

# 基于优化三维变形模型参数的人脸重建方法

张红颖\*, 杨维民, 王汇三

中国民航大学电子信息与自动化学院, 天津 300300

**摘要** 针对基于三维(3D)变形模型的人脸图像重建方法对人脸特征点检测不准确导致的重构模型形状表达能力不稳健问题,提出了一种优化 3D 变形模型参数的重建方法。首先,通过改进的位置映射图网络准确提取和定位人脸特征点,并以此为基础得到初始模型参数。然后,为了提高模型的精准度和泛化能力,融合基于回归方法得到的参数获取优化的模型参数。最后,对 3D 变形模型进行优化,得到最终的人脸模型。用真实人脸作为实验数据的结果表明,本方法能够实现精确的 3D 人脸重建。

**关键词** 机器视觉; 三维人脸重建; 三维变形模型; 人脸特征点; 模型参数

**中图分类号** TP391.4

**文献标志码** A

**doi:** 10.3788/LOP202158.2015008

## Face Reconstruction Method Based on Optimized Three-Dimensional Morphable Model Parameters

Zhang Hongying\*, Yang Weimin, Wang Huisan

College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China

**Abstract** This paper proposes a reconstruction method for optimized three-dimensional (3D) morphable model parameters to address the instability of the reconstructed model's shape expression ability due to the inaccurate detection of facial feature points in the face-image reconstruction method based on the 3D morphable model. First, an improved position map regression network is used to extract and locate the facial feature points accurately and on this basis, the initial model parameters are calculated. The parameters obtained using the regression method are then integrated to obtain the optimised model parameters and improve the model's accuracy and generalisation capabilities. Finally, to obtain the final face model, the 3D morphable model is optimized. This approach can produce accurate 3D face reconstruction using real faces as experimental data.

**Key words** machine vision; three-dimensional face reconstruction; three-dimensional morphable model; facial feature points; model parameters

**OCIS codes** 150.6910; 100.6890; 100.3008

## 1 引言

根据单张人脸图像重建出三维(3D)人脸,是 3D 人脸重建领域中的一个热点及难点问题,目前已在多个领域中得到了广泛的应用<sup>[1-3]</sup>。3D 人脸重建方法可分为基于阴影恢复形状(SFS)的方法<sup>[4]</sup>、基于 3D 变形模型(3DMM)匹配的方法<sup>[5]</sup>以及基于回

归的方法<sup>[6]</sup>。其中,3DMM 方法具有面部表达能力强、面部重建完整的优势,被广泛应用于计算机视觉<sup>[5]</sup>、计算机图形学<sup>[7]</sup>、人类行为分析<sup>[8]</sup>等领域。

基于 3DMM 匹配的方法,将二维(2D)人脸图像上的特征点作为输入信息,结合 3D 人脸特征点求解模型参数,然后通过建立 3D 人脸的线性组合重构出 3D 人脸。该方法可以重构出完整的人脸模

收稿日期: 2020-08-03; 修回日期: 2020-09-13; 录用日期: 2020-09-23

基金项目: 国家重点研发计划(2018YFB1601200)、四川省青年科技创新研究团队专项(2019JDTD0001)

通信作者: \*carole\_zhang0716@163.com

型且可塑性强,但重建模型的精度与泛化能力依赖于高质量的模型参数,且模型参数的质量受特征点位置的影响较大。为了提高特征点的精准度,进而提高模型参数质量,Dou 等<sup>[9]</sup>提出了一种双耦合结构学习方法,可自动选取 2D 和 3D 特征点,使提取特征点时具有更高的容错率,但该方法在计算模型参数时容易陷入局部最优,导致重建模型的精度较差。Romdhani 等<sup>[10]</sup>为了降低特征点对模型参数的影响,在重建过程中加入了像素强度、人脸边缘等特征并对重建过程进行优化。但该方法加入的特征易受噪声影响,且计算量过大。近年来,随着深度学习的发展,人们结合神经网络提出了基于回归的方法,通过单张 2D 人脸图像回归出模型参数。如 Zhu 等<sup>[11]</sup>提出的 3D 密集人脸对齐(3DDFA)方法,通过建立 3DMM 与 2D 图像之间的关联,设计了两种视图特征和损失函数,然后采用双通道卷积神经网络(CNN)回归模型参数。但在回归参数过程中,该方法用圆柱体近似代替 3DMM 将 3D 图像投影到 2D 图像上,提取的特征信息冗余较多,模型参数的质量较差;且该方法采用平均人脸模型迭代更新模型参数,导致重建后的人脸模型泛化能力不足。综上所述,利用 3DMM 构建出的人脸模型精准度与泛化能力受模型参数质量的影响较大。

针对上述问题,本文将基于 3DMM 匹配的方法与 3DDFA 方法相结合,提出了一种优化模型参数的方法。首先,用改进的位置映射图网络(PRN)提取人脸特征点,并在此基础上利用 3DMM 方法求解模型参数。然后,将改进 PRN 回归出的模型参数与 3DDFA 方法回归出的模型参数进行融合,进一步得到优化的模型参数。最后,用优化的模型参数对 3DMM 进行优化,得到精准的人脸模型。

