed information regarding 3D face face from sparse points is of great challenge. Meanwhile, face dense point space generating from few sample faces leads to the incompleteness of the reconstructed face dense point space; therefore, we consider the probability of face shape and construct a probabilistic space, with the face model established in this work described below Fig. 1.

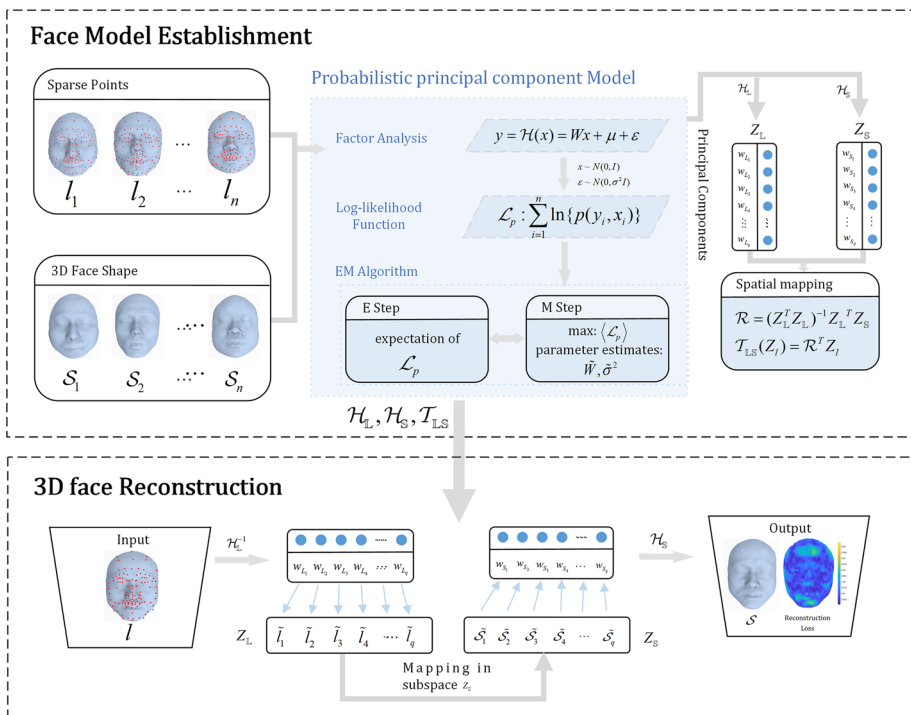


Fig. 1 The pipeline of our method which includes two parts: face model establishment and 3D face reconstruction

3.2 Pipeline

The pipeline of our method includes two modules: face model establishment and 3D face reconstruction, in which we establish our 3D face model through few sampled faces and reconstruct the corresponding 3D face shape using input sparse points.

In the face model establishment, we exploit PPCA method to extract the principal axes from the landmarks dataset \mathbb{L} and the corresponding 3D face shape dataset \mathbb{S} respectively. Subsequently, we express the landmarks and the 3D face shape in a low-dimensional manner and construct corresponding subspaces. We further obtain the transformation mapping between two subspaces by least square method. With the mappings of dimension reduction and spatial transformation, our 3D face model can be established.

In the 3D face reconstruction, we accept landmarks as an input. With the established face model, we obtain the low-dimensional representation of the input landmarks. Then, using transformation mapping, we can obtain its corresponding representation in the face point cloud subspace. Finally, the dense point representation of the face can be generated by PPCA reconstruction method.

4 Methodology

Based on sparse points, our learning-based method can generate the corresponding 3D face shape approximately. To simulate 3D face shape in practical, we consider the noise of the sampled faces, construct the probabilistic form of the landmarks and 3D face shape, and apply PPCA to generate their low-dimensional representations.

4.1 Construction of probabilistic space

Generally, 3D face shapes are various in the area of eyes, nose and mouth. To express 3D face shapes in the real world as abundant as possible, we take the probability of 3D face shapes into consideration and provide reconstructed faces a more detailed description; therefore, the 3D face shape can be described as:

$$\mathcal{S} = W_{\mathbb{S}}Z_{\mathbb{S}} + \bar{\mathcal{S}} + \epsilon_{\mathbb{S}} \quad (3)$$

where $\epsilon_{\mathbb{S}}$ denotes the residual of the scanned 3D face, $Z_{\mathbb{S}}$ denotes latent variables in 3D face shape subspace, and the matrix $W_{\mathbb{S}}$ relates 3D face shape and its low-dimensional representation. In fact, the majority of noises are Gaussian processes, and we regard latent variables $Z_{\mathbb{S}}$ as the simplest Gaussian distribution $Z_{\mathbb{S}} \sim N(0, I)$ and the noise term $\epsilon_{\mathbb{S}} \sim N(0, \sigma_{\mathbb{S}}^2 I)$. Thus the conditional probability distribution of \mathcal{S} is $\mathcal{S} | Z_{\mathbb{S}} \sim N(W_{\mathbb{S}}Z_{\mathbb{S}} + \bar{\mathcal{S}}, \sigma_{\mathbb{S}}^2 I)$. The probability density function of 3D face shape \mathcal{S} is given as follows.

$$p(\mathcal{S}, W_{\mathbb{S}}, \sigma_{\mathbb{S}}^2) = \frac{1}{(2\pi)^{||\mathcal{S}||/2} |C_{\mathbb{S}}|^{1/2}} * \exp\left\{-\frac{1}{2}(\mathcal{S} - \bar{\mathcal{S}})^T C_{\mathbb{S}}^{-1} (\mathcal{S} - \bar{\mathcal{S}})\right\} \quad (4)$$

where $||\mathcal{S}||$ represents the number of vertices of 3D face point cloud, $C_{\mathbb{S}} = W_{\mathbb{S}}W_{\mathbb{S}}^T + \sigma_{\mathbb{S}}^2 I$ denotes the 3D face shape covariance, and therefore, the probability distribution of \mathcal{S} is $N(\bar{\mathcal{S}}, W_{\mathbb{S}}W_{\mathbb{S}}^T + \sigma_{\mathbb{S}}^2 I)$.

