



基于深度三维形变的单张 3D 人脸重建算法

摘要

三维形变模型(3DMM)作为人脸重建的重要方式,在3D建模、图像合成等领域有着广泛的应用。由于受训练数据类型、数量以及主成分等因素影响,3DMM存在过约束的现象,不能提供足够的灵活性来表示高频变形。本文将三维形变模型嵌入到深度神经网络中,为提升3D人脸重建的表示能力提供了新的思路。为了提升网络学习效率,本文构造了一种双通路神经网络,实现了在全局路径和局部路径之间的平衡。通过在学习目标和网络结构两方面改进非线性3DMM,提出了一种比线性或以往的非线性模型更能捕捉到更高层次细节的模型。算法对比与仿真实验表明,本文算法在3D人脸重建上的归一化平均误差更低,所生成的3D人脸模型鲁棒性好、重构准确,实现了较好的3D人脸重建性能。

关键词

3D人脸重建;三维形变模型;深度神经网络;损失函数

中图分类号 TP391

文献标志码 A

收稿日期 2021-09-25

资助项目 河南省科技厅科技计划(162102310240);河南省高等学校重点科研项目(16A460013)

作者简介

杜召彬,男,硕士,讲师,主要研究领域为数字图像处理。duzhaobin19750317@163.com

¹ 郑州职业技术学院 软件工程系,郑州,450121

² 郑州轻工业大学 软件学院,郑州,450103

0 引言

长期以来,3D人脸重建技术一直是计算机视觉和光学成像等领域研究热点,在人脸识别^[1-2]、视频编辑^[3]和虚拟化妆^[4]等方向有着广泛的应用前景^[5-8]。

目前,自动化程度最高、应用最广泛的3D重建方法是基于统计模型的三维人脸形变模型(3D Morphable Model, 3DMM)^[9],近年有许多学者致力于这方面研究^[10-11]。例如:文献[12]针对非正面人脸转化为正面人脸图像通常会出现转换前后语义不一致的问题,即局部区域扭曲现象,提出了一种基于三维形变模型(3DMM)的人脸姿态矫正方法,从而通过该方法可以生成自然的人脸正面图像;文献[13]首先对2D人脸图像标准化,使用光照信息和对称纹理重建人脸自遮挡区域的纹理,然后依据2D-3D点对应关系从标准化的2D人脸图像获取相应的3D人脸纹理,结合人脸形状和纹理信息,最终得到3D人脸重建结果;文献[14]提出一种基于multi-level损失函数的弱监督学习模型,结合传统三维人脸形变模型(3DMM),直接从大量无标签的二维人脸图像中学习3D人脸特征信息,从而提出了基于单张二维人脸图像的3D人脸重建算法。

针对传统3DMM方式存在过约束,以及不能提供足够的灵活性来表示高频变形的问题,本文将非线性3DMM与神经网络相结合,提出了一种基于深度3DMM的3D人脸重建算法。

1 基于深度3DMM的3D人脸重建

1.1 线性三维形变模型

三维形变模型(3DMM)是一种使用面部几何形状和反射率来表示人脸的参数化模型。本文借鉴文献[15-16]的方法,利用主成分分析法(Principal Component Analysis, PCA)来描述3D人脸空间:

$$S = \bar{S} + G\alpha, \quad (1)$$

式中: $S \in \mathbf{R}^{3Q}$ 是拥有 Q 个顶点的3D人脸形状; $\bar{S} \in \mathbf{R}^{3Q}$ 是平均人脸形状; G 是3D人脸形状PCA主成分^[14]; $\alpha \in \mathbf{R}^L$ 是对应于3D人脸形状PCA主成分的形状参数。

将脸部的反射率 $A \in \mathbf{R}^{3Q}$ 定义在平均人脸形状 \bar{S} 内,其中 \bar{S} 描述了顶点 Q 的RGB颜色。可将脸部的反射率表示为脸部反射率PCA主成分的线性组合:

$$A = \bar{A} + R\beta, \quad (2)$$

式中: \bar{A} 是脸部的平均反射率; R 是脸部反射率 PCA 主成分; $\beta \in \mathbf{R}^{l_A}$ 是对应于脸部反射率 PCA 主成分的反射率参数.

