

基于 3D-2D 映射的大姿态人脸识别

李晓峰, 游志胜

(四川大学计算机学院, 成都 610065)

摘要: 为了解决真实场景下大姿态人脸识别准确率低的问题, 本文从数据增广的角度, 围绕三维人脸信息数据进行大姿态人脸识别关键技术的研究与探索, 提出了一个基于 3D-2D 映射的大姿态人脸识别算法框架. 区别于当前基于 3D 点云数据的人脸识别算法, 本文提出的方法主要利用 3D 人脸丰富姿态信息, 通过 3D-2D 映射的姿态数据扩充方法, 进而训练特定姿态人脸特征提取模型, 然后集成到统一大姿态人脸识别框架中. 本文提出方法的关键是利用注册 3D 人脸图像信息来辅助 2D 人脸多姿态识别, 很容易集成现有的 2D 人脸识别方法到提出的框架中. 实验表明, 本文提出的方法在无约束真实场景下能够有效提升大姿态人脸识别准确率, 同时不显著增加计算负担, 为当前基于 3D 信息解决 2D 人脸识别问题提供了新思路.

关键词: 人脸识别; 大姿态人脸识别; 三维对二维

中图分类号: TP391 **文献标识码:** A **DOI:** 10.19907/j.0490-6756.2022.042003

Large-pose face recognition based on the 3D-2D mapping

LI Xiao-Feng, YOU Zhi-Sheng

(College of Computer Science, Sichuan University, Chengdu 610065, China)

Abstract: To solve the problem of low accuracy of large pose face recognition in real scenes, a large pose face recognition algorithm framework is proposed based on the 3D-2D mapping focusing on the exploration of the key technologies of large pose face recognition from the perspective of data augmentation. Different from the current face recognition algorithm based on 3D point cloud data, the proposed method mainly uses 3D face data to enrich the posture information via the data expansion method by 3D-2D mapping. The specific posture face feature extraction models are trained and integrated into the unified large pose face recognition framework. The key of the proposed method is to use registered 3D face information to assist 2D face multi-pose recognition. It is easy to integrate existing 2D face recognition methods into the proposed framework. Experiments show that the method proposed in this paper can effectively improve the accuracy of large pose face recognition in unconstrained real scenes without increasing the significant computational load and provides a new idea for the current 3D information to solve the 2D face recognition problem.

Keywords: Face recognition; Large-pose face recognition; 3D to 2D

1 引言

当前基于 2D 人脸识别方法主要能克服小姿

态偏转情况. 但头部的刚性旋转角度过大会伴随自遮挡效应, 使得人脸大量细节特征信息丢失. 姿态的变化也会导致面部纹理位置发生非线性变化, 使

收稿日期: 2021-08-23

基金项目: 四川省应用基础研究(2021YJ0079); 国家自然科学基金(U20A20161)

作者简介: 李晓峰(1973-), 男, 副教授, 研究方向为计算机视觉, 模式识别. E-mail: lixiaofengflc@qq.com

得部分语义信息丢失. 此外, 姿态变化通常还会与其他干扰因素同时发生, 如非均匀光照等, 从而造成识别率极大降低. 克服大姿态人脸识别面临的挑战, 已成为无约束场景下人脸识别技术研究的主要方向.

为了解决大姿态条件下人脸识别率低的问题, 相关研究学者已经提出了一系列解决方法, 依据数据类别主要分为三大类: (1) 基于二维可见光人脸的识别方法^[1-6]; (2) 基于三维人脸的识别方法^[7-9]; (3) 基于三维人脸对二维人脸的识别方法^[10-12]. 二维可见光人脸图像对姿态、光照^[13, 14]等因素较为敏感, 当人脸姿态超过一定范围、图像高曝光以及低光照场景, 识别准确性会大幅下降. 当前基于 2D 人脸识别的大姿态解决方案主要分为两种, 一种是通过姿态人脸生成正面人脸, 之后对正面人脸进行特征提取和识别; 另一种则是基于不变表征的学习方法. Zhang 等^[15]提出的解耦表征学习方法通过对比两张图片, 剔除姿态、光照等与识别无关的因素, 提取出具有姿态鲁棒性的人脸特征. 赵澜涛等^[16]提出基于双路卷积神经网络的算法, 通过将人脸投影到高维特征空间达到分离姿态信息的效果, 从而对姿态人脸进行识别. Luan 等^[17]引入了生成对抗网络到解耦表征学习框架中, 通过生成器生成姿态人脸, 通过判别器分类姿态人脸, 从而提取出具有姿态鲁棒性的人脸特征.

鉴于三维人脸模型对姿态、光照等因素相对不敏感, 能够更加客观地反应人脸的几何结构信息, 因此具备更强的身份判别能力, 在识别准确率和鲁棒性等方面都优于二维人脸识别, 使得三维人脸识别相对二维人脸识别占据很大优势. 但仅靠三维人脸识别依然存在采集环境设备成本高、要求用户配合程度高、数据处理算法复杂等多方面问题, 相对于二维人脸识别面临更多的应用限制. 相比而言, 本文提出的方法仅仅利用 3D 人脸丰富的姿态信息, 基于现有的深度人脸特征提取方法来提升大姿态人脸识别准确率, 这可以很好地利用现有二维人脸识别的基础, 降低三维人脸识别应用的门槛, 还可以充分利用三维人脸的全脸信息.