## 2 基于 3DMM 的重建方法

基于 3DMM 的重建方法将 3DMM 映射到 2D 人脸图像上,通过减小映射点与 2D 图像特征点之间的误差对 3DMM 进行优化,从而实现 3D 人脸重建。其中,3DMM 由形状向量和纹理向量组成,形状向量  $\mathbf{S} = (X_1, Y_1, Z_1, \dots, X_n, Y_n, Z_n)^T \in \mathbf{R}^{3 \times n}$  为笛卡儿坐标系中的 3D 坐标;纹理向量  $\mathbf{T} = (R_1, G_1, B_1, \dots, R_n, G_n, B_n)^T \in \mathbf{R}^{3 \times n}$  为 3DMM 中对应点的 RGB(Red, Green, Blue)三通道像素值  $R$ 、 $G$ 、 $B$ 。将  $m$  个训练样本的形状向量与纹理向量映射到主成分分析(PCA)空间,并将其线性组合成形状模型  $\mathbf{S}_{\text{Model}}$  和纹理模型  $\mathbf{T}_{\text{Model}}$ ,可表示为

$$\mathbf{S}_{\text{Model}} = \bar{\mathbf{S}} + \sum_{i=1}^{m-1} \alpha_i \mathbf{s}_i, \quad (1)$$

$$\mathbf{T}_{\text{Model}} = \bar{\mathbf{T}} + \sum_{i=1}^{m-1} \beta_i \mathbf{t}_i, \quad (2)$$

式中,  $\bar{\mathbf{S}}$  为平均面部形状模型,  $\bar{\mathbf{T}}$  为平均面部纹理模型,  $\mathbf{s}_i$  为 3DMM 对应的形状向量,  $\mathbf{t}_i$  为 3DMM 对应的纹理向量,  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{m-1})$  为形状参数,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_{m-1})$  为纹理参数。  $\boldsymbol{\alpha}$  的分布满足

$$p(\boldsymbol{\alpha}) \sim \exp \left[ -\frac{1}{2} \sum_{i=1}^{m-1} (\alpha_i / \sigma_i)^2 \right], \quad (3)$$

式中,  $\sigma_i$  为形状协方差矩阵的特征值,  $\boldsymbol{\beta}$  的分布与  $\boldsymbol{\alpha}$  一致。人脸面部存在表情变化,因此,在形状模型中加入表情系数,将人脸模型的线性表示扩充为

$$\mathbf{S}_{\text{Model}} = \bar{\mathbf{S}} + \sum_{i=1}^{m-1} \alpha_{i,\text{id}} \mathbf{s}_{i,\text{id}} + \sum_{i=1}^{n-1} \alpha_{i,\text{exp}} \mathbf{e}_{i,\text{id}}, \quad (4)$$

式中,  $\mathbf{s}_{i,\text{id}}$  为中性表情的第  $i$  个形状特征向量,  $\boldsymbol{\alpha}_{\text{id}} = (\alpha_{1,\text{id}}, \dots, \alpha_{m-1,\text{id}})$  为形状参数,  $\mathbf{e}_{i,\text{id}}$  为第  $i$  个表情的特征向量,  $\boldsymbol{\alpha}_{\text{exp}} = (\alpha_{1,\text{exp}}, \dots, \alpha_{n-1,\text{exp}})$  为表情参数。在构建模型后,利用正交比例投影将 3D 人脸模型投影到 2D 人脸图像上,可表示为

$$\mathbf{X}_{\text{projection}} = s \times \mathbf{P} \times \mathbf{R} \times \mathbf{S}_{\text{Model}} + \mathbf{t}_{2D}, \quad (5)$$

式中,  $\mathbf{X}_{\text{projection}}$  为模型中 3D 顶点投影到 2D 平面的向量,即投影到 2D 平面后对应的坐标点,  $s$  为比例系数,  $\mathbf{P}$  为投影矩阵  $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ ,  $\mathbf{R}$  为旋转矩阵,  $\mathbf{t}_{2D}$  为位移矩阵。

为了使 3DMM 更接近真实人脸,求解投影的 2D 平面向量与给定的 68 个 2D 对应点坐标  $\mathbf{X}$  的最小误差,可表示为

$$\underset{s, \mathbf{R}, \boldsymbol{\alpha}_{\text{id}}, \boldsymbol{\alpha}_{\text{exp}}, \mathbf{t}_{2D}}{\operatorname{argmin}} \|\mathbf{X}_{\text{projection}} - \mathbf{X}\|^2 + \lambda \sum_{i=1}^n \left( \frac{\gamma_i}{\sigma_i} \right)^2, \quad (6)$$

式中,  $\gamma$  为 PCA 系数,包括参数  $\boldsymbol{\alpha}_{\text{id}}$  和  $\boldsymbol{\alpha}_{\text{exp}}$ ,  $\lambda$  为常数。可以发现,基于 3DMM 匹配的重建方法需要精准的人脸特征点  $\mathbf{X}$ ,且其重建精准度依赖于高质量的模型参数  $\boldsymbol{\alpha}_{\text{id}}$  和  $\boldsymbol{\alpha}_{\text{exp}}$ 。