于是,可以运用 3DMM 来合成新的面部视图.首先,利用弱透视投影模型进行图像处理:

$$V = RS, \quad (3)$$

$$g(S, m) = V^{2D} = fPV + t_{2d} = M(m) \begin{bmatrix} S \\ 1 \end{bmatrix}, \quad (4)$$

式中: R 是由 3D 旋转角所构成的旋转矩阵; $g(S, m)$ 是 3D 旋转顶点 V 向二维位置 V^{2D} 的投影函数; $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ 是投影矩阵; $M(\cdot)$ 是拥有 6 个自由度、大小为 24 的投影矩阵函数; m 是矩阵 6 自由度的向量参数; f 是比例因子; t_{2d} 是平移向量.

1.2 双通路神经网络

基于全局的 3DMM 模型通常对噪声和不匹配具有较强鲁棒性,但却存在过约束的现象,不能提供足够的灵活性来表示高频变形.为了有效改进这方面的不足,本文构建了双通路神经网络来实现对几何形状和反射率的自动解码.

发挥局部路径与全局路径相结合及其在图像合成中的成功经验^[17],应用于 3D 人脸重建.解码器的总体结构如图 1 所示,其中一个由分步卷积层和 5 个上采样步骤组成的全局路径用于全局推断,另一个由若干子网络构成的局部路径用于生成人脸不同部位的细节,包括眼睛、鼻子和嘴等.局部路径的引入,在以下两个方面具有明显优势:一是由于上采样步骤较少,局部路径能够更好地表达图像信息的高频细节;二是局部路径中每个子网络都可以学习部

分特定的过滤器,这比全域路径的计算效率更高.

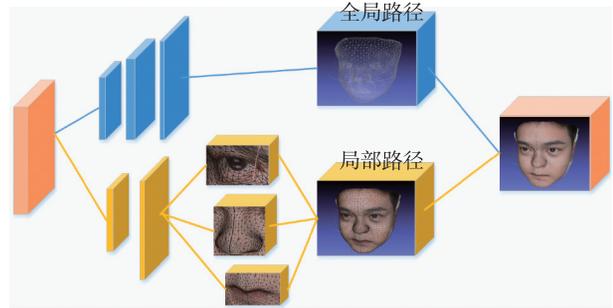


图 1 双通路神经网络示意

Fig. 1 Schematic diagram of dual-path neural network

如图 1 所示,本文将 4 个局部路径综合输出到单一的特征张量中,然后通过连续的卷积层来整合全局路径与局部路径信息,并生成最终的反射率/形状.此外,不同于其他 3D 人脸重建方式^[18-20],本文算法不设固定的关键信息点,3DMM 将通过局部路径与全局路径相结合的方式生成人脸反射率和 3D 人脸空间形状.

1.3 非线性三维形变模型

在线性 3DMM 中,对每个形状、反射率分量的分解均可以看成是系数和 PCA 主成分之间的矩阵乘法.但从神经网络的角度来看,可将其看作是一个没有激活函数的全连通层的神经网络^[21-22].为了增加模型的表达能力,本文将浅层网络扩展到深层架构,设计了一种将 3DMM 模型与深度神经网络相结合的非线性三维形变模型.