针对上述问题, 本文提出一种基于 3D-2D 映射的大姿态人脸识别算法框架. 该框架试图挖掘三维人脸的丰富姿态信息来扩充训练数据, 进而构建特定姿态特征提取模型. 因此提出方法主要优势如下: (1) 利用三维人脸信息来辅助 2D 人脸模型训练; (2) 简化模型设计的同时, 不增加额外的训练负担, 避免直接三维人脸特征提取的复杂工作.

2 本文方法

本文提出的基于 3D-2D 映射的大姿态人脸识别算法框架如图 1 所示.

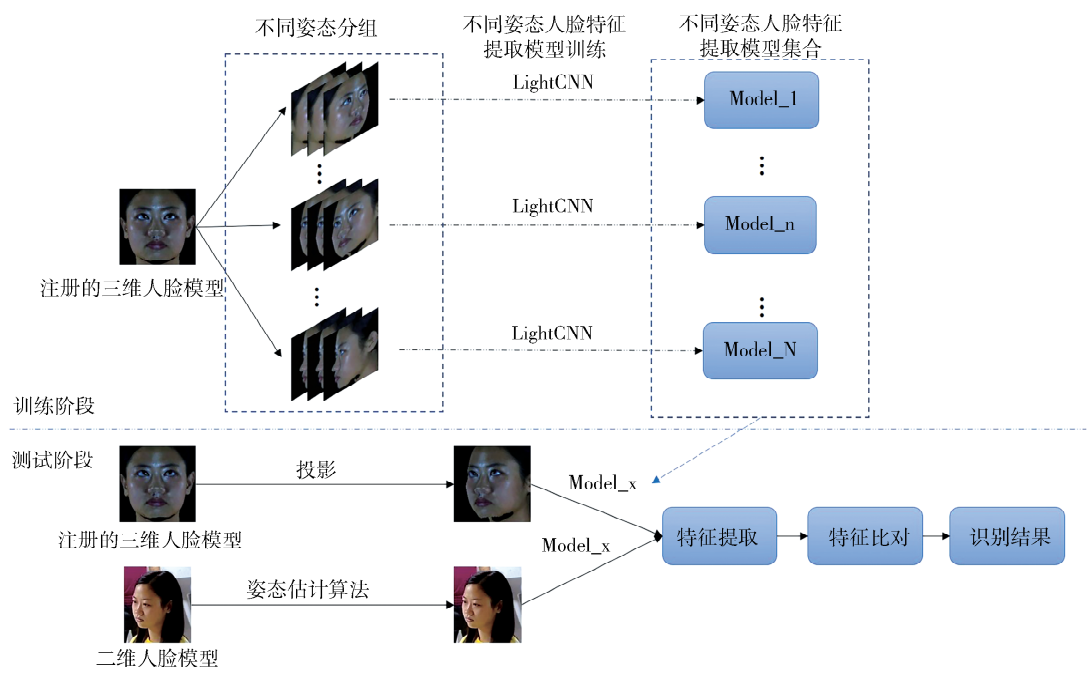


图 1 基于 3D-2D 映射的大姿态人脸识别算法流程示意图
Fig. 1 The flowchart of large pose face recognition algorithm based on 3D-2D mapping

图 1 中的框架主要分为两个阶段: 训练阶段和测试阶段(模型推理阶段). 在训练阶段, 该方法首先构建多姿态人脸训练数据库. 具体步骤是首先获取三维人脸数据, 然后将三维人脸数据进行左右偏转角、俯仰角等不同角度投影, 得到不同姿态的二维人脸数据集. 为建立鲁棒性的姿态特征提取模型, 需对获取的丰富姿态进行姿态分类和分组, 针对不同姿态数据集分别训练得到相应姿态人脸特征提取模型, 最终得到多姿态人脸特征提取模型集合. 在测试阶段, 输入二维人脸图像通过姿态估计算法得到姿态信息, 利用姿态信息在多姿态人脸特征提取模型集合中进行相应姿态模型选择, 用于将注册的三维人脸模型按照姿态信息进行投影的二维纹理图像和输入图像进行人脸特征提取, 进而做特征比对和识别.

2.1 3D-2D 人脸姿态数据生成

本文首先将三维人脸模型按照角度投影到二维图像空间, 之后基于二维多姿态人脸纹理识别模型进行人脸识别. 因此, 3D-2D 人脸姿态数据生成是本文提出框架的重要组成部分. 考虑到 RGB 图像更符合视觉习惯且三维人脸模型投影到二维空间存在信息丢失的问题, 本文采用多角度投影方法, 将三维人脸模型按角度依次投影为多张不同视角的 RGB 图像, 分别是俯仰角(*pitch*), 左右偏转角(*yaw*), 左右偏转角范围为 $[-90^\circ, +90^\circ]$ 、俯仰角范围为 $[-60^\circ, +60^\circ]$.

使用三维人脸点云所在柱面中心轴作为垂直 Y 轴, 取鼻尖点与 Y 轴垂直相交所在直线为 Z 轴, 将所有三维人脸统一坐标系, 如图 2a 所示, 投影方向如图 2b 所示进行定义. 则投影方向 θ 等同于将三维模型在三维空间变形之后向 $z + \cot\left(\frac{\pi}{2} - \theta\right) \cdot x = 0$ 所表示的平面投影.