4.2 Dimensional reduction of PPCA method

With the known 3D face shape probability distribution, the estimation of the coefficient matrix W_S and variance σ_S^2 can be generated through its corresponding log-likelihood estimation and (5) provides its logarithm formulation.

$$\begin{aligned} & \ln\{p(S|W_S, \sigma_S^2)\} \\ &= -\frac{\|S\|}{2} \ln(2\pi) - \frac{1}{2} \ln |C_S| - \frac{1}{2} (S - \bar{S})^T C_S^{-1} (S - \bar{S}) \end{aligned} \tag{5}$$

$$\begin{aligned} \mathcal{L}_p &= \ln\{p(S|W_S, \sigma_S^2)\} \\ &= \sum_{n=1}^N \ln\{p(S_n|W_S, \sigma_S^2)\} \\ &= -\frac{N\|S\|}{2} \ln 2\pi - \frac{N}{2} \ln |C_S| - \frac{1}{2} \sum_{n=1}^N (S_n - \bar{S})^T C_S^{-1} (S_n - \bar{S}) \\ &= -\frac{N}{2} \{ \|S\| \ln(2\pi) + \ln |C_S| + tr(C_S^{-1} Cov_S) \} \end{aligned} \tag{6}$$

where N denotes the number of sample faces, $Cov_S = \frac{1}{N} \sum_{n=1}^N (S_n - \bar{S})(S_n - \bar{S})^T$ denotes the sample covariance matrix of the 3D face shape. The derivation of the log-likelihood formulation with respect to W_S is as follows.

$$\frac{\partial \mathcal{L}_p}{\partial W_S} = N(C_S^{-1} Cov_S C_S^{-1} W_S - C_S^{-1} W_S) \tag{7}$$

From (7), the stationary points satisfy (8).

$$Cov_S C_S^{-1} W_S = W_S \tag{8}$$

The estimated formulation of the coefficient matrix is given as equation (9) and the maximum likelihood estimator of variance σ_S^2 is given as (10).

$$W_{S,ML} = U_q (\Lambda_q - \sigma_S^2 I)^{\frac{1}{2}} R \tag{9}$$

$$\sigma_{S,ML}^2 = \frac{1}{\|S\| - q} \sum_{j=q+1}^{\|S\|} \lambda_j \tag{10}$$

where q represents the retained principal components and U_q consists of q eigenvectors of covariance Cov_S . Λ_q represents a diagonal matrix generated by $\lambda_1, \dots, \lambda_q$ that are the eigenvalues corresponding to q eigenvectors, and R denotes an arbitrary $q \times q$ orthogonal rotation matrix. The maximum likelihood estimated formulation of the coefficient matrix and variance are given by (9), (10). Specifically, we can utilize the EM algorithm to estimate them, then the corresponding latent variables can be obtained.

3D landmarks are selected from the point cloud of a face, therefore the probability of landmarks should be considered.

$$l = W_{\mathbb{L}}Z_l + \bar{l} + \epsilon_{\mathbb{L}} \quad (11)$$

In (11), \bar{l} and Z_l represent the average landmarks of a face and the latent variables in landmarks subspace, $\epsilon_{\mathbb{L}}$ denotes the residual of landmarks l , while the coefficient matrix $W_{\mathbb{L}}$ also represents the relation of the landmarks and its low-dimensional representation. By utilizing the above-described PPCA method, we can also generate the corresponding low-dimensional representation of the landmarks.

4.3 EM based parameter estimation

Generally, EM algorithm is used to estimate the parameters in the maximum likelihood function, which can be implemented by dividing into two steps. In the E step, based on the observed data, we determine the expectation of likelihood function of the latent variable distribution. For example, taking the observed data and latent data in landmarks space, we obtain the estimation of maximum likelihood $\mathcal{L}_p = \sum_{n=1}^N \ln\{p(S_n, Z_{\mathbb{S}})\}$ based on the current estimation. In the M step, we maximize the expectation of likelihood function and update the estimated parameters for achieving the maximum of the likelihood function. Related estimation of the parameters can be updated as (12),(13).

$$\tilde{W}_{\mathbb{S}} = Cov_{\mathbb{S}}W_{\mathbb{S}}(\sigma_{\mathbb{S}}^2I + M_{\mathbb{S}}^{-1}W_{\mathbb{S}}^TCov_{\mathbb{S}}W_{\mathbb{S}})^{-1} \quad (12)$$

$$\tilde{\sigma}_{\mathbb{S}}^2 = \frac{1}{||S||}tr(Cov_{\mathbb{S}} - Cov_{\mathbb{S}}W_{\mathbb{S}}M_{\mathbb{S}}^{-1}\tilde{W}_{\mathbb{S}}^T) \quad (13)$$

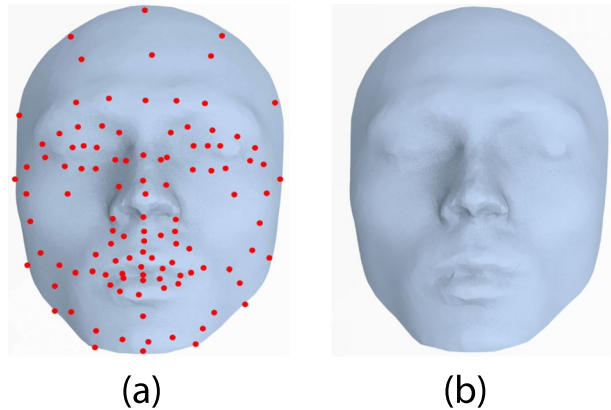
The E and M steps iterate continuously to achieve convergence [33], until the estimated parameters converge to a fixed value. Then, the mappings $\mathcal{H}_{\mathbb{L}} : Z_{\mathbb{L}} \rightarrow \mathbb{L}$ and $\mathcal{H}_{\mathbb{S}} : Z_{\mathbb{S}} \rightarrow \mathbb{S}$ can be determined. $\mathcal{H}_{\mathbb{L}}$ and $\mathcal{H}_{\mathbb{S}}$ map the landmarks in subspace $Z_{\mathbb{L}}$ and the 3D face shape in subspace $Z_{\mathbb{S}}$ into their original space, respectively. The principal components are extracted to obtain a low-dimensional representation and the subspaces of landmarks $Z_{\mathbb{L}}$ and face point cloud $Z_{\mathbb{S}}$ are obtained. Then, we should evaluate the correlation between the subspaces $Z_{\mathbb{L}}$ and $Z_{\mathbb{S}}$. We utilize linear spatial mapping to transfer the shape vectors in the landmarks subspace into the face point cloud subspace, and the 3D face model is established. When the variance $\sigma_{\mathbb{S}}^2$ is close to 0, the probability model converges to the traditional PCA model.

4.4 Unsupervised training on face datasets

Existing unsupervised methods usually exploit the correlation between the landmarks and the 3D face shape directly to facilitate 3D face reconstruction, which is full of challenge. Therefore, we find an indirect way to learning the correlation between these two spaces so as to decrease the complexity of face reconstruction problem.