如图 2 所示,本文使用两个深度神经网络分别将形状、反射率参数解码为 3D 人脸形状和反射率.为了使框架端到端可训练,这些参数由编码器网络

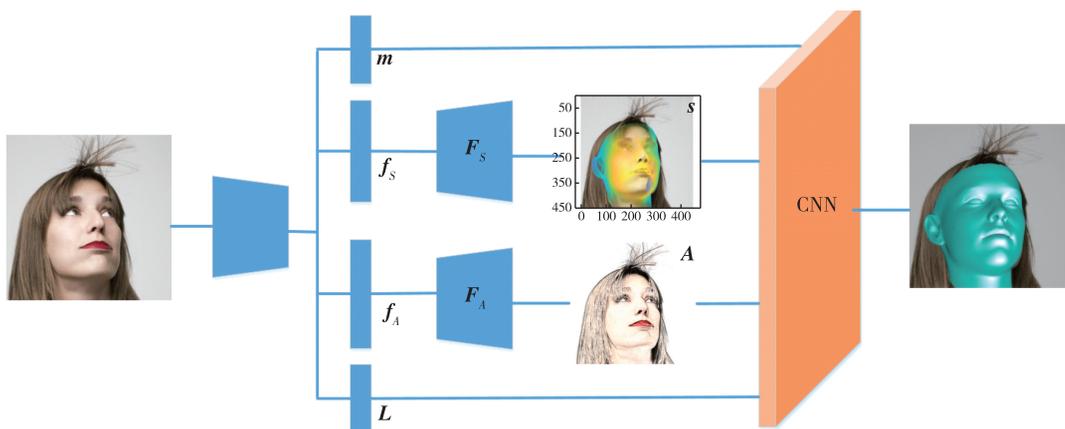


图 2 基于深度 3DMM 的 3D 人脸重建架构示意

Fig. 2 Schematic of 3D face reconstruction based on deep 3DMM

估计.在基于物理的渲染层的辅助下,3个深度网络联合起来,最终自动实现单张3D人脸重建的目标.

对于给定一组 n 张二维人脸图像 $\{I_i\}_{i=1}^K$, 通过该框架学习一种编码器 $E: I \rightarrow m, L, f_s, f_A$ 用来估计投影 m 、照明参数 L 、形状参数 $f_s \in \mathbf{R}^l$ 以及反射率参数 $f_A \in \mathbf{R}^l$; 学习一种3D形状解码器 $F_S: f_s \rightarrow S$ 将形状参数解码为3D形状 $S \in \mathbf{R}^{3Q}$; 学习一种反射率解码器 $F_A: f_A \rightarrow A$ 将反射率参数解码为真实的反射率 $A \in \mathbf{R}^{3Q}$, 用 m, L, S 和 A 渲染图像以更好地近似原始图像.数学上,其目标函数为

$$\arg \min_{E, D_S, D_A} \sum_I K(\hat{I}, I), \quad \hat{I} = R(E_M(I), E_L(I), F_S(E_S(I)), F_A(E_A(I))), \quad (5)$$

式中: $K(\cdot)$ 是3D重建的损失函数,将其定义为

$$K = K_{\text{rec}} + \lambda_L K_{\text{lan}} + \lambda_{\text{reg}} K_{\text{reg}}, \quad (6)$$

式中: K_{rec} 是三维形变重建损失,以描述图像估计 \hat{I} 与图像输入 I 之间的相似度; K_{lan} 是三维形变标注损失,以强化图像重建中的几何约束; K_{reg} 是三维形变正则化损失,以有助于寻求合理的解决方案; $\lambda_L, \lambda_{\text{reg}}$ 是权重系数.

1) 三维形变重建损失

对于重建损失有很多的方式,其中最直接的方式是用 l_1 或 l_2 损耗在像素空间中比较图像.为了更好地处理异常值,本文采用了鲁棒性较强的 $l_{2,1}$, 其中RGB颜色空间的距离基于 l_2 , 求和则基于 l_1 范数以加强稀疏性^[23-24]:

$$K_{\text{rec}}^i = \frac{1}{|\nu|} \sum_{q \in \nu} \|\hat{I}(q) - I(q)\|_2, \quad (7)$$

式中: ν 是二维人脸图像网格覆盖的像素集合.

图像 $\hat{I}(q)$ 和 $I(q)$ 之间的亲密度也可以在特征空间中通过感知损失得到加强:

$$K_{\text{rec}}^f = \frac{1}{|C|} \sum_{j \in C} \frac{1}{W_j H_j C_j} \|\varphi_j(\hat{I}) - \varphi_j(I)\|_2^2, \quad (8)$$

式中: C 是神经网络层 φ 上的一个子集; $\varphi_j(I)$ 是处理 I 时得到的第 j 层维数为 $W_j \times H_j \times C_j$ 的激活函数.

于是,最终的三维形变重建损失是由图像和特征重建损失组成的加权平均值:

$$K_{\text{rec}}(\hat{I}, I) = K_{\text{rec}}^i(\hat{I}, I) + \lambda_f K_{\text{rec}}^f(\hat{I}, I), \quad (9)$$

式中: λ_f 是权重系数.