为了便于计算三维人脸模型上点到投影面的距离, 先将模型沿坐标轴平移. 假设模型上有 N 个点, 第 i 个点的形状信息为 $S_0^i(x_0^i, y_0^i, z_0^i)$, 对应的纹理信息为 $T^i(R^i, G^i, B^i)$. 将三维人脸模型沿着 X 轴和 Z 轴方向平移的距离为 h_x, h_z , 则

$$x_1^i = x_0^i + h_x, z_1^i = z_0^i + h_z \quad (1)$$

其中

$$\begin{cases} h_x > \max_{i=1, \dots, N} x_0^i - \min_{i=1, \dots, N} x_0^i \\ h_z > \max_{i=1, \dots, N} z_0^i - \min_{i=1, \dots, N} z_0^i \end{cases} (\theta > 0);$$

$$\begin{cases} h_x > \min_{i=1, \dots, N} x_0^i - \max_{i=1, \dots, N} x_0^i \\ h_z > \min_{i=1, \dots, N} z_0^i - \max_{i=1, \dots, N} z_0^i \end{cases} (\theta \leq 0) \quad (2)$$

假设三维人脸模型的投影方向为 θ , 则第 i 个三维坐标点 $S_1^i(x_1^i, y_1^i, z_1^i)$ 在 $z + \cot\left(\frac{\pi}{2} - \theta\right) \cdot x = 0$ 上的投影 $S_2^i(x_2^i, y_2^i, z_2^i)$ 可计算为

$$\begin{cases} x_2^i = \frac{\tan\left(\frac{\pi}{2} - \theta\right) \cdot [\tan\left(\frac{\pi}{2} - \theta\right) \cdot x_1^i - z_1^i]}{1 + [\tan\left(\frac{\pi}{2} - \theta\right)]^2} \\ y_2^i = y_1^i \\ z_2^i = \frac{z_1^i - \tan\left(\frac{\pi}{2} - \theta\right) \cdot x_1^i}{1 + [\tan\left(\frac{\pi}{2} - \theta\right)]^2} \end{cases} \quad (3)$$

将投影结果旋转到 YZ 面, 假设 $S_3^i(x_3^i, y_3^i, z_3^i)$ 是 $S_2^i(x_2^i, y_2^i, z_2^i)$ 旋转后结果, 则:

$$\begin{bmatrix} x_3^i \\ y_3^i \\ z_3^i \end{bmatrix} = \begin{bmatrix} \cos\left(\frac{\pi}{2} - \theta\right) & 0 & \sin\left(\frac{\pi}{2} - \theta\right) \\ 0 & 1 & 0 \\ -\sin\left(\frac{\pi}{2} - \theta\right) & 0 & \cos\left(\frac{\pi}{2} - \theta\right) \end{bmatrix} \begin{bmatrix} x_2^i \\ y_2^i \\ z_2^i \end{bmatrix} \quad (4)$$

假设投影结果的行为 *nrows*, 列为 *ncols*, 点 $S_3^i(x_3^i, y_3^i, z_3^i)$ 在平面上表示第 (a_i, h_i) 个像素点, 则投影图像为

$$\begin{cases} a_i = \frac{(\max_{i=1, \dots, N} y_0^i - y_0^i) \cdot nrows}{\max_{i=1, \dots, N} y_0^i - \min_{i=1, \dots, N} y_0^i} \\ h_i = \frac{(\max_{i=1, \dots, N} z_3^i - z_3^i) \cdot ncols}{\max_{i=1, \dots, N} z_3^i - \min_{i=1, \dots, N} z_3^i} \end{cases} \quad (5)$$

像素点 (a_i, h_i) 的纹理值即:

$$\begin{aligned} \text{img}(a_i, h_i, 1) &= R_i, \text{img}(a_i, h_i, 2) = G_i, \\ \text{img}(a_i, h_i, 3) &= B_i \end{aligned} \quad (6)$$

假设 $S_1^m(x_1^m, y_1^m, z_1^m)$ 和 $S_1^n(x_1^n, y_1^n, z_1^n)$ 投影所得坐标为 $S_2^m(x_2^m, y_2^m, z_2^m)$ 和 $S_2^n(x_2^n, y_2^n, z_2^n)$, 则如果式 (7) 成立, 根据式 (8) 可以判断该二维点的纹理信息 (R, G, B) .

$$x_2^m = x_2^n, y_2^m = y_2^n, z_2^m = z_2^n \quad (7)$$

则:

$$\begin{cases} R = R^m, G = G^m, B = B^m, \\ \text{当 } z_1^m + \tan \theta \cdot x_1^m \geq z_1^n + \tan \theta \cdot x_1^n, \\ R = R^n, G = G^n, B = B^n, \text{ 其他} \end{cases} \quad (8)$$