To reconstruct the 3D face as accurately as possible, the sparse representation of the face that we choose must be able to represent the features of a 3D face. Landmarks can provide rich semantic information and descriptions of faces such as the corners of the eyes and mouth, the nose and the contour of the face. Based on the DLIB SDK [16] and MPEG-4 [25] standards, we select 113 landmarks from the face point cloud to make our method more efficient. A sample face and its sparse representation are shown in Fig. 2. In our method, we first train the landmarks dataset and face point cloud dataset under the

Fig. 2 The representation of a face from FaceWarehouse. (a) the sparse representation of a face. (b) the groundtruth of a face



condition of no feedback and aim to map the 3D landmarks to the face point cloud space. Therefore, we extract the principal features from the datasets and obtain the optimized mapping between the sparse points space and the dense point cloud space.

Feature-preserved learning Our framework takes the sparse points of a face as input and can be easily trained to maximize the retained details from the landmarks space and the face point cloud space. From Fig. 2 we can observe that the selected feature points contain information about the forehead, eyes, nose, mouth and contour of 3D face shape. Since our selected landmarks can contain more face features, the relationship between the spaces of the landmarks and the face point cloud can become more accurate. More details are illustrated in Section 5.

Prior to the reconstruction of 3D face shape based on input sparse points, a 3D face model should be established by the learning-based method. For the face model establishment module, we first establish the correlation between the landmarks and the face point cloud. Therefore, we preserve the features from the original space and analyze their relationship. Different from the traditional PCA method, the PPCA method considers the probability of spaces and preserves more details when extracting principal components. We utilize the PPCA method to preserve the features from the landmarks space and the face point cloud space. The EM algorithm introduced above can help estimate the coefficient matrices W_L , W_S and variances σ_L^2 , σ_S^2 . As convergence criteria set in [10], we use the criteria that the change in the transformation matrix elements must be less than 10^{-4} when we employ the EM algorithm. The process of the first module can be summarized as Algorithm 1, and $C = WW^T + \sigma^2I$, $M = W^TW + \sigma^2I$, d denotes the point number of sample face shape.

Algorithm 1 3D model establishment.

Input: Training datasets \mathbb{L} and \mathbb{S} ;

Output: Estimation of parameters $W_{\mathbb{L}}, W_{\mathbb{S}}, \sigma_{\mathbb{L}}^2, \sigma_{\mathbb{S}}^2$;

- 1: Extract principal components from spaces \mathbb{L} and \mathbb{S} ;
 - 2: Initialize parameters $W_{\mathbb{L}}, W_{\mathbb{S}}$ as random matrices, and $\sigma_{\mathbb{L}}^2, \sigma_{\mathbb{S}}^2$ are equal to 1;
 - 3: Update parameters as following:

$$\tilde{W} = SW(\sigma^2 I + M^{-1}W^T S W)^{-1}$$

$$\tilde{\sigma}^2 = \frac{1}{d} \text{tr}(S - S W M^{-1} \tilde{W}^T)$$
 - 4: **if** $|\tilde{W} - W| > 10^{-4}$ **then**
 $W = \tilde{W}$, iterate as step 3;
end if
 - 5: Determine the landmarks subspace $Z_{\mathbb{L}}$ and 3D face shape subspace $Z_{\mathbb{S}}$ by estimated parameters;
 - 6: **return** $W_{\mathbb{L}}, W_{\mathbb{S}}, \sigma_{\mathbb{L}}^2, \sigma_{\mathbb{S}}^2$;
-

3D face model establishment In Algorithm 1, we have learned information from the datasets and determined the estimations of the coefficient matrices and variances. The subspaces of the landmarks and 3D face shape can be generated, respectively. Since they are both q -dimensional, we utilize linear mapping to determine the correlation between the subspaces of the 3D face point cloud and the landmarks. The mapping between the low-dimensional representations of the landmarks and the 3D face shape can be described by (14).

$$\mathcal{T}_{\mathbb{L}\mathbb{S}}(Z_l) = (Z_{\mathbb{L}}^T Z_{\mathbb{L}})^{-1} Z_{\mathbb{L}}^T Z_{\mathbb{S}} Z_l \tag{14}$$

These estimated parameters $W_{\mathbb{L}}, W_{\mathbb{S}}, \sigma_{\mathbb{L}}^2, \sigma_{\mathbb{S}}^2$ and spatial transformation mapping $\mathcal{T}_{\mathbb{L}\mathbb{S}}$ can construct the mapping $\phi : \mathbb{L} \rightarrow \mathbb{S}$ and determine the 3D face model.

4.5 3D face reconstruction with landmarks

The process of establishing a 3D face model is an offline learning process, then, we can automatically determine the reconstructed face with the input landmarks. Existing supervised methods obtain the shape estimation with the feedback of the training loss, however, their methods require a large number of sample faces and datasets with high quality, which results in high-complexity of training step. The input sparse points l contain redundant information. Therefore, the reconstructed face obtained directly from the sparse points does not necessarily obtain the best reconstruction performance. To eliminate the influence of redundant information, we first extract its features based on the established 3D face model. We can map the input sparse landmark l into its low-dimensional representation Z_l in the subspace of landmarks as (15).

$$Z_l = (W_{\mathbb{L}}^T W_{\mathbb{L}} + \sigma_{\mathbb{L}}^2 I)^{-1} W_{\mathbb{L}} (l - \bar{l}) \tag{15}$$

However, before we obtain the reconstructed dense point representation of a face, the landmarks in its subspace must be transformed into the face point cloud subspace $Z_{\mathbb{S}}$. Based on the transformation mapping $\mathcal{T}_{\mathbb{L}\mathbb{S}}$, we obtain its representation in the 3D face shape subspace. Finally, we utilize the reconstruction method of PPCA to determine our

reconstructed face point cloud. Since our method is based on spatial mapping, more details can be preserved when transforming the landmarks into the face point cloud. The process of reconstructing a 3D face shape is summarized in Algorithm 2.

Algorithm 2 3D face reconstruction.

Input: The 3D landmarks l ;

Output: The reconstructed 3D face shape \mathcal{S} ;

- 1: Extract features from l using mapping \mathcal{H}_L ;
 - 2: Transform Z_l into the subspace of 3D face shape:

$$Z_s = \mathcal{T}_{LS}(Z_l);$$
 - 3: Reconstruct the 3D face shape \mathcal{S} with the PPCA method:

$$\mathcal{S} = W_S Z_s + \bar{S};$$
 - 4: **return** \mathcal{S} ;
-

5 Experiments and data analysis

In this section, we first determine our experimental setup for face model training and validation, including the evaluation metric and databases that we apply in our work. To combine PCA and FA, we have implemented PPCA and obtain our 3D face model based on it. We compare the effectiveness of our method with 3DMM method [2], PCA-based method and PRNet-based method [7]. The 3DMM method is a classic method for reconstruction by deforming the 3D face shape. We also implement our method over different datasets to prove its adaptability and several experiments over datasets with noise provide evidence for its stability in 3D face shape reconstruction.