2) 三维形变标注损失

为了有助于实现更好的三维形变,本文构建了一种标注损失,测量估计和二维图像与三维形变之间标注的欧氏距离:

$$K_{\text{lan}} = \left\| M(m) \begin{bmatrix} S(:, d) \\ 1 \end{bmatrix} - U \right\|_2^2, \quad (10)$$

式中: $U \in \mathbf{R}^{2 \times 68}$ 是事前标注的二维图像标记信息; d 是存储与标记的二维图像信息对应的68个3D顶点索引向量.由于三维形变标注损失只影响与关键点相关的顶点^[25-26], 本文将三维形变标注损失仅应用于编码器 E , 以防止学习不需要的信息.

3) 三维形变正则化损失

为了确保3D重建的合理性,并能真实地反映形态、反射率和光照等因素,本文添加了一些正则化损失项:

$$K_{\text{reg}} = \lambda_{\text{sym}} K_{\text{sym}} + \lambda_{\text{con}} K_{\text{con}} + \lambda_{\text{smo}} K_{\text{smo}}, \quad (11)$$

式中: K_{sym} 是反射率对称约束; K_{con} 是反射率保持约束; K_{smo} 是形状平滑约束; $\lambda_{\text{sym}}, \lambda_{\text{con}}, \lambda_{\text{smo}}$ 是权重系数.

① 反射率对称约束

借鉴Tran等^[25]利用顶点的UV空间关系将 S 和 A 分别表示为 A^{UV} 和 S^{UV} . 考虑人脸的对称性,在此设定了反射率对称约束:

$$K_{\text{sym}} = \|A^{UV} - f_{\text{lip}}(A^{UV})\|_1, \quad (12)$$

式中: $f_{\text{lip}}(\cdot)$ 是水平图像翻转函数.

② 反射率保持约束

使用反射率对称约束可以帮助修正人脸图像中的全局阴影.然而,面部细节如酒窝等,也将嵌入反射率通道中.为了进一步消除反射率通道的阴影,遵循假定反射率为分段常数的视网膜皮层理论^[27-28], 本文在其梯度的两个方向上施加稀疏性约束:

$$K_{\text{con}} = \sum_{\nu_j^{UV} \in N_i} \omega(\nu_i^{UV}, \nu_j^{UV}) \|A^{UV}(\nu_i^{UV}) - A(\nu_j^{UV})\|_2^p, \quad (13)$$

式中: N_i 是像素 ν_i^{UV} 的一组4像素邻域; $\omega(\nu_i^{UV}, \nu_j^{UV}) = \exp(-\alpha \|c(\nu_i^{UV}) - c(\nu_j^{UV})\|)$ 是权重系数.

③ 形状平滑约束

对于三维形变,本文通过对所有顶点位置添加拉普拉斯正则化来实现平滑约束:

$$K_{\text{smo}} = \sum_{\nu_i^{UV} \in S^{UV}} \left\| S^{UV}(\nu_i^{UV}) - \frac{1}{|N_i|} \sum_{\nu_j^{UV} \in N_i} S^{UV}(\nu_j^{UV}) \right\|_2, \quad (14)$$

2 实验分析

本文算法训练中选择了两个公开可用的人脸图像数据集:AFLW数据集^[29]是一个二维人脸图像数据集,每个图像有多达21处标记,每个标记有一个可见性标签.文献[29]选择了一个偏航角均衡分布

的 AFLW 子集,包括 3 901 张用于训练的图像和 1 299 张用于测试的图像,本文也使用相同的子集对算法进行训练和测试. AFW 数据集^[30]包含 205 张图像,每张人脸图像有 6 处标记,每个标记有一个可见性标签.对于每张人脸图像,提供检测到的包围框.由于 AFW 数据集图像数量相对较少,本文仅使用该数据集进行测试.

本文算法仿真使用 CPU: Intel Core (TM) i5-8250U; RAM: 16 GB; 显卡: Geforce GTX 独立 GPU 的硬件设备,基于 Pytorch 深度学习框架实现,采用 ResNet-50 网络模型,将最后一层全连接层设为 314,神经网络采用 Adam 方法优化参数,学习率设为 0.001, batch size 设为 16, 10 个 epoch 共训练 101 360 次迭代, 3D 重建的损失函数权重 λ_L, λ_{reg} 分别设为 2.6×10^{-2} 和 4×10^{-3} , 三维形变正则化损失项中的 $\lambda_{sym}, \lambda_{con}, \lambda_{smo}$ 分别设为 0.25、0.45 和 0.3.