## 3 基于优化模型参数的人脸重建

为了提高重建模型的精准度与泛化能力,使重建图像更接近真实人脸,提出了一种优化模型参数的重建方法。首先,利用改进的 PRN 提取人脸特征点。然后,将特征点求解的模型参数与回归的参数进行融合,得到优化的模型参数。最后,根据模型参数对 3DMM 进行优化,最大程度重建真实的人脸。本方法的 3D 人脸重建过程如图 1 所示。

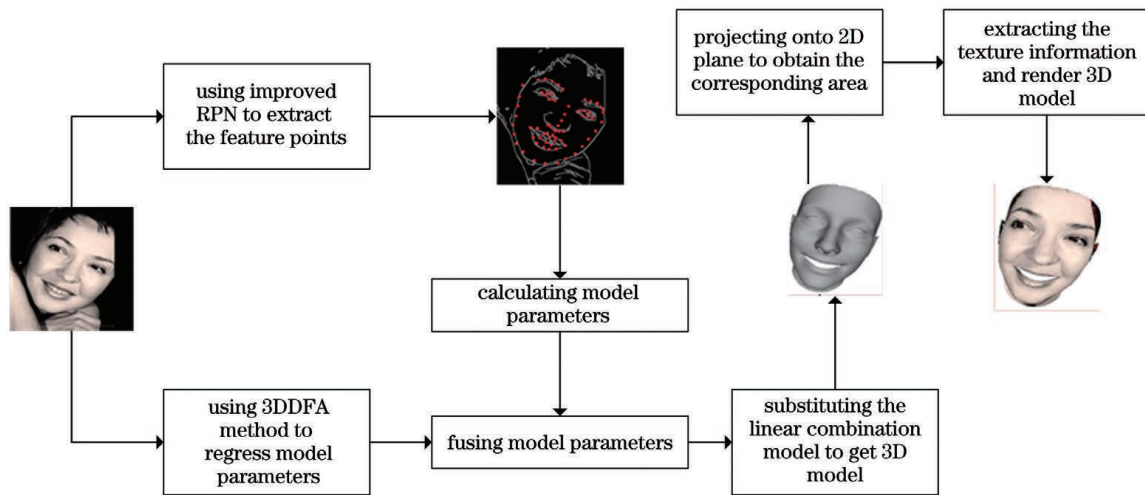


图 1 本方法的流程图

Fig. 1 Flow chart of our method

### 3.1 人脸特征点提取

当人脸图像为正面人脸图时,传统方法能很好地提取人脸特征点,如主动形状模型(ASM)<sup>[12]</sup>,ASM提取的特征点如图2所示。但在人脸姿态变化情况下,ASM无法对特征点进行准确提取。PRN中的纹理颜色坐标(UV)位置图记录了模型的3D坐标点,可以有效定位特征点,提高特征点提取的精度,PRN与ASM提取的特征点如图3所示。可以发现,ASM在人脸出现姿态变化时,特征点提取效果较差,特别是对鼻子和人脸轮廓特征

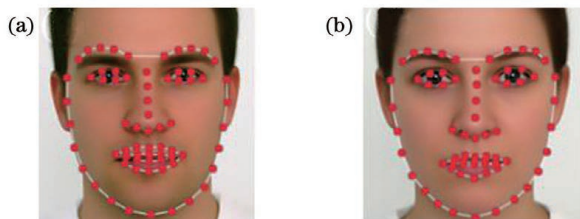


图 2 ASM提取的正面人脸特征点。(a)正面人脸1;  
(b)正面人脸2

Fig. 2 Positive face feature points extracted by ASM.

(a) Positive face 1; (b) positive face 2

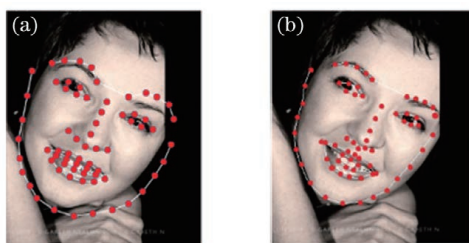


图 3 不同方法提取的特征点。(a)ASM;(b)PRN

Fig. 3 Feature points extracted by different methods.

(a) ASM; (b) PRN

点的提取。而PRN提取的特征点更贴近于实际中的位置,因此,本方法采用PRN提取人脸的68个特征点。

从图3(b)可以发现,PRN提取的特征点中部分特征点存在较大偏差。为了提高特征点的精度,本方法结合人脸边缘特征对已提取的特征点进行校准。Canny算子不易受噪声干扰,在光照条件不理想的环境下,也可以对人脸边缘信息进行提取,因此,采用Canny算子进行边缘提取,并利用高斯滤波进一步减少噪声的影响,得到处理后的人脸边缘特征如图4所示。将带有人脸特征点与人脸边缘的图像叠加在一起,以校准特征点,结果如图5所示。



图 4 基于Canny算子提取的人脸边缘特征。(a)人脸图像;(b)人脸边缘

Fig. 4 Facial edge features extracted based on Canny operator. (a) Face image; (b) face edge