Datasets We test our proposed method on Chinese craniofacial database [5, 6] and FaceWarehouse database [3]. Chinese craniofacial database contains the head slice image data of Shaanxi Xianyang hospital obtained by using a CT scanning system and consists of the craniofacial data of 208 volunteers from the Han nationality in northern China. There are 93 women and 115 men among these volunteers, with their ages ranging from 19 to 75. Each 3D face shape in its dense point representation contains 2709 vertices. These 208 3D face shapes in the dataset have been already registered to enable more effective evaluation. FaceWarehouse contains 150 3D faces from different ethnic backgrounds using Kinect's rgb-d camera. The ages of the individuals range from 7 to 80 and each individual has 47 different facial expressions. Specifically, the dense point representation of each 3D face shape in FaceWarehouse consists of 6508 vertices. We use 25 sample faces in Chinese craniofacial dataset and 150 sample faces in FaceWarehouse. For each sample face, we select 113 landmarks and construct the corresponding landmarks datasets.

Evaluation metric To evaluate the reconstruction performance, we choose the difference between the groundtruth \mathcal{S}_g and the output reconstructed face \mathcal{S}_r as reconstruction loss. Specifically, we employ the average Euclidean distance of the corresponding face points to measure the different reconstruction performance characteristics and take it as the value of the objective function. The reconstruction loss can be represented as follows.

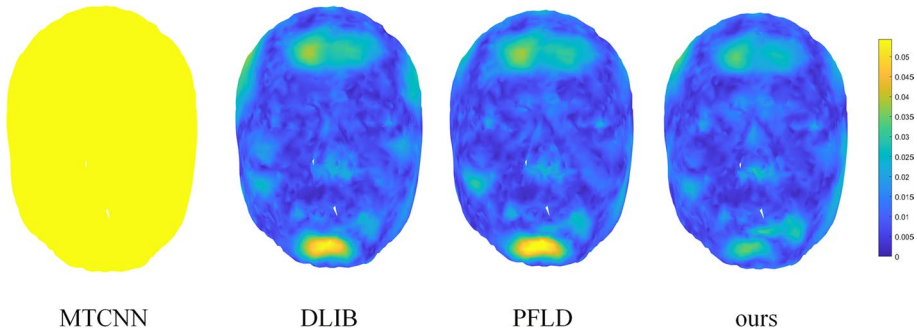


Fig. 3 The comparison of different selected landmarks

Table 1 The comparison of different methods on different datasets

Method	MTCNN	DLIB	PFLD	Ours
Mean Error	0.1225	0.0117	0.0108	0.0100

$$\mathcal{L}_e = \frac{1}{\|S\|} * \sum_{i=1}^{\|S\|} |S_{g,i} - S_{r,i}| \tag{16}$$

where $S_{g,i}$ denotes the i^{th} vertex in groundtruth face, and $S_{r,i}$ denotes the i^{th} vertex in reconstructed face. Since the reconstruction loss is defined as (16), a lower reconstruction loss indicates our method performs better and the obtained mapping ϕ is more appropriate.

The following experiments were carried out on a laptop with Intel® Core™ i5-10210U CPU @ 1.6-GHz processors, 16 GB RAM, 64-bit Windows 10 operating system, and MATLAB R2019b software.

5.1 Landmarks selection

Landmarks contain the shape features of a 3D face that may be closely related to the dense point representation of a face. Therefore, the selection of the landmarks will influence the performance of our method. To locate the feature points appropriately, we have referenced the DLIB SDK and MPEG-4 standard for landmarks detection. DLIB SDK for landmarks detection contains 68 feature points and the MPEG-4 standard contains 84 feature points. We select 113 landmarks to analyze the reconstruction loss with respect to the groundtruth. The 113 selected landmarks are arranged in groups such as eyes, cheeks, nose and mouth. For these groups, there are 15 feature points in the left eye area, 15 feature points in the right eye area, 19 feature points in the nose area, and 27 feature points in the mouth area. Moreover, 29 feature points are selected to describe the contour of face shape and 8 feature points contain information about the eyebrows and forehead.

To analyze the influence of the landmarks selection on the reconstructed face, we compare our selection with previous works on landmark selection, such as MTCNN(Multi-task convolutional neural network) [36], DLIB SDK and PFLD(Practical Facial Landmark Detector) [12]. Figure 3 illustrates the better performance of 113 selected landmarks and

Table 1 shows the reconstruction error of different landmark selection methods. Specifically, as we can observe, the reconstructed performance based on the MTCNN landmark selection shows a yellow face that demonstrates a greater reconstruction loss, and reconstruction based on DLIB SDK and PFLD shows greater reconstruction loss for the forehead and chin. By contrast, when we select 113 feature points as the sparse points representation of a face in the landmarks space, a better overall reconstruction performance is obtained.

5.2 Implementation details

In this section, we first consider the effectiveness of our method by comparing it with non-probabilistic methods. The proposed method takes 3D sparse landmarks as the input data, and 3D face point cloud as the output. Further experiments on the training datasets with noise prove the adaptability of the proposed method.

We use the Basel Face Model (BFM) [21] for comparison with our method and this model is the most commonly currently used 3DMM. The face model was trained over BFM and in this experiment, we test 3DMM by inputting 68 landmarks to reconstruct the 3D face shape. For the PRNet-based method, the model was already trained on 300W-LP [38], and therefore, we only test this method on FaceWarehouse. The PCA-based method and the proposed method take 30 samples of FaceWarehouse randomly and use the corresponding 68 landmarks as the training set, and the remaining samples comprise the test dataset. We set the number of latent variables $q = 28$ and initialize coefficient matrix W as a random $d \times q$ matrix, and variance $\sigma^2 = 1$.

We utilize leave-one-out cross-validation to verify the effectiveness and adaptability of our method compared with the traditional non-probabilistic method (the PCA method and PCA&LFA model[17] for example). To emphasize the effectiveness of the extraction of principal information, our experiments select 25 sample faces from the Chinese craniofacial database and divide them into 25 groups. Each sample face in the database belongs to one group. Every time we select 24 groups to train and the rest of the sample faces are the test samples which reconstruct 3D faces using our 113 landmarks. We also perform similar experiments on FaceWarehouse, where we select 30 sample faces from FaceWarehouse randomly as our experimental dataset. Prior to establishing our face model, we set the number of latent variable $q = 23$ for Chinese craniofacial database and $q = 28$ for FaceWarehouse, preserving as much information as possible in the low-dimensional space. In the step of the extraction of principal components, we initialize coefficient matrix W as a random $d \times q$ matrix, and variance $\sigma^2 = 1$.