以二维人脸图像作为输入,运用本文算法进行 3D 人脸重建,重建效果如图 3 所示.由图 3 可知,本文算法能够准确获取二维人脸五官信息,3D 重建效果图能够还原二维人脸特征,达到了预期效果.

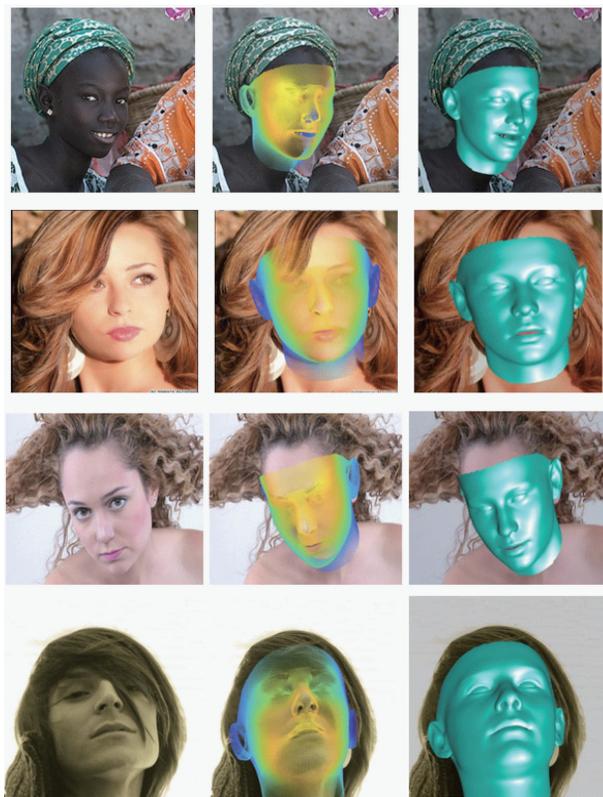


图 3 本文算法 3D 人脸重建效果
Fig. 3 3D face reconstruction renderings based on the proposed algorithm

为了公平比较本文算法相对于传统 3DMM 算法的改进优势,针对三组不同视角的二维人脸图像,分别使用两种算法进行 3D 人脸重建,3D 模型如图 4 所示.使用传统 3DMM 算法^[25,31]进行重建,由于单张二维图像数据类型单一、训练数量以及线性基数量较少,得到的 3D 人脸纹理来自 3D 模型的平均人脸纹理^[32],加之传统 3DMM 算法存在的过约束现象,不能提供足够的灵活性来表示高频变形,使得 3D 重建图像的整体平滑度不够,纹理失真明显,缺少目标人脸的真实细节信息.使用本文算法进行重建,通过局部路径与全局路径相结合的方式,能够改善传统 3DMM 算法高频变形表达能力的不足,获取面部区域的纹理细节,能更好地表现人脸眼睛、鼻子、嘴巴等区域的细节信息,最大程度地还原二维人脸细节,同时更好地保留人脸图像的光照信息.



图 4 算法改进仿真分析

Fig. 4 Simulation analysis of algorithm improvement

为了检验本文算法,通过对比仿真 3DDFA^[33]和 DeFA 算法^[34],来分析本文算法在 3D 人脸重建方面的比较优势.3D 重建算法间的性能比较,通常采用归一化平均误差(Normalized Mean Error, NME)作为指标,归一化平均误差是指根据人脸包围盒的大小进行归一化,如式(15)^[35]所示:

$$I_{NME} = \frac{1}{T} \sum_{k=1}^N \frac{\| \mathbf{u}_k - \mathbf{y}_k \|_2}{d}, \quad (15)$$

式中: N 为测试图像标记点个数; d 为人脸真实的包围盒长(l) 宽(h) 之积的平方根,计算方式为 $d =$

$\sqrt{l \cdot h}$; u_k 和 y_k 分别为 3D 重建图像点坐标和测试集上的标签.