人脸边缘特征包含了精准的特征点信息,为了使提取出的特征点尽可能贴近人脸边缘,设计了一个 $3 \times 3$ 的检测器,以检测距离特征点最近的边缘位置,检测器的结构如图6所示。其中, $(x_n, y_n)$ 为待调整的特征点, $(x_{fo}, y_{fo})$ 为临近特征点, $(x_e, y_e)$ 为边缘坐标点。

在叠加图中,人脸边缘和特征点存在相对应

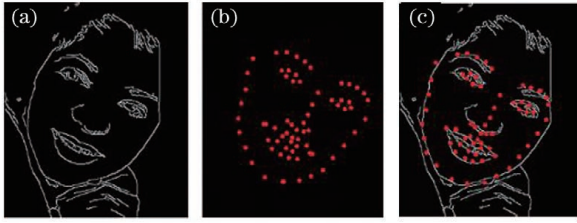


图 5 人脸边缘与人脸特征点的叠加。(a)人脸边缘;  
(b)人脸特征点;(c)叠加图

Fig. 5 Superposition of face edge and face feature points. (a) Face edge; (b) face feature points; (c) superimposed image

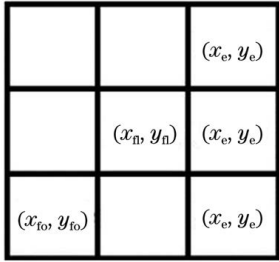


图 6 边缘特征检测器的结构

Fig. 6 Structure of the edge feature detector

的像素坐标,因此,利用欧氏距离计算两者之间的距离。为了防止检测器检测到错误边缘,令待调整特征点与邻近特征点的距离最大,待调整特征点与边缘特征的距离最小,满足的约束条件可表示为

$$d_f^j = \max \sqrt{(x_n^j - x_{fo}^j)^2 + (y_n^j - y_{fo}^j)^2}, \quad (7)$$

$$d_e^j = \min \sqrt{(x_n^j - x_e^j)^2 + (y_n^j - y_e^j)^2}, \quad (8)$$

式中,  $1 \leq j \leq 68, 1 \leq i \leq n$ ,  $d_f^j$  为待调整特征点到邻近特征点的距离,  $d_e^j$  为待调整特征点到边缘的距离。由于检测器在执行检测任务时会对所有特征点进行遍历,速度较慢,可通过缩小检测器的检测范围,提高执行效率。因此,增加约束条件

$$\begin{cases} f(x, y) = (x - x_n^j)^2 + (y - y_n^j)^2 + \frac{d_{\min}^2}{4}, & (9) \\ f(x, y) \leq 0 \end{cases}$$

$$d_{\min} = \min(d_{j,j-1}, d_{j,j+1}), \quad (10)$$

式中,  $d_{\min}$  为待调整特征点与相邻两个特征点的最小距离,当  $i$  分别为 1 和 68 时,  $d_{\min} = d_{j,j+1}$  和  $d_{j,j-1}$ 。检测器的执行步骤如下。

1) 输入待调整特征点  $(x_n, y_n)$ , 取第  $j$  个特征点作为被调整特征点; 然后根据 (9) 式、(10) 式计算所需检测的范围。

2) 在计算的检测范围中, 计算特征点  $(x_n, y_n)$  与邻近特征点、边缘特征的距离。

3) 根据 (7) 式、(8) 式选出满足条件的边缘特征点  $(x_e, y_e)$ , 并代替特征点  $(x_n, y_n)$ 。

4) 重复步骤 1)~步骤 3), 校准所有特征点。

5) 校准完所有特征点后结束执行, 并输出校准后特征点。校准后的效果如图 7 所示, 可以发现, 校准后的特征点比校准前更精准。

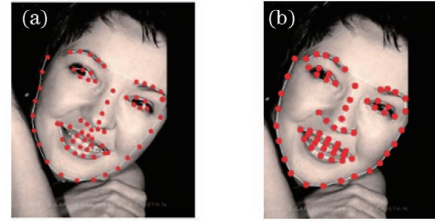


图 7 校准前后的特征点。(a)校准前;(b)校准后  
Fig. 7 Feature points before and after calibration.

(a) Before calibration; (b) after calibration

### 3.2 模型参数融合

为了提高 3DMM 的参数质量, 使重建的人脸特征表达能力更强, 结合 3DDFA 方法回归模型参数, 并将该参数与基于 3DMM 匹配方法求解的参数进行融合, 得到精度更高的参数, 从而提高模型的性能。首先, 将所有参数放在集合  $p = [s, R, t_{2D}, \alpha_{id}, \alpha_{exp}]^T$  中, 以回归 3DMM 的参数。在回归过程中, 3DDFA 方法采用的网络模型在提取人脸特征时会损失部分信息, 且在训练过程中占用的内存较高。而 MobileNet-V2<sup>[13]</sup> 通过去掉线性整流单元 (ReLU) 函数中的 Narrow layer 保留了特征的多样性, 增强了网络的表达能力。同时, 引入残差结构, 通过先升维再降维的操作和增强梯度的传播方式减少训练时所需的内存。因此, 本方法引入 MobileNet-V2 对参数进行回归, 参数更新过程可表示为