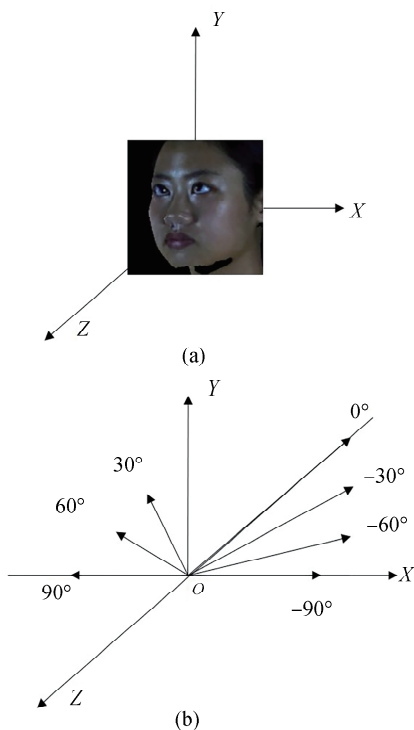


图 2 三维人脸坐标系
Fig. 2 3D facial coordinate system

基于 3D-2D 人脸姿态数据生成结果如图 3 所示。

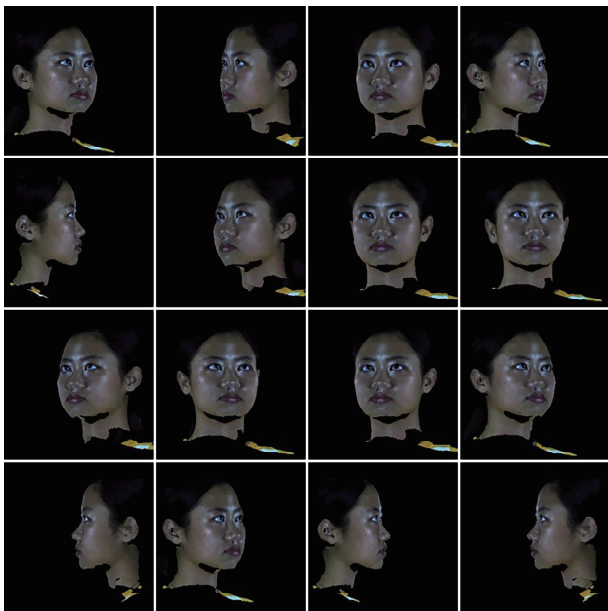


图 3 基于 3D-2D 人脸姿态数据生成结果示例
Fig. 3 Projection results based on 3D-2D mapping

2.2 基于特定姿态的模型训练

通过基于 3D-2D 映射的多姿态人脸数据生成算法得到不同姿态的数据集,考虑到实际场景下人脸姿态分布等信息,将偏转角划分为以下 7 个区

间: $[-90^\circ, -75^\circ]$, $[-75^\circ, -60^\circ]$, $[-60^\circ, -30^\circ]$, $[-30^\circ, +30^\circ]$, $[+30^\circ, +60^\circ]$, $[+60^\circ, +75^\circ]$, $[+75^\circ, +90^\circ]$. 俯仰角按照投影间隔为 20° 划分为 6 个区间. 因此,投影所得人脸数据共分 42 种姿态数据集.

本文采用轻量级模型 LightCNN-9^[18] 进行人脸特征提取,该模型通过拓展 maxout 激活函数,实现对噪声的过滤和有用信号的保留,从而获取更好的特征图 MFM(Max-Feature-Map). 如图 4 所示为本文采用的 MFM2/1 原理示意图. 与 maxout 激活函数不同的是,MFM 使用足够的隐藏神经元来近似凸函数,仅抑制少量神经元使 CNN 模型变得轻小和健壮. MFM 在 CNN 中起到类似生物中局部特征选择的角色,在不同位置挑选不同滤波器学到最优特征. 虽然 MFM 层的梯度是稀疏的,但是伴有 MFM 的 CNN 可以获得更紧凑的表征. MFM 通过激活前面卷积层的特征图最大值获取更多富有竞争力的节点,所以 MFM 可以实现特征选择并加速生成稀疏连接.

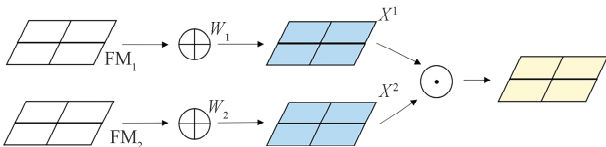


图 4 MFM 2/1 通过竞争关系抑制神经元原理图
Fig. 4 MFM 2/1 suppresses a neuron by a competitive relationship

本文采用 ArcFace^[9,20] 基础上改进对特征向量归一化和加性角度间隔,提高类间可分性,同时加强类内紧度和类间差异. 如式(9)所示, x_i 表示第 i 个样本对应的特征向量,属于第 y_i 类. W_j 是权重矩阵 W 的第 j 列. 输入图像的批量大小为 N . θ_j 是权重 W_j 和特征 x_i 之间的角度, s 为特征尺度, m 为附加的角边距惩罚项.

$$L = \frac{1}{N} \sum_i -\log \frac{e^{s(\cos\theta_{y_i,i}+m)}}{e^{s(\cos(\theta_{y_i,i}+m)} + \sum_{j \neq y_i} e^{s\cos(\theta_j,i)}} \quad (9)$$

其中:

$$W = \frac{W^*}{\|W^*\|},$$

$$x = \frac{x^*}{\|x^*\|},$$

$$\cos(\theta_j,i) = W_j^T x_i.$$

通过不同姿态数据集进行相应的姿态人脸特征提取模型训练,最终得到 7 个姿态人脸特征提取

模型集合供后续人脸识别过程中的人脸特征提取选择。

2.3 大姿态人脸比对

考虑到为了得到鲁棒性的姿态人脸特征提取模型,本文划分了不同的姿态范围分别针对特定姿态进行模型训练。考虑在实际模型部署过程中,同时用所有训练好的模型进行特征提取和比对极度耗时,而且我们也很难利用已有的3D人脸信息。为此,本文构建了一个基于姿态估计的大姿态人脸识别测试框架。该框架主要分为三部分:(1)待测试图像姿态估计及模型选择和特征提取;(2)3D注册人脸模型特定姿态投影及模型选择和特征提取;(3)人脸姿态特征比对及识别。