对比实验中,分别将 3 种算法的 3D 人脸重建模型与二维样本图像的 PCA 主成分进行密集对齐,以点对面距离归一化平均误差作为横坐标输入,以处理不同二维图像的累计数量作为纵坐标输入,拟合出算法累积归一化平均误差变化曲线,对比仿真结果如图 5 所示.图中,本文算法对应归一化平均误差曲线的均值为 2.03%,3DDFA 算法对应归一化平均误差曲线的均值为 2.46%,DeFA 算法对应归一化平均误差曲线的均值为 4.36%.可以看出,本文算法的 3D 重建归一化平均误差(NME)比 3DDFA 算法^[33]的 3D 重建结果降低 17.48%,比 DeFA 算法^[34]的 3D 重建结果降低 53.44%,应用效果有明显的改善.

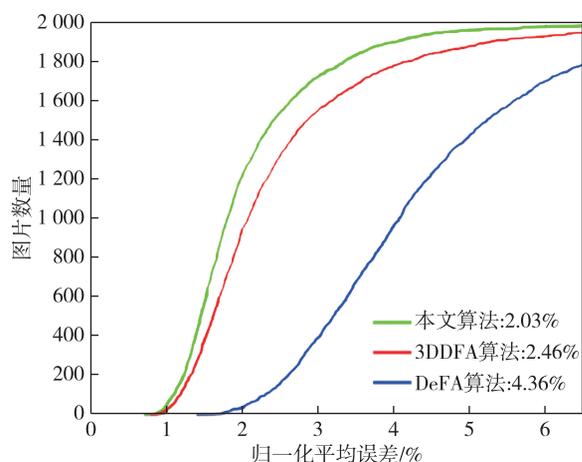


图 5 算法比对归一化平均误差

Fig. 5 Comparison of normalized average errors

图 6 为本文算法与 3DDFA 算法、DeFA 算法的可视化结果.可以看出,由于 3DDFA 和 DeFA 算法都在 3D 重建中引入了一些伪迹^[25,32],使得人脸重建看起来不自然,相比较本文算法的 3D 重建形状更为平滑、自然.

3 结论

针对传统 3DMM 迭代拟合过程中存在过约束、不能提供足够的灵活性来表示高频变形等问题,本文在分析线性三维形变模型的基础上,构建双通路神经网络实现了几何形状和反射率的解码,将浅网络扩展到深层架构,设计了一种将 3DMM 模型与神经网络相结合的非线性三维形变模型,从而提出了基于深度 3DMM 的 3D 人脸重建算法.算法采取局部路径与全局路径相结合的方式,使得每个子网络都

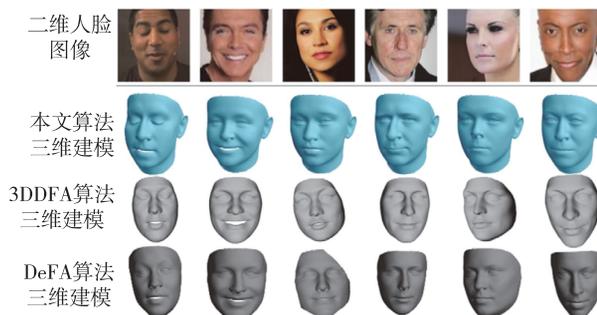


图 6 3D 人脸重建效果对比

Fig. 6 Comparison of 3D face reconstruction performances

可以学习特定细节信息,在比全域路径获得更高计算效率的同时,能够更好地表示人脸面部的高频细节信息.从算法比对和仿真实验结果可知,本文算法误差更低,进一步改善了传统 3DMM 算法在高频变形表达能力方面的不足,通过模型拟合可以实现高保真的 3D 人脸重建.

参考文献

References

[1] 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. DOI: 10. 1109/AFGR.2008. 4813376

[2] Masi I, Trần A T, Hassner T, et al. Do we really need to collect millions of faces for effective face recognition? [C] // European Conference on Computer Vision. October 11 - 14, 2016, Amsterdam, Netherlands. DOI: 10. 1007/978-3-319-46454-1_35

[3] Garrido P, Valgaerts L, Sarmadi H, et al. VDub: modifying face video of actors for plausible visual alignment to a dubbed audio track [J]. Computer Graphics Forum, 2015, 34(2): 193-204

[4] Li C, Zhou K, Lin S. Simulating makeup through physics-based manipulation of intrinsic image layers [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 7 - 12, 2015, Boston, MA, USA. IEEE, 2015: 4621-4629