To verify the stability of our method, we add Gaussian noise and Poisson noise on the Chinese craniofacial database and FaceWarehouse to simulate the datasets with noise. Since the original datasets are smooth without noise, we add noise of different intensities to the training set under the condition that do not add noise to landmarks. For Gaussian noise, we add 1,3 and 5% noise with the variance values of 0.01 and 0.1, respectively. For Poisson noise, we add 1,3, and 5% noise with $\Lambda = 30$ and $\Lambda = 20$, respectively. Prior to the iterative process, we initialize the coefficient matrix W as a random $d \times q$ matrix, and variance $\sigma^2 = 1$. We train our face model based on these datasets with noise, and reconstruct their corresponding 3D face shapes from the landmarks.

5.3 Experimental results

In the comparison with non-probabilistic methods, 120 3D face shapes are reconstructed from the landmarks. The output of the 3DMM method consists of 53,215 vertices and the output of the PRNet-based method consists of 43,867 vertices. Since the reconstructed faces obtained by different methods consist of different number of vertices, we remove parts of the reconstructed faces prior to measuring the distance between the reconstructed face and the groundtruth. The rigid iterative closest point (ICP) algorithm is applied to register reconstructed 3D face and groundtruth, to better compare the reconstruction performance. We use the average projection distance of the groundtruth point to the reconstructed mesh as the reconstruction loss. Table 2 shows the comparison of our method with 3DMM method, PCA-based method and PRNet-based method.

The quantitative results of the average reconstruction loss, the variance and the Root Mean Squared Error (RMSE) are presented in Table 2, and we visualize the reconstruction loss of some reconstructed faces in Fig. 4. The outputs of 3DMM show little differences between each other, possibly due to the expressionless input landmarks. The PRNet-based method shows a slightly worse performance than our method. From the reconstructed faces obtained by our method, more details can be observed compared to the faces reconstructed by other methods. For instance, the areas of nose, mouth and eyes are clearer. Though there are less sample faces, the application of PPCA finds out general features and extracts more detailed information from small sample size datasets. This is due to FA can preserve information greatly influence observed data, meanwhile, PCA can preserve detailed information, and the combination of FA and PCA can preserve important information in small size datasets. According to the quantitative metrics and the reconstructed results shown in Table 2 and Fig. 4, our proposed method is effective.

We further evaluate the reconstruction performance of the proposed method on different datasets as shown in Table 2. Figures 5, 6 show the reconstruction loss of the faces in the Chinese craniofacial database and FaceWarehouse, respectively. The blue and orange lines in Figs. 5, 6 represent the face reconstruction performance using our method and the traditional non-probabilistic method (PCA method), respectively. As shown in Fig. 5, our method has similar performance to that of the traditional non-probabilistic method. Among

Table 2 The comparison of different methods on different datasets

Method	Reconstruction Error		
	Mean	Variance	RMSE
	FaceWarehouse		
3DMM [22]	1.70×10^{-2}	1.49×10^{-5}	1.75×10^{-2}
PCA	6.24×10^{-3}	2.54×10^{-6}	6.44×10^{-3}
PRNet [7]	9.84×10^{-3}	3.81×10^{-6}	1.00×10^{-2}
ours	5.04×10^{-3}	1.35×10^{-6}	5.18×10^{-3}
	Chinese craniofacial database		
PCA	1.02×10^{-2}	7.73×10^{-6}	1.06×10^{-2}
ours	1.00×10^{-2}	7.27×10^{-6}	1.04×10^{-2}
	FaceWarehouse		
PCA	2.39×10^{-2}	2.30×10^{-5}	2.44×10^{-2}
ours	2.09×10^{-2}	1.53×10^{-5}	2.13×10^{-2}

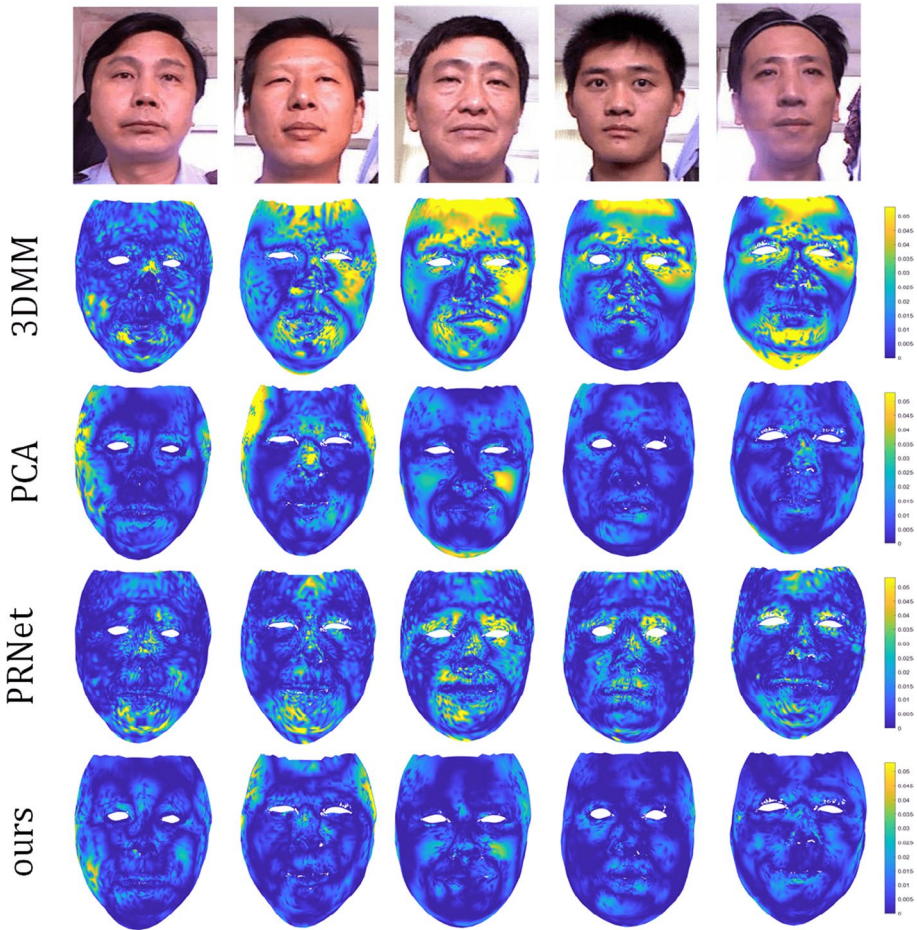


Fig. 4 Different methods for face reconstruction on FaceWarehouse. The first row shows original 2D image of the reconstructed faces. The second row to the fifth row show the reconstruction error of 3DMM, PCA, PRNet and our method, respectively