基于以上三个步骤,本文提出的大姿态人脸比对算法框架相对于常规的2D人脸特征比对框架而言,在不增加额外计算复杂的前提下,显著提升人脸姿态识别的准确性,同时又极大利用了现有的三维人脸图像信息。

3 实验

3.1 数据集

考虑到本文提出的算法框架主要依赖于3D人脸数据,而当前开源人脸识别数据集很难满足实际的训练要求。本文在训练过程中构建了一个混合3D-2D人脸识别数据库 Hybrid 3D+2D (Face Recognition DataSet, HFRD)。在测试过程中为了分别评估本文提出的人脸识别算法框架的泛化性和有效性,本文分别在实验室场景下和无约束自然场景下验证所提算法框架的性能。进一步,本文分别构建了约束场景下的人脸识别基准集 (Constrained Face Recognition Ground-Truth DataSet, CFRGT) 和无约束条件下的人脸识别基准集 (Unconstrained Face Recognition Ground-Truth DataSet, UCFRGT)。

(1) HFRD 数据库:该数据库首先融合了1000个个体的高精度三维人脸数据以及对应的多姿态二维人脸图像,并按角度进行姿态类别分组。考虑到1000个训练数据量严重不足,数据库又融入了如300W-LP^[21]和Multi-pie^[22]等公开人脸数据集,并对这些公开人脸数据集的人脸姿态进行人工筛选、标记和分组。通过融合不同数据库样本图像,同时采取数据增强来扩充人脸样本数据,最终达到共计1万人的约50万二维人脸纹理图像训练样本。样本如图5所示。

(2) CFRGT 数据库:该数据库在实验室环境下采集,包含388个个体的高精度三维人脸数据信息和对应的多角度二维人脸姿态图像。其中,三维人脸图像采用川大智胜公司的精度为0.1 mm的高精度三维人脸全脸照相机获取,二维人脸姿态图像采用相机弧度拍摄,收集人脸 *pitch* 和 *yaw* 角度方向人脸姿态信息,其中 $Pitch \in [-60^\circ, +60^\circ]$, $yaw \in [-90^\circ, +90^\circ]$ 。该数据用于评估一定约束条件下的基于3D-2D映射的大姿态人脸识别算法的性能。数据样本如图6所示。

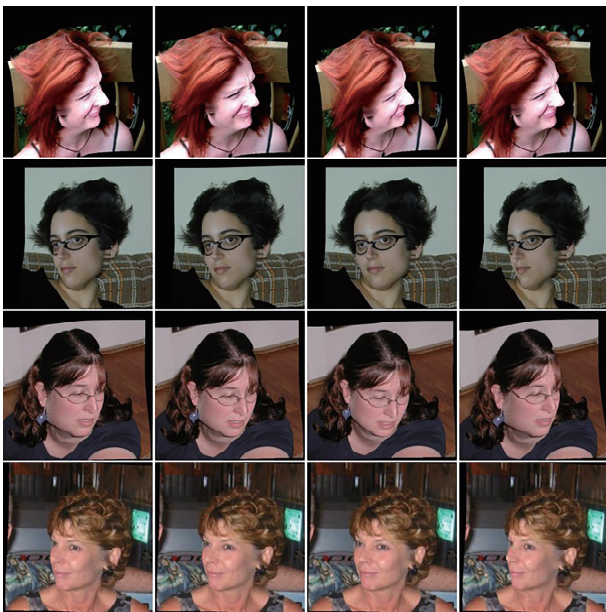


图5 HFRD数据样本示例

Fig. 5 Examples of HFRD data images

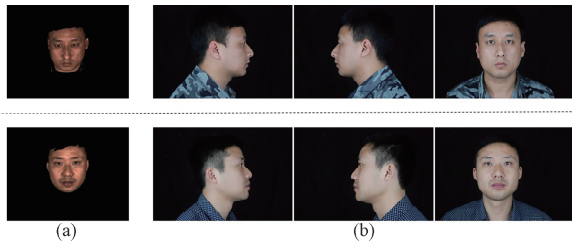


图6 CFRGT数据样本示例:(a)三维纹理投影图像;
(b)二维姿态图像

Fig. 6 Examples of CFRGT data images; (a) 3D texture; (b) 2D pose image

(3) UCFRGT:该数据采集了107个个体的高精度三维人脸图像信息和非约束条件下的二维人脸姿态图像。其中,高精度三维人脸图像采用川大智胜公司的高精度三维人脸全脸照相机获取,无约束二维人脸图像采集使用监控摄像头获取,在动态监控场景下随机获取每个注册人脸的130张二维人脸姿态图像。该数据用于评估非约束条件下的基

于 3D-2D 映射的大姿态人脸识别算法的性能. 数据样本如图 7 所示.