$$\Delta p^k = M^k(p^k, I), \quad (11)$$

式中,  $M$  为网络模型,  $k$  为迭代次数,  $0 \leq k \leq K, I$  为输入图像。回归后的参数包含比例系数  $s$ 、旋转矩阵  $R$  和位移矩阵  $t_{2D}$ 。基于 3DMM 匹配方法的重建中, 直接利用校准后的特征点与正交比例投影求解参数  $\alpha_{id}$  和  $\alpha_{exp}$ , 舍弃了投影到 2D 平面的向量与给定的 68 个 2D 对应点  $X$  的最小误差, 从而缩短重建时间。首先, 将  $\sum_{i=1}^{m-1} \alpha_{i,id} s_{i,id}$  置为 0, 将除  $\alpha_{id}$  和  $\alpha_{exp}$  以外的参数代入 (5) 式, 得到参数  $\alpha_{id}$

$$\alpha_{id} = \left( \sum_{i=1}^{m-1} s_{i,id} \right)^{-1} \times \left[ R^{-1} \times P^{-1} \times s^{-1} \times (X_{\text{projection}} - t_{2D}) - \bar{S} \right]. \quad (12)$$

同理, 将  $\sum_{i=1}^{n-1} \alpha_{i, \text{exp}} e_{i, \text{id}}$  置为 0, 代入参数得到  $\alpha_{\text{exp}}$ , 将两种方法求解出的  $\alpha_{\text{id}}$ 、 $\alpha_{\text{exp}}$  进行融合, 可表示为

$$\alpha_{\text{new}} = (1 - z) \times \alpha_{\text{rg}} + z \times \alpha_{\text{mt}}, \quad (13)$$

式中,  $z$  为融合系数, 其被约束为非负数, 且  $0 \leq z \leq 1$ ,  $\alpha_{\text{rg}}$  为回归得到的模型参数,  $\alpha_{\text{mt}}$  为匹配得到的参数。为了更好地融合不同模型的参数, 通过最小化误差函数求取最优的融合系数。首先, 构造误差函数

$$E(z) = \frac{1}{o} \sum_{i=1}^o \|S_t(i) - S_p(z, i)\|^2, \quad (14)$$

式中,  $S_t(i)$  为数据库中人脸模型的 3D 点,  $S_p(z, i)$  为融合参数后的 3D 点,  $o$  为 3D 点的个数。然后, 用梯度下降法计算误差函数的最小值, 并取其对应的融合系数, 得到优化后的模型参数  $\alpha_{\text{id\_new}}$  和  $\alpha_{\text{exp\_new}}$ 。最后, 将参数代入线性表达的 3D 人脸模型  $S_{\text{Model}}$  中, 得到最终的 3D 人脸模型, 可表示为

$$S_{\text{newModel}} = \bar{S} + \left( \sum_{i=1}^{m-1} \alpha_{\text{id\_new}} s_{i, \text{id}} + \sum_{i=1}^{n-1} \alpha_{\text{exp\_new}} e_{i, \text{id}} \right). \quad (15)$$

由上述分析可知, 用自适应融合系数重建后的模型泛化能力较强, 且重建出的模型精准度更高。

### 3.3 模型渲染

为了使重建后的人脸模型在视觉上更接近真实人脸, 结合 2D 人脸图像中的纹理信息对模型进行渲染。首先, 将 3D 模型映射到 2D 图像上, 建立 2D 坐标与 3D 坐标的对应关系, 如图 8 所示。可以发现, 映射到 2D 图像上的模型刚好覆盖了人脸区域。然后, 利用(5)式求出由 3D 模型映射到 2D 图像中的对应坐标  $X_{\text{projection\_new}}$ , 并在投影区域内提取纹理信息。按建立好的坐标对应关系, 将提取出的纹理信息赋值给人脸模型中的顶点, 可表示为

$$T_{\text{Model}} = \sum_{i=1} [X_{\text{Model}}^i, P(X_{\text{projection\_new}}^i)]. \quad (16)$$

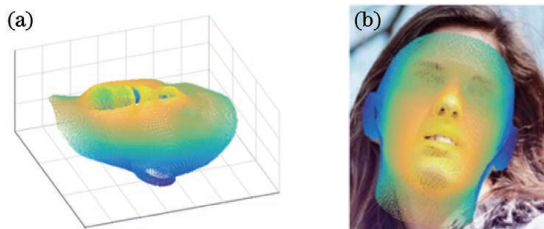


图 8 投影结果。(a)3D 人脸模型;(b)投影图像  
Fig. 8 Projection results. (a) 3D face model;  
(b) projected image

将模型中的三角面进行连接, 得到渲染后的 3D 人脸模型, 如图 9 所示。可以发现, 渲染后的人脸模型比渲染前的人脸模型在视觉上更接近真实人脸。

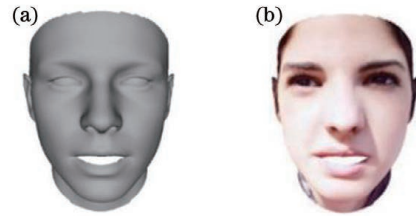


图 9 模型渲染的结果。(a)渲染前;(b)渲染后  
Fig. 9 Result of model rendering. (a) Before rendered;  
(b) after rendered