[5] 张彦雯, 胡凯, 王鹏盛. 三维重建算法研究综述 [J]. 南京信息工程大学学报(自然科学版), 2020, 12(5): 591-602
ZHANG Yanwen, HU Kai, WANG Pengsheng. Review of 3D reconstruction algorithms [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2020, 12(5): 591-602

[6] 傅泽华, 龚勋, 李天瑞. 二维及三维多模人脸数据库构建 [J]. 数据采集与处理, 2017, 32(3): 595-603
FU Zehua, GONG Xun, LI Tianrui. 2D & 3D multi-modal face database [J]. Journal of Data Acquisition and Processing, 2017, 32(3): 595-603

- [7] 黄菲,高飞,朱静洁,等.基于生成对抗网络的异质人脸图像合成:进展与挑战[J].南京信息工程大学学报(自然科学版),2019,11(6):660-681
HUANG Fei,GAO Fei,ZHU Jingjie, et al.Heterogeneous face synthesis via generative adversarial networks: progresses and challenges[J].Journal of Nanjing University of Information Science & Technology (Natural Science Edition),2019,11(6):660-681
- [8] 刘庆华,李智.一种基于HOG与LBP双特征融合模型的人脸年龄估计方法[J].江苏科技大学学报(自然科学版),2021,35(3):50-55
LIU Qinghua,LI Zhi.Faceage estimation method based on HOG and LBP dual feature fusion model[J].Journal of Jiangsu University of Science and Technology (Natural Science Edition),2021,35(3):50-55
- [9] 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,1999.DOI:10.1145/311535.311556
- [10] 杨玮玥,张倩,王斌,等.3D建模的多姿态人脸合成方法[J].中国科技论文,2018,13(14):1573-1577,1602
YANG Weiyue,ZHANG Qian,WANG Bin, et al.Multi-pose face synthesis method by 3D modeling[J].China Sciencepaper,2018,13(14):1573-1577,1602
- [11] 敖建锋,苏泽锴,刘传立,等.基于点云深度学习的3D目标检测[J].激光与红外,2020,50(10):1276-1282
AO Jianfeng,SU Zekai,LIU Chuanli, et al.3D object detection based on point cloud deep learning[J].Laser & Infrared,2020,50(10):1276-1282
- [12] 张海燕,文翔,臧怀娟,等.基于3DMM的人脸姿态校正算法[J].电子测量与仪器学报,2019,33(7):99-105
ZHANG Haiyan,WEN Xiang,ZANG Huaijuan, et al.Face pose correction algorithm based on 3DMM[J].Journal of Electronic Measurement and Instrumentation,2019,33(7):99-105
- [13] 阳瑜,吴小俊.基于人脸标准化的纹理和光照保持3D人脸重构[J].模式识别与人工智能,2019,32(6):557-568
YANG Yu,WU Xiaojun.Texture and illumination preserving 3D face reconstruction based on face normalization [J]. Pattern Recognition and Artificial Intelligence,2019,32(6):557-568
- [14] 吴越,董兰芳.基于弱监督学习的三维人脸形状与纹理重建[J].计算机系统应用,2020,29(11):183-189
WU Yue,DONG Lanfang.3D face shape and texture reconstruction based on weakly supervised learning [J]. Computer Systems & Applications,2020,29(11):183-189
- [15] 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,1999
- [16] Liu X M.Video-based face model fitting using adaptive active appearance model [J]. Image and Vision Computing,2010,28(7):1162-1172
- [17] Huang R,Zhang S, Li T Y, et al.Beyond face rotation: global and local perception GAN for photorealistic and identity preserving frontal view synthesis [C] // 2017 IEEE International Conference on Computer Vision (ICCV).October 22-29,2017,Venice,Italy.IEEE,2017:2458-2467
- [18] Karras T,Aila T,Laine S, et al.Progressive growing of GANs for improved quality, stability, and variation [J]. arXiv e-print,2017,arXiv:1710.10196
- [19] Tran L, Yin X, Liu X M. Disentangled representation learning GAN for pose-invariant face recognition [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).July 21-26,2017,Honolulu,HI,USA.IEEE,2017:1283-1292
- [20] Tran L, Yin X, Liu X M. Representation learning by rotating your faces [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2019,41(12):3007-3021
- [21] Tewari A,Zollhöfer M,Garrido P, et al.Self-supervised multi-level face model learning for monocular reconstruction at over 250 Hz [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition.June 18-23,2018,Salt Lake City,UT,USA.IEEE,2018:2549-2559
- [22] 刘冰雁,叶雄兵,周赤非,等.基于改进DQN的复合模式在轨服务资源分配[J].航空学报,2020,41(5):256-264
LIU Bingyan, YE Xiongbing, ZHOU Chifei, et al. Allocation of composite mode on-orbit service resource based on improved DQN [J]. Acta Aeronautica et Astronautica Sinica,2020,41(5):256-264