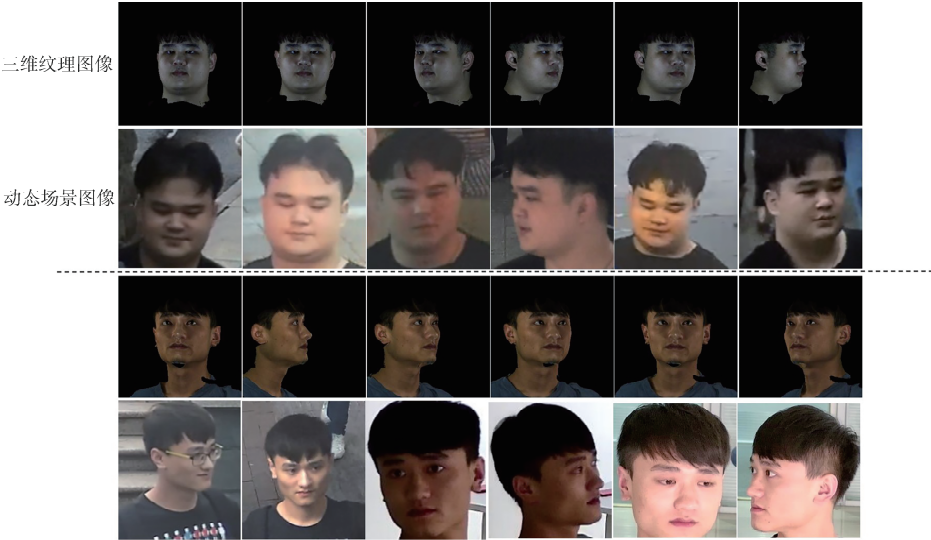


图 7 UCFRGT 数据样本示例: (a) 三维纹理投影图像; (b) 实际场景抓拍图像
Fig. 7 UCFRGT data image example: (a) 3D texture projection image; (b) snapshot of scene

3.2 评估指标

针对本文提出的 3D-2D 人脸识别框架,区别于常规的人脸识别算法评估基准,本文采用了新的评估指标来衡量算法性能,即排序提升率、TOP-1 和 TOP-5 命中率. 其定义如下:

- (1) 排序提升率= 结果排前率-结果靠后率;
- (2) 结果排前率=结果排前总数/测试样本总数) * 100%;
- (3) 结果靠后率=(结果靠后总数/测试样本总数) * 100%.

采用排序提升率指标主要用来反映三维方法是否对结果排序产生了较大的影响,是否有改善整体的趋势. 该方法能够有效地反映引入三维信息对人脸识别结果的改善程度. TOP-1 和 TOP-5 命中率指标定义如下:

- (1) TOP-1: 输入一张 query 图像,该人在 Rank1 出现,则视为命中;
- (2) TOP-1 命中率:命中总数/query 总数;
- (3) TOP-5: 输入一张 query 进去,该人在 Rank5 或之前出现,视为命中;
- (4) TOP-5 命中率:命中总数/query 总数.

3.3 对比方法及评估设置

为进一步比较提出方法的综合性能,本文采用了人脸姿态鲁棒性表征的识别方法作为对比方法. 为方便后续介绍,本章提出的基于 3D-2D 映射的大姿态人脸识别(3D-2D Face Pose Robust Recog-

nition, 3D-2D_FPRR)作为基准算法,考虑到对比方法仅采用了 2D 人脸图像来进行训练,用 2D_FPRR 指代对比方法.

为进一步衡量算法对于不同姿态识别性能,本文对测试集进行了划分. 考虑人脸姿态 yaw 方向的变化对人脸影响最大,本文针对 yaw 方向姿态变化进行划分为 $[-90^{\circ}, -75^{\circ}]$, $[-75^{\circ}, -60^{\circ}]$, $[-60^{\circ}, -30^{\circ}]$, $[-30^{\circ}, +30^{\circ}]$, $[+30^{\circ}, +60^{\circ}]$, $[+60^{\circ}, +75^{\circ}]$, $[+75^{\circ}, +90^{\circ}]$ 等 7 个区间测试集合,并针对这些测试结果进行实验的评估. 即图 1 中 $N=7$.

此外,本章为了衡量提出方法能够应用于大姿态人脸识别构建了 4 个测试数据库,分别是基础三维测试库、三维测试库、基础二维测试库和二维测试库. 本文还构建了一个查询数据库用于动态识别场景应用下对识别方法的测试. 每个数据库详细信息如表 1 所示.

表 1 算法测试数据集
Tab. 1 Algorithm test data set

Item	数据库描述	库名称
A	只采用三维人脸模型投射的二维图像入库	基础三维库
B	三维投影库+2D 人像库	三维测试库
C	采集人脸对应的二维图像库	基础二维测试库
D	采集人像的二维图像库+其他 2D 图像	二维测试库
E	动态抓拍图像	查询数据库

3.4 实验结果

3.4.1 CFRGT 数据集 表 2 给出了提出方法的整体测试比对结果. 实验结果表明, 通过引入高精度三维人脸信息可以明显提升算法性能指标. 在

表 2 CFRGT 数据库方法评估结果

Tab. 2 Evaluation results of CFRGT database method

测试库	3D-2D_FPRR		2D_FPRR	
	TOP-1/%	TOP-5/%	TOP-1/%	TOP-5/%
基础二维测试库	87.87	93.92	77.85	87.89
基础三维测试库	93.55	93.42	87.52	91.09
提升率	5.68	2.50	9.67	3.20
二维测试库	75.60	86.61	61.60	70.61
三维测试库	85.68	91.76	75.63	81.86
提升率	10.08	5.15	14.03	11.25

3D-2D_FPRR 算法的前提下, 基础三维测试库比基础二维测试库的 TOP-1 命中率提升 5.68%, TOP-5 命中率提升 2.5%, 三维测试库相对于二维测试库的 TOP-1 命中率提升 10.08%, TOP-5 命中率提升 5.15%. 在 2D_FPRR 算法的前提下, 基础三维测试库比基础二维测试库的 TOP-1 命中率提升 9.67%, TOP-5 命中率提升 3.2%, 三维测试

库相对于二维测试库的 TOP-1 命中率提升 14.03%, TOP-5 命中率提升 11.25%. 本节进一步计算了提出方法的三维排前率、三维靠后率和排序提升率, 具体结果见表 3.