Fig. 5 Face reconstruction loss on the Chinese craniofacial database

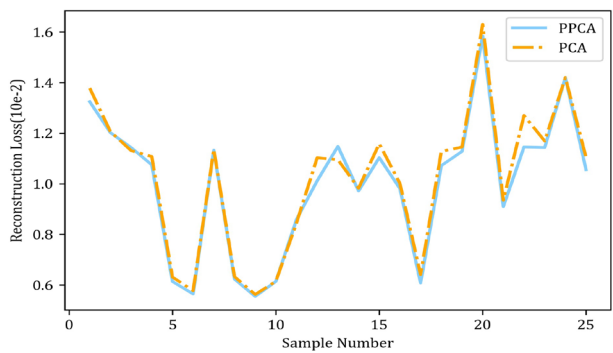
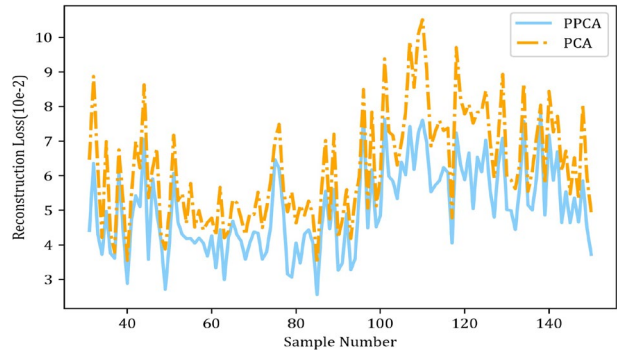


Fig. 6 Face reconstruction loss on FaceWarehouse



25 faces in the Chinese craniofacial database, the average reconstruction loss of our method is 1.00×10^{-2} , and while for the method based on PCA, it is 1.02×10^{-2} . Since the PPCA method is derived from the PCA method, when the estimated parameter σ^2 approaches 0, the PPCA method transforms into the PCA method. For the reconstruction of faces in FaceWarehouse, the estimated parameters σ_L^2 and σ_S^2 are 1.016×10^{-5} and 9.86×10^{-6} , respectively. Figure 6 demonstrates that the reconstruction loss of our method is lower than that of the method based on PCA in most cases, so that our method shows better performance on FaceWarehouse, reducing the average reconstruction loss by 12.47%, which indicates the feasibility of our method.

We evaluate the performance of our method on the datasets with different intensities of noise. As shown in Figs. 7 and 8, when the noise is added to training faces, the heat map of reconstructed loss contains more yellow areas. Upon adding 5% Gaussian noise

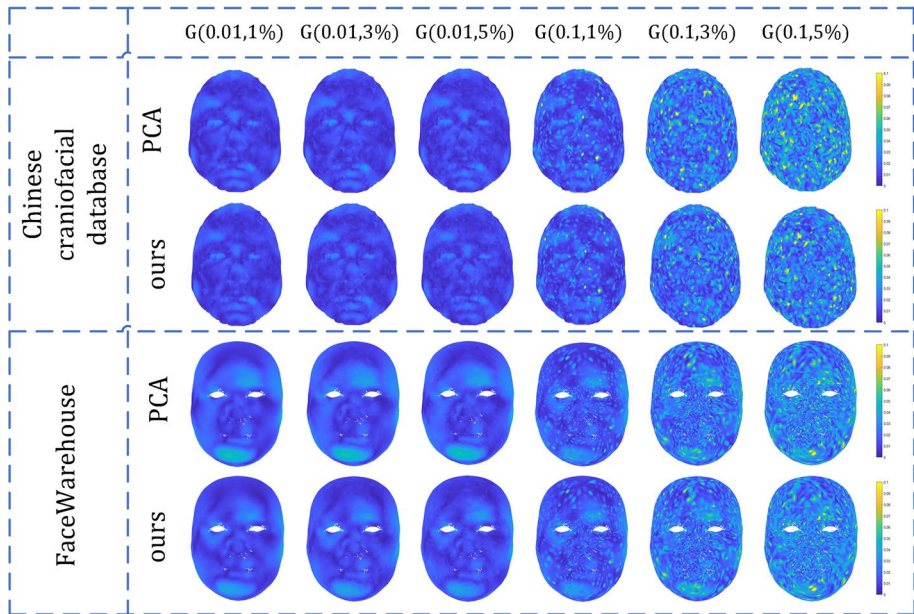


Fig. 7 Reconstructed faces from the databases with added Gaussian noise of different intensities. $G(\sigma^2, p)$ denotes datasets added p Gaussian noise with variance of σ^2 and mean of 0

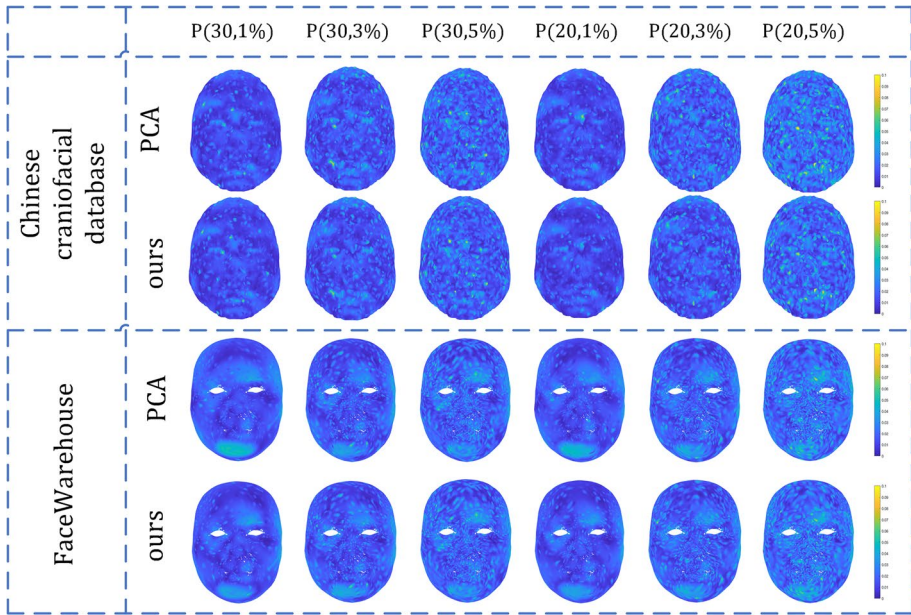
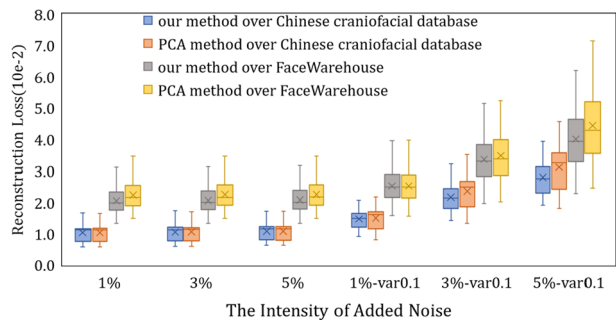