- [23] Thies J,Zollhofer M,Stamminger M, et al.Face2Face: real-time face capture and reenactment of RGB videos [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).June 27-30,2016,Las Vegas,NV.IEEE,2016.DOI:10.1109/CVPR.2016.262
- [24] Thies J,Zollhöfer M,Stamminger M, et al.FaceVR: real-time facial reenactment and eye gaze control in virtual reality [J]. arXiv e-print,2016,arXiv:1610.03151
- [25] Tran L,Liu X M.Nonlinear 3d face morphable model [J]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018:7346-7355
- [26] Tran L,Liu X M.On learning 3D face morphable model from in-the-wild images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2019,43(1):157-171
- [27] Meka A,Zollhöfer M,Richardt C, et al.Live intrinsic video [J]. ACM Transactions on Graphics (TOG),2016,35(4):1-14
- [28] Shu Z X,Yumer E, Hadap S, et al. Neural face editing with intrinsic image disentangling [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).July 21-26,2017,Honolulu,HI.IEEE,2017
- [29] Köstinger M,Wohllhart P,Roth P M, et al.Annotated facial landmarks in the wild:a large-scale,real-world database for facial landmark localization [C] // 2011 IEEE International Conference on Computer Vision Workshops. November 6-13,2011,Barcelona,Spain.IEEE,2011:2144-2151
- [30] 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. IEEE, 2012: 2879-2886
- [31] Tran A T, Hassner T, Masi I, et al. Extreme 3D face reconstruction: looking past occlusions [J]. arXiv e-print, 2017, arXiv: 1712. 05083
- [32] 阳瑜, 吴小俊. 基于人脸标准化的纹理和光照保持 3D 人脸重构[J]. 模式识别与人工智能, 2019, 32(6): 557-568
YANG Yu, WU Xiaojun. Texture and illumination preserving 3D face reconstruction based on face normalization [J]. Pattern Recognition and Artificial Intelligence, 2019, 32(6): 557-568
- [33] Zhu X Y, Lei Z, Liu X M, et al. Face alignment across large poses: a 3D solution [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27–30, 2016, Las Vegas, NV, USA. IEEE, 2016: 146-155
- [34] Liu Y J, Jourabloo A, Ren W, et al. Dense face alignment [C] // 2017 IEEE International Conference on Computer Vision. October 22–29, 2017, Venice, Italy. IEEE, 2017: 1619-1628
- [35] Zhu X Y, Lei Z, Liu X M, et al. Face alignment across large poses: a 3D solution [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27–30, 2016, Las Vegas, NV, USA. IEEE, 2016: 146-155

3D face reconstruction from single image based on deep 3D morphable model

DU Zhaobin¹ CUI Xiao²

¹ Software Engineering Department, Zhengzhou Technical College, Zhengzhou 450121

² College of Software Engineering, Zhengzhou University of Light Industry, Zhengzhou 450103

Abstract 3D morphable model (3DMM) has been widely used in 3D modeling, image synthesis and related fields. However, it is perplexed by over-constraint due to the influence from size, types, and principal components of training data, thus cannot provide enough flexibility to represent high-frequency deformation. Here, we embed the 3DMM into deep neural network to improve its representation ability in 3D face reconstruction. A dual-path neural network is constructed and improved in efficiency of network learning, which achieves balance between global path and local path. Then the nonlinear 3DMM is improved in both learning objectives and network structure, so as to capture more details than linear or previous nonlinear models. The comparison and simulation experiments show that the proposed algorithm has lower normalized average error in 3D face reconstruction, and the generated 3D face model has good robustness and accurate details.

Key words 3D face reconstruction; 3D morphable model (3DMM); deep neural network; loss function