表 3 CFRGT 数据库方法性能比较

Tab. 3 Performance comparison of CFRGT database methods

测试库	三维靠后率/%	三维排前率/%	排序提升率/%
3Dvs2D 基础测试库	7.10	20.33	13.23
3Dvs2D 测试库	3.13	63.58	60.45

基于 CFRGT 数据集对不同姿态人脸正确识别率进行统计和分析的结果如表 4 所示. 实验结果表明, 本文提出的算法框架可以在不同姿态人脸数据的识别率上有不同程度提升. 对于 $[-90^{\circ}, -75^{\circ}]$, $[-75^{\circ}, -60^{\circ}]$, $[+60^{\circ}, +75^{\circ}]$, $[+75^{\circ}, +90^{\circ}]$ 等大姿态情况下人脸识别率提升明显, 最高达 23%. 相对于仅仅依赖跨姿态人脸特征鲁棒性提取方法, 三维人脸对二维人脸的识别方法不仅引入三维人脸信息辅助特定姿态特征提取, 同时也将多模型特征提取方法集成到人脸识别框架中, 为解决大姿态人脸识别问题提供了一个有效的解决方案.

表 4 CFRGT 数据集人脸姿态角度与人脸正识率统计

Tab. 4 Face pose angle and face recognition rate statistics in CFRGT dataset

Method	Pose angles(yaw)						
	$[-90^{\circ}, -75^{\circ}]$	$[-75^{\circ}, -60^{\circ}]$	$[-60^{\circ}, -30^{\circ}]$	$[-30^{\circ}, +30^{\circ}]$	$[+30^{\circ}, +60^{\circ}]$	$[+60^{\circ}, +75^{\circ}]$	$[+75^{\circ}, +90^{\circ}]$
2D_FPRR/%	42	55	65	87	66	48	40
3D-2D_FPRR/%	65	71	80	93	75	66	56
提升率/%	23	16	15	6	9	18	16

3.4.2 UCFRGT 数据 针对无约束条件下的人脸识别, 其姿态变化相对于实验室环境下采集的人脸姿态样本, 其场景变化更为丰富. 对于基于姿态鲁棒性特征提取方法来说, 特征学习更为困难. 表 5 给出了在该数据库上的识别方法的测试结果.

表 5 UCFRGT 数据库方法评估结果

Tab. 5 Evaluation results of UCFRGT database method

测试库	3D-2D_FPRR		2D_FPRR	
	TOP-1/%	TOP-5/%	TOP-1/%	TOP-5/%
基础二维测试库	77.85	87.89	67.82	77.81
基础三维测试库	87.52	91.09	71.52	80.03
提升率	9.67	3.20	3.70	2.22
二维测试库	43.60	54.61	33.50	45.65
三维测试库	75.63	81.86	53.61	70.21
提升率	32.03	27.25	20.11	24.56

表 6 UCFRGT 数据库方法性能比较

Tab. 6 Performance comparison of UCFRGT database methods

测试库	三维靠后率/%	三维排前率/%	排序提升率/%
3Dvs2D 基础测试库	8.10	15.33	7.23
3Dvs2D 测试库	4.13	53.58	49.45

实验结果表明, 在 3D-2D_FPRR 算法的前提下, 基础三维测试库比基础二维测试库的 TOP-1 命中率提升 9.67%, TOP-5 命中率提升 3.2%, 三维测试库相对于二维测试库的 TOP-1 命中率提升 32.03%, TOP-5 命中率提升 27.25%. 在 2D_FPRR 算法的前提下, 基础三维测试库比基础二维测试库的 TOP-1 命中率提升 3.7%, TOP-5 命中率提升 2.22%, 三维测试库相对于二维测试库的 TOP-1 命中率提升 20.21%, TOP-5 命中率提升 24.56%. 表 6 同样展示了提出方法的三维靠后率、

三维排前率和排序提升率. 表 7 进一步给出了提出的方法在各个 γ_{aw} 姿态角度下的人脸识别正确率

结果. 基于 3D 人脸的跨姿态人脸识别方法的性能正确识别率明显提升, 识别率最高提升 22%.