## 4 实验结果与分析

实验步骤: 首先, 对输入的 2D 人脸图像进行特征点提取, 并用提取后的特征点求解模型参数。然后, 将基于 3DMM 匹配求解的模型参数与基于回归方法求解的模型参数进行自适应融合, 并用融合后的参数进行模型重建。最后, 提取 2D 人脸图像中的纹理图像对 3D 人脸模型进行渲染, 得到最终的 3D 人脸模型。实验包括融合系数分析和结果对比两部分, 使用的数据库包括 AFW<sup>[14]</sup>、IBUG<sup>[15]</sup>、LFPW<sup>[16]</sup>。

### 4.1 融合系数分析

采用自适应方式对模型参数进行融合, 以自动求解出最优融合系数。随机从数据库中选取 5 组人脸模型, 并计算不同融合系数下数据库人脸模型和重建人脸模型之间的归一化均方差 (NMSE), 可表示为

$$\begin{cases} X_{\text{MSE}} = \frac{1}{n} \sum_{t=1}^n (D_t - R_t)^2 \\ N = \frac{X_{\text{MSE}} - M_{\text{min}}}{M_{\text{max}} - M_{\text{min}}} \end{cases}, \quad (17)$$

式中,  $X_{\text{MSE}}$  为均方差,  $D$  为真实值,  $R$  为重建值,  $n$  为总数据量,  $t$  为其中的某一数据,  $M_{\text{min}}$  和  $M_{\text{max}}$  分别为最小误差和最大误差,  $N$  为 NMSE。计算分别使用融合系数  $\lambda_{\text{id}}$  和  $\lambda_{\text{exp}}$  重建出 3D 人脸模型的归一化均方差, 结果如图 10 所示。其中,  $N$  值越小, 表明模型的重建精准度越高。可以发现, 随着  $\lambda_{\text{id}}$  的增大 ( $\alpha_{\text{mt}}$  参数的占比增大),  $N$  逐渐减小, 且  $\lambda_{\text{id}}$  在 0.4~0.5 之间时, 重建出的人脸模型误差最小, 但  $\lambda_{\text{id}}$  超过 0.5 后重建误差逐渐增大。当  $\lambda_{\text{exp}}$  在 0.6~0.65 之间时, 重建出的人脸模型误差最小, 当  $\lambda_{\text{exp}}$  超过 0.65 后重建误差逐渐增大。这表明在整

体外观上,回归方法要优于匹配方法,但在细节刻画上,匹配方法优于回归方法。对比发现,单独使用

$\lambda_{id}$  时的  $N$  小于单独使用  $\lambda_{exp}$  的  $N$ , 这表明参数  $\alpha_{id}$  对重建人脸模型精度的影响较大。

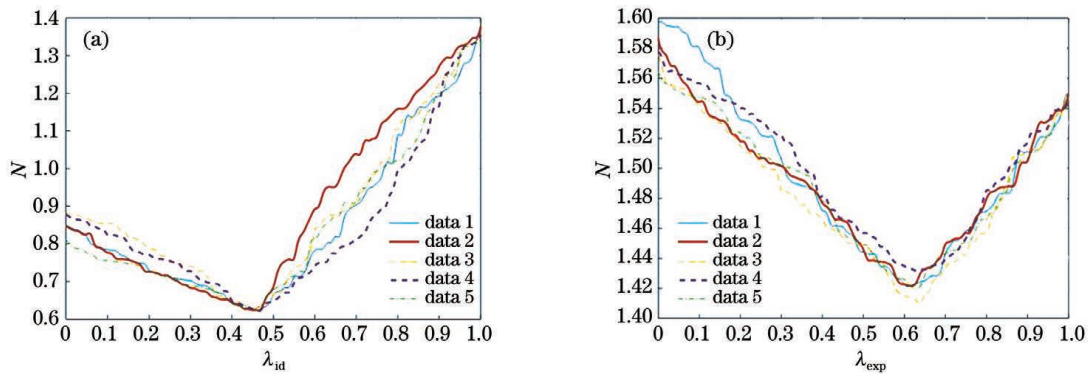


图 10 不同系数下的 NMSE。(a)  $\lambda_{id}$ ; (b)  $\lambda_{exp}$

Fig. 10 NMSE with different coefficients. (a)  $\lambda_{id}$ ; (b)  $\lambda_{exp}$

将两个融合系数同时代入人脸模型,并从三个数据库中随机选取 50 组人脸模型,计算  $N$ , 结果如图 11 所示。其中,插图为基于回归方法和融合后重建的人脸误差。可以发现,相比用单独参数重建的人脸,通过融合参数求解的人脸模型精度有所提高,且误差稳定在 0.60 到 0.62 之间,验证了本方法的可靠性。

#### 4.2 实验结果

首先,用本方法对人脸图像进行重建,并与真实人脸进行对比。图 12 为由单张人脸图像重建出的 3D 人脸模型,同时展示了  $0^\circ$ 、 $30^\circ$ 、 $60^\circ$ 、 $90^\circ$  的人脸模型。其中,本方法生成的 3D 人脸模型具有 53490 个 3D 顶点,160470 个三角形。