Fig. 8 Reconstructed faces from the databases with added Poisson noise of different intensities. $P(\lambda, p)$ denotes datasets added p Poisson noise with λ

with variance of 0.1 to training faces, the specific reconstruction loss obtained using our method is 2.81×10^{-2} , while it is 3.30×10^{-2} for the PCA-based method, corresponding to the reduction of the reconstruction by 14.85%. Upon adding 5% Poisson noise with $\lambda = 20$ on training face, the specific reconstruction loss obtained using our method is 2.12×10^{-2} , while it is 2.37×10^{-2} for PCA-based method, corresponding to the reduction of the reconstruction loss by 10.46%, showing more blue area in the heat maps when compared to traditional non-probabilistic method. In other words, the accuracy for the reconstruction of 3D faces of our method is higher than that of the traditional non-probabilistic method. Different from the traditional non-probabilistic method, our method constructs the face probabilistic space based on the probability distribution and preserves more principal information when establishing the 3D face model. Even when it is based on few sample faces, our method provides better reconstruction results, particularly for the details of nose, mouth and eyes.

Fig. 9 The reconstruction loss over databases added noise of different intensities



We also plot error boxplots as shown in Fig. 9. Among the 25 sample faces in the Chinese craniofacial database, with the increasing added noise intensity, the reconstruction loss also increases, and our method tends to give rise to lower reconstruction loss. When adding different types of noise to the training datasets, our method shows a better performance than the PCA-based reconstruction method. These reconstruction differences arise because our method considers the probabilistic distribution of the face when extracting the principal components from the sparse points and face point cloud spaces. Besides, we can also observe that our method prefer to have a more stable performance on faces with noise as the variance of reconstruction error is smaller than other method. Therefore, our method has low sensitivity for the noise, and the reconstructed faces are more accurate and detailed.

We increase the proportion of the added noise on the training database for further experiments. We add 10% Gaussian noise to the Chinese craniofacial database and FaceWarehouse and compare the resulting faces. In Figs. 10, 11, we observe that the reconstructed face using our method has a smaller difference between fitted face and the groundtruth

Fig. 10 The reconstruction loss using two methods over Chinese craniofacial database added 10% Gaussian noise

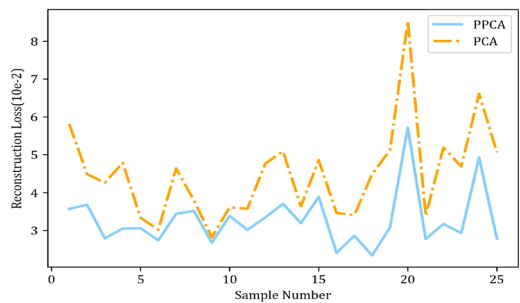


Fig. 11 The reconstruction loss using two methods over FaceWarehouse added 10% Gaussian noise

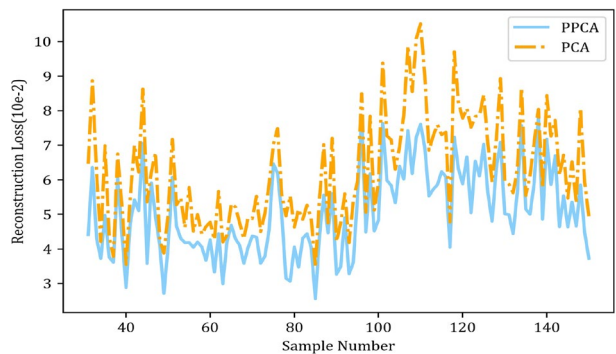


Table 3 The reconstruction loss of over Chinese craniofacial database with noise of different intensities

Intensity of Noise	PPCA	PCA
Var0.01(30%)	1.30×10^{-2}	1.30×10^{-2}
Var0.01(50%)	1.42×10^{-2}	1.50×10^{-2}
Var0.1(10%)	3.54×10^{-2}	4.50×10^{-2}
Var0.1(30%)	6.18×10^{-2}	8.19×10^{-2}
Var0.1(50%)	7.84×10^{-2}	1.09×10^{-1}

when compared with other method, and the average reconstruction loss is reduced by 27.02% and 18.88% respectively. Compared with other method, our proposed method constructs the probabilistic space and can reduce the influence of noise when training face datasets. Due to this, as the noise intensity increases, our reconstruction loss is smaller than other method. Table 3 demonstrates the specific reconstruction loss on Chinese craniofacial database with noise of different intensities. Furthermore, the complexity of our method in face model establishment module is related to the number of sampled faces and the number of their vertices, and it is linear in 3D face reconstruction module. Since the implementation of our method is based on EM algorithm, which accomplishes its convergence through estimating parameters iteratively, there exists difficulty in quantitative complexity analysis for our method. We would like to research on its quantitative complexity analysis in our future work.

6 Conclusion

In this paper, we propose a method to reconstruct the dense point representation of 3D face based on its sparse representation with few sample faces. Compared with previous work, our method constructs the probabilistic spaces of landmarks and 3D face shape. The combination of PCA and FA is effective on extracting principal components which are corresponding to latent variables. The application of PPCA method helps the low-dimensional subspaces of sparse points and face point cloud preserve more details, then we establish our 3D face model. In the experiments, we demonstrate that the combination of PCA and FA is effective and endow our method with adaptability in the reconstruction of the 3D face shape. Further experiments on datasets with noise indicate that our method is more stable than the traditional non-probabilistic method. However, in face model establishment module, the use of an iterative approach give rise to the high complexity of our method. It is left for further research to reduce the time cost to establish face model.

Acknowledgements This work is partially supported by National Natural Science Foundation of China (No.61972041, No.62072045), the National Key Cooperation between the BRICS of China (No.2017YFE0100500), National Key R&D Program of China (No.2017YFB1002604, No.2017YFB1402105).

