

多信息融合的多姿态三维人脸面部五官标志点定位方法

郭 哲 张艳宁 林增刚

(陕西省语音与图像信息处理重点实验室 西北工业大学计算机学院 西安 710129)

摘 要 针对三维人脸模型面部五官标志点定位对姿态变化非常敏感的问题,提出了一种基于多信息融合的多姿态三维人脸五官标志点定位方法.首先对二维人脸纹理图像采用仿射不变的 Affine-SIFT 方法进行特征点检测,再利用映射关系将其投影到三维空间,并采用局部邻域曲率变化最大规则和迭代约束优化相结合的方法对面部五官标志点进行精确定位.在 FRGC2.0 和自建 NPU3D 数据库的实验结果表明,文中方法无需对姿态和三维数据的格式进行预先估计和定义,算法复杂度低,同时对人脸模型的姿态有着较强的鲁棒性,与现有五官标志点定位方法相比,有着更高的定位精度.

关键词 三维人脸模型;标志点定位;多信息融合;多姿态;迭代约束优化

中图法分类号 TP391 DOI号: 10.3724/SP.J.1016.2012.00163

Multi-Pose 3D Facial Landmark Localization Based on Multi-Model Information

GUO Zhe ZHANG Yan-Ning LIN Zeng-Gang

(Shaanxi Key Laboratory of Speech & Image Information Processing (SAIIP),
School of Computer Science, Northwestern Polytechnical University, Xi'an 710129)

Abstract This paper proposes a novel multi-pose 3D facial landmark localization method based on multi-model information. In the presented method, affine invariant Affine-SIFT was utilized to 2D face texture image for feature point detection, the detected points were then mapped into the corresponding 3D face model. For 3D facial surface, the local neighbor curvature maximum change and iterative constraint optimization were combined to complete the facial landmark localization. The proposed method does not need estimate and define the posture of the face model and the format of 3D face mesh, therefore is more suitable for practical application. Experimental results on FRGC2.0 and NPU3D face database show the proposed method is robust to face pose change, and has higher localization accuracy compared with the existing methods.

Keywords 3D face model; landmark localization; multi-model; multi-pose; iterative constraint optimization

1 引 言

人脸五官标志点定位是人脸建模、表情分析及人脸识别中的关键技术,准确高效的面部五官特征

提取能够为上述应用奠定良好的基础.相比二维数据而言,三维人脸数据最大的优势在于能够获取人脸显式的三维空间形状结构描述,克服传统二维数据易受光照、视图影响的缺点^[1-3].

在三维人脸五官标志点定位中,姿态的影响是

收稿日期:2011-01-18;最终修改稿收到日期:2011-02-14.本课题得到国家自然科学基金(60872145)、国家“八六三”高技术研究发展计划项目基金(2009AA01Z315)、西北工业大学基础研究基金(JC201122)资助.郭 哲,女,1984年生,博士研究生,主要研究方向为模式识别. E-mail: guozhe@mail.nwpu.edu.cn.张艳宁(通信作者),女,1967年生,博士,教授,博士生导师,主要研究领域为信号处理、图像信息处理、模式识别. E-mail: ynzhang@nwpu.edu.cn.林增刚,男,1977年生,博士,讲师,主要研究方向为模式识别.

无法避免的,因坐标轴的变化导致同一个体不同姿态下模型点的对应坐标各不相同.实用的人脸五官标志点定位系统应该是在用户不配合的情况下使用^[4],此时人的头部会以多种姿态的形式出现,因此进行标志点定位必须考虑头部姿态的变化,多姿态面部五官标志点定位也一直是人脸识别研究的难点.

与二维图像不同,三维数据大都以非均匀分布的三角网格来存储,由此造成一些传统特征提取方法如模板匹配法^[5]等较难直接应用于三维数据中.而二维人脸特征点定位的主流方法——主动形状模型(Active Shape Model, ASMs^[6])、主动表象模型(Active Appearance Model, AAMs^[7])以及一系列的改进方法^[8-11],由于采用形状模型建模的思路,需要手工标定大量的人脸特征点来进行模型训练,因而较难应用于五官轮廓特征不明显的三维人脸数据.

为解决基于三维数据的人脸面部标志点定位, Gordon^[12]首先利用曲率特征将三维人脸表面进行分割,并假设人脸对称平面严格垂直来削减姿态变化,从而完成面部特征点的提取.该方法仅在包含 24 个深度数据的小型数据库中进行了测试. Lu 等人^[13]采用子空间描述的鼻子侧面轮廓模型来筛选出最佳鼻