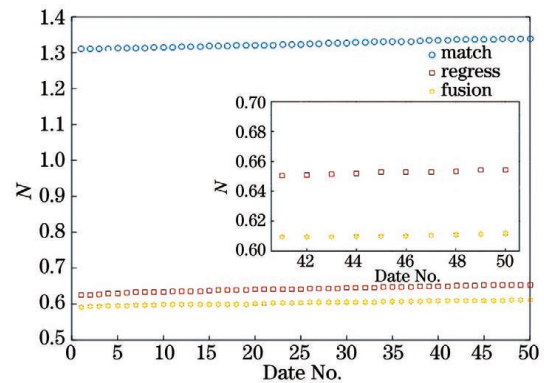


图 11 不同方法重建图像的 NMSE

Fig. 11 NMSE of reconstructed image by different methods

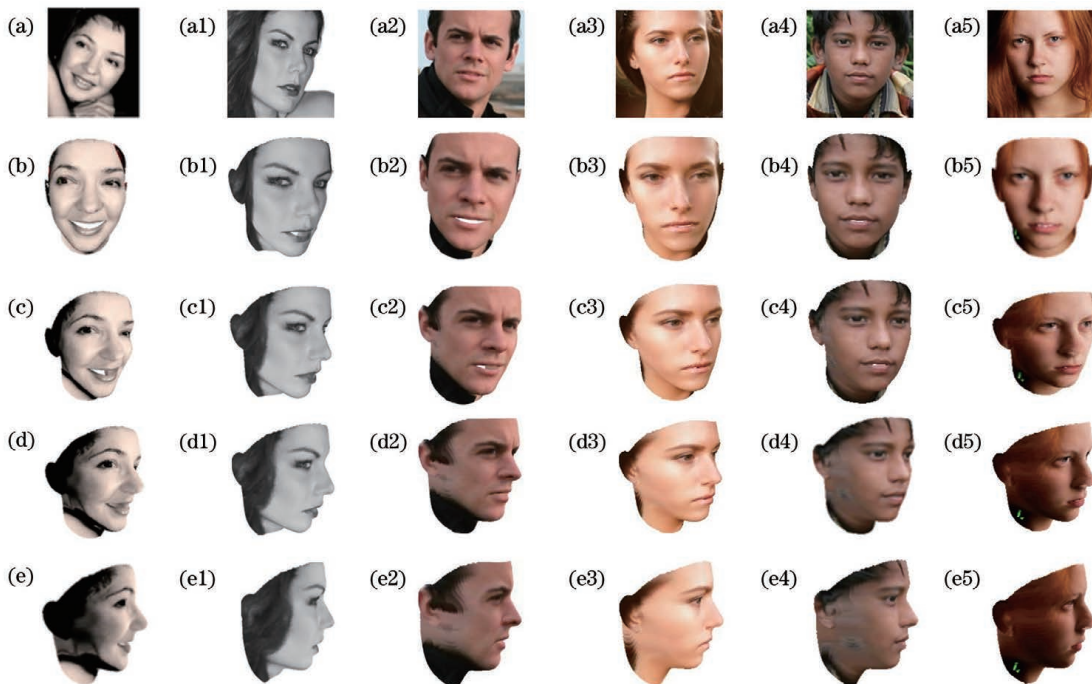


图 12 3D 人脸重建结果。(a) 2D 图像; (b)  $0^\circ$ ; (c)  $30^\circ$ ; (d)  $60^\circ$ ; (e)  $90^\circ$

Fig. 12 3D face reconstruction results. (a) 2D image; (b)  $0^\circ$ ; (c)  $30^\circ$ ; (d)  $60^\circ$ ; (e)  $90^\circ$

为了评估本方法的性能,用 AFW、IBUG、LFPW 数据库对本方法、3DDFA、VRN (Volumetric regression network)<sup>[17]</sup> 和 3DMM-CNN<sup>[18]</sup> 方法进行定性的对比分析,不同方法的重建结果如图 13 所示。可以发现,本方法和 3DDFA 方法重建结果的

差别较小,但明显优于 3DMM-CNN 方法的重建结果,在面部刻画方面的泛化能力更强,更加贴合真实人脸。VRN 方法的重建结果比本方法粗糙,但该方法使用体素数据,在视觉上比本方法更接近真实人脸。



图 13 不同方法的重建结果。(a)2D 图像;(b)VRN 方法;(c)3DDFA 方法;(d)3DMM-CNN 方法;(e)本方法  
Fig. 13 Reconstruction results of different methods. (a) 2D image; (b) VRN method; (c) 3DDFA method;  
(d) 3DMM-CNN method; (e) our method

计算不同方法重建图像的精度,并用迭代最近点算法找到与真实点距离最近的点,然后计算重建后的 3D 坐标与真实 3D 坐标之间的 NMSE,结果如表 1 所示。可以发现,本方法的重建精度总体上优于其他三种对比方法。定量分析与定性分析结果均表明,本方法可提高模型的形状表达能力,实现更精准的人脸重建。