References

1. Aji OP, Purnama IKE, Yuniarno EM (2019) Facial model deformation based on landmarks using laplacian. In: 2019 International conference on computer engineering, network, and intelligent multimedia (CENIM), IEEE, pp 1–6
2. Blanz V, Vetter T (1999) A morphable model for the synthesis of 3d faces. In: Proceedings of the 26th annual conference on Computer graphics and interactive techniques, pp 187–194
3. Cao C, Weng Y, Zhou S, Tong Y, Zhou K (2013) Facewarehouse: A 3d facial expression database for visual computing. *IEEE Transactions on Visualization and Computer Graphics* 20(3):413–425
4. Chen Y, Wu F, Wang Z, Song Y, Ling Y, Bao L (2020) Self-supervised learning of detailed 3d face reconstruction. *IEEE Transactions on Image Processing* 29:8696–8705
5. Deng Q, Zhou M, Shui W, Wu Z, Ji Y, Bai R (2011) A novel skull registration based on global and local deformations for craniofacial reconstruction. *Forensic Science International* 208(1–3):95–102
6. Duan F, Huang D, Tian Y, Lu K, Wu Z, Zhou M (2015) 3d face reconstruction from skull by regression modeling in shape parameter spaces. *Neurocomputing* 151:674–682

7. Feng Y, Wu F, Shao X, Wang Y, Zhou X (2018) Joint 3d face reconstruction and dense alignment with position map regression network. In: Proceedings of the european conference on computer vision (ECCV), pp 534–551
8. Ferrari C, Berretti S, Pala P, Del Bimbo A (2018) Learning 3dmm deformation coefficients for rendering realistic expression images. In: International conference on smart multimedia, Springer, pp 320–333
9. Geraci M, Farcomeni A (2016) Probabilistic principal component analysis to identify profiles of physical activity behaviours in the presence of non-ignorable missing data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 65(1):51–75
10. Gruszczynski M, Klos A, Bogusz J (2019) A filtering of incomplete gnss position time series with probabilistic principal component analysis. In: Geodynamics and earth tides observations from global to micro scale, Springer, pp 247–273
11. Gu M, Shen W (2020) Generalized probabilistic principal component analysis of correlated data. *Journal of Machine Learning Research* 21(13):1–41
12. Guo X, Li S, Yu J, Zhang J, Ma J, Ma L, Liu W, Ling H (2019) Pflfd: a practical facial landmark detector. arXiv:190210859
13. Hu X, Wang Y, Zhu F, Pan C (2016) Learning-based fully 3d face reconstruction from a single image. In: 2016 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, pp 1651–1655
14. Hubert M, Rousseeuw PJ, Vandenberghe K (2005) Robpca: a new approach to robust principal component analysis. *Technometrics* 47(1):64–79
15. Jhanani R, Harshitha S, Kalaichelvi T, Subedha V (2020) Mobile application for human facial recognition to identify criminals and missing people using tensor flow pp 16–20
16. King DE (2009) Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research* 10:1755–1758
17. Knothe R, Romdhani S, Vetter T (2006) Combining pca and lfa for surface reconstruction from a sparse set of control points. In 7th international conference on automatic face and gesture recognition (FGR06), pp 637–644, <https://doi.org/10.1109/FGR.2006.31>
18. Liu F, Zeng D, Zhao Q, Liu X (2016) Joint face alignment and 3d face reconstruction. In: European conference on computer vision, Springer, pp 545–560
19. Liu P, Han X, Lyu M, King I, Xu J (2020) Learning 3d face reconstruction with a pose guidance network. In: Proceedings of the asian conference on computer vision
20. Mredhula L, Dorairangaswamy M (2016) An effective filtering technique for image denoising using probabilistic principal component analysis (ppca). *Journal of Medical Imaging and Health Informatics* 6(1):194–203
21. Paysan P, Knothe R, Amberg B, Romdhani S, Vetter T (2009a) A 3d face model for pose and illumination invariant face recognition. In: 2009 sixth IEEE international conference on advanced video and signal based surveillance, IEEE, pp 296–301
22. Paysan P, Knothe R, Amberg B, Romdhani S, Vetter T (2009b) A 3d face model for pose and illumination invariant face recognition. In: 2009 Sixth IEEE international conference on advanced video and signal based surveillance, IEEE, pp 296–301
23. Schölkopf B, Smola A, Müller KR (1997) Kernel principal component analysis. In: International conference on artificial neural networks, Springer, pp 583–588
24. Shah SMS, Batool S, Khan I, Ashraf MU, Abbas SH, Hussain SA (2017) Feature extraction through parallel probabilistic principal component analysis for heart disease diagnosis. *Physica A: Statistical Mechanics and its Applications* 482:796–807
25. Tekalp AM, Ostermann J (2000) Face and 2-d mesh animation in mpeg-4. *Signal Processing: Image Communication* 15(4–5):387–421
26. Thomas D (2020) Real-time simultaneous 3D head modeling and facial motion capture with an RGB-D camera. arXiv:2004.10557
27. Tipping ME, Bishop CM (1999) Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 61(3):611–622
28. Tran L, Liu X (2018) Nonlinear 3d face morphable model. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)
29. Tran L, Liu F, Liu X (2019) Towards high-fidelity nonlinear 3d face morphable model. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)
30. Tuan Tran A, Hassner T, Masi I, Medioni G (2017) Regressing robust and discriminative 3d morphable models with a very deep neural network. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 5163–5172

31. Vaddi R, Manoharan P (2018) Probabilistic pca based hyper spectral image classification for remote sensing applications. In: International conference on intelligent systems design and applications, Springer, pp 863–869
32. Wold S, Esbensen K, Geladi P (1987) Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* 2(1–3):37–52
33. Wu CJ (1983) On the convergence properties of the em algorithm. *Ann Stat* :95–103
34. Xiao Q, Han L, Liu P (2014) 3d face reconstruction via feature point depth estimation and shape deformation. In: 2014 22nd international conference on pattern recognition, IEEE, pp 2257–2262
35. Yi H, Li C, Cao Q, Shen X, Li S, Wang G, Tai YW (2019) Mmface: A multi-metric regression network for unconstrained face reconstruction. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 7663–7672
36. Zhang K, Zhang Z, Li Z, Qiao Y (2016) Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters* 23(10):1499–1503
37. Zhong Y, Pei Y, Li P, Guo Y, Ma G, Liu M, Bai W, Wu W, Zha H (2020) Face denoising and 3d reconstruction from a single depth image. In: 2020 15th IEEE international conference on automatic face and gesture recognition (FG 2020), pp 117–124, <https://doi.org/10.1109/FG47880.2020.00005>
38. Zhu X, Lei Z, Liu X, Shi H, Li SZ (2016) Face alignment across large poses: A 3d solution. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR)
39. Zou H, Hastie T, Tibshirani R (2006) Sparse principal component analysis. *Journal of Computational and Graphical Statistics* 15(2):265–286

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.