表 7 UCFRGT 数据集上人脸正识率与姿态角度统计

Tab. 7 Face recognition rate and attitude angle statistics on UCFRGT dataset

Method	Pose angles(γ_{aw})						
	$[-90^{\circ}, -75^{\circ}]$	$[-75^{\circ}, -60^{\circ}]$	$[-60^{\circ}, -30^{\circ}]$	$[-30^{\circ}, +30^{\circ}]$	$[+30^{\circ}, +60^{\circ}]$	$[+60^{\circ}, +75^{\circ}]$	$[+75^{\circ}, +90^{\circ}]$
2D_FPRR/%	38	45	60	82	65	41	32
3D-2D_FPRR/%	54	65	73	85	75	60	54
提升率/%	16	20	13	3	10	19	22

4 结 论

本文提出的基于 3D-2D 映射的大姿态人脸识别算法框架, 利用三维人脸图像信息来辅助完成二维大姿态人脸识别, 为大姿态人脸识别难题提供了新的解决方案. 首先, 我们基于注册的 3D 人脸图像利用 3D-2D 映射的姿态数据生成算法达到数据扩充的目标; 然后, 针对不同姿态训练对应姿态的人脸特征提取模型得到模型集合并集成到大姿态人脸识别框架. 实验结果表明, 本文提出的算法框架可以避开直接三维人脸特征提取的复杂工作, 同时又充分利用了三维人脸图像的姿态信息. 在大姿态场景下, 相比仅采用了二维人脸图像的人脸识别算法的人脸正识率提升 20% 以上, 可以很好地解决无约束真实场景下的大姿态人脸识别.

参考文献:

[1] Taskiran M, Kahraman N, Erdem C E. Face recognition: Past, present and future (a review)[J]. Digit Signal Process, 2020, 106: 102809.

[2] Guo J, Zhu X, Zhao C, *et al.* Learning meta face recognition in unseen domains [C]// Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020.

[3] 杜文超, 邓宗平, 赵启军, 等. 基于局部形状约束网络的人脸对齐[J]. 四川大学学报: 自然科学版, 2017, 54: 953.

[4] 胡静, 陶洋. 基于 RPCA 的群稀疏表示人脸识别方法[J]. 重庆邮电大学学报: 自然科学版, 2020, 32: 459.

[5] 蔡晓云, 尹贺峰. 基于结构化局部约束低秩表示的人脸识别[J]. 江苏大学学报: 自然科学版, 2020, 41: 256.

[6] 孙俊, 才华, 朱新丽. 基于双重注意力机制的深度

人脸表示算法[J]. 吉林大学学报: 理学版, 2021, 59: 883.

[7] Bhople A R, Shrivastava A M, Prakash S. Point cloud based deep convolutional neural network for 3D face recognition [J]. Multimed Tools Appl, 2021, 80: 30237.

[8] Sharma S, Kumar V. Voxel-based 3D face reconstruction and its application to face recognition using sequential deep learning [J]. Multimed Tools Appl, 2020, 79: 17303.

[9] Shi L, Wang X, Shen Y. Research on 3D face recognition method based on LBP and SVM [J]. Optik-Int J Light Elect Opt, 2020, 220: 165157.

[10] Di H, Ardabilian M, Wang Y, *et al.* Automatic asymmetric 3D-2D face recognition [C]// Proceedings of the International Conference on Pattern Recognition. Istanbul, Turkey: IEEE Computer Society, 2010.

[11] Kakadiaris I A, Toderici G, Evangelopoulos G, *et al.* 3D-2D face recognition with pose and illumination normalization [J]. Comput Vis Image Und, 2016, 154: 137.

[12] Xiang X, Le H A, Dou P, *et al.* Evaluation of a 3D-aided Pose invariant 2D face recognition system [C]// Proceedings of the International Joint Conference on Biometrics. Denver, Colorado, USA: IEEE, 2017.

[13] Us A, Km B, Hd C. Illumination normalization techniques for makeup-invariant face recognition [J]. Comput Electr Eng, 2021, 89: 106921.

[14] Vishwakarma V P, Dalal S. A novel non-linear modifier for adaptive illumination normalization for robust face recognition [J]. Multimed Tools Appl, 2020, 79: 11503.

[15] Zhang T, Wang H, Dong Q, *et al.* Deep disentangling siamese network for frontal face synthesis under neutral illumination [J]. IEEE Signal Proc Let,

2018, 25: 1344.

[16] 赵澜涛, 林家骏. 基于双路 CNN 的多姿态人脸识别方法[J]. 华东理工大学学报: 自然科学版, 2019, 45: 466.

[17] Luan T, Xi Y, Liu X. Disentangled representation learning GAN for pose-invariant face recognition [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Hawaii, America: IEEE, 2017.

[18] Wu X, He R, Sun Z, *et al.* A light CNN for deep face representation with noisy labels[J]. IEEE T Inf Foren Sec, 2018, 13: 2884.

[19] Deng J, Guo J, Xu E N, *et al.* ArcFace: additive angular margin loss for deep face recognition [C]// Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Hawaii, America: IEEE, 2019.

[20] Liu W, Wen Y, Yu Z, *et al.* SphereFace: deep hypersphere embedding for face recognition [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Hawaii, America: IEEE, 2017.

[21] Zhu X, Zhen L, Liu X, *et al.* Face alignment across large poses: A 3D solution [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, America: IEEE, 2016.

[22] Gross R, Matthews I, Cohn J, *et al.* Multi-PIE [J]. Image Vision Comput, 2010, 28: 807.

引用本文格式:

中 文: 李晓峰, 游志胜. 基于 3D-2D 映射的大姿态人脸识别[J]. 四川大学学报: 自然科学版, 2022, 59: 042003.

英 文: Li X F, You Z S. Large-pose face recognition based on the 3D-2D mapping [J]. J Sichuan Univ: Nat Sci Ed, 2022, 59: 042003.