表 1 不同方法重建图像的 NMSE

Method	VRN	3DDFA	3DMM-CNN	Ours
AFW	0.698	0.634	1.223	0.617
IBUG	0.623	0.644	1.056	0.545
LFPW	0.669	0.653	1.326	0.656

## 5 结 论

为了使 3DMM 更精准地表达出真实人脸模型,

对人脸特征点进行了准确提取,并提出了一种优化模型参数的方法。对基于 3DMM 匹配和回归方法求解的模型参数进行融合,提高了参数质量,从而实现了更真实的人脸重建效果。实验结果表明,相比 VRN、3DDFA、3DMM-CNN 方法,本方法在模型泛化能力及精度上的性能更好,且能得到真实的 3D 人脸模型。但本方法依然存在一些不足,如在纹理提取过程中,存在遮挡时对部分纹理特征的提取效果不佳,因此,后续工作中还需对这些问题进行改进。

## 参 考 文 献

- [1] Huang S, Hu Y, Gong C L, et al. High precision three-dimensional face reconstruction algorithm based on gradient light image [J]. Acta Optica Sinica, 2020, 40(4): 0410001.  
黄硕, 胡勇, 巩彩兰, 等. 基于梯度光图像的高精度

- 三维人脸重建算法[J]. 光学学报, 2020, 40(4): 0410001.
- [2] Ge B Z, Zhao D, Tian Q G, et al. Three-dimensional human face laser scanning based on three-line structured light[J]. Chinese Journal of Lasers, 2010, 37(6): 1594-1598.  
葛宝臻, 赵丹, 田庆国, 等. 三光带激光三维人脸扫描方法[J]. 中国激光, 2010, 37(6): 1594-1598.
- [3] Lu J W, Chen H T, Ma X P, et al. 3D printing mask attacks detection based on multi-feature fusion[J]. Laser & Optoelectronics Progress, 2019, 56(3): 031002.  
陆经纬, 陈鹤天, 马肖攀, 等. 基于多特征融合的3D打印面具攻击检测[J]. 激光与光电子学进展, 2019, 56(3): 031002.
- [4] Horn B K P, Brooks M J. The variational approach to shape from shading [J]. Computer Vision, Graphics, and Image Processing, 1986, 33(2): 174-208.
- [5] Blanz V, Vetter T. A morphable model for the synthesis of 3D faces [C] // Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques-SIGGRAPH '99, July 1, 1999, New York, USA. New York: ACM Press, 1999: 187-194.
- [6] Masi I, Trần A T, Hassner T, et al. Do we really need to collect millions of faces for effective face recognition?[M] // Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9909: 579-596.
- [7] Aldrian O, Smith W A P. Inverse rendering of faces with a 3D morphable model[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(5): 1080-1093.
- [8] Amberg B, Knothe R, Vetter T. Expression invariant 3D face recognition with a morphable model [C] // 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition, September 17-19, 2008, Amsterdam, Netherlands. New York: IEEE Press, 2008: 1-6.
- [9] Dou P F, Wu Y H, Shah S, et al. Robust 3D face shape reconstruction from single images via two-fold coupled structure learning and off-the-shelf landmark detectors [C] // Proceedings of the British Machine Vision Conference 2014, September, 2014, Nottingham. British: British Machine Vision Association, 2014: 1-13.
- [10] Romdhani S, Vetter T. Estimating 3D shape and texture using pixel intensity, edges, specular highlights, texture constraints and a prior [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), June 20-25, 2005, San Diego, CA, USA. New York: IEEE Press, 2005: 986-993.
- [11] Zhu X Y, Liu X M, Lei Z, et al. Face alignment in full pose range: a 3D total solution [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(1): 78-92.
- [12] Guo Y Q, He R, Zheng W S, et al. Active shape model based on sparse representation [M] // Zheng W S, Sun Z N, Wang Y H, et al. Biometric recognition. Lecture notes in computer science. Heidelberg: Springer, 2012, 7701: 94-103.
- [13] Sandler M, Howard A, Zhu M L, et al. Inverted residuals and linear bottlenecks: mobile networks for classification, detection and segmentation [EB/OL]. (2018-01-13) [2020-08-01]. <https://arxiv.org/abs/1801.04381v1>.
- [14] Zhu X X, Ramanan D. Face detection, pose estimation, and landmark localization in the wild [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press, 2012: 2879-2886.
- [15] Sagonas C, Tzimiropoulos G, Zafeiriou S, et al. 300 faces in-the-wild challenge: the first facial landmark localization challenge [C] // 2013 IEEE International Conference on Computer Vision Workshops, December 2-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2013: 397-403.
- [16] Belhumeur P N, Jacobs D W, Kriegman D J, et al. Localizing parts of faces using a consensus of exemplars [C] // CVPR 2011, June 20-25, 2011, Colorado Springs, CO, USA. New York: IEEE Press, 2011: 545-552.
- [17] Jackson A S, Bulat A, Argyriou V, et al. Large pose 3D face reconstruction from a single image via direct volumetric CNN regression [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 1031-1039.
- [18] Tran A T, Hassner T, Masi I, et al. Regressing robust and discriminative 3D morphable models with a very deep neural network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1493-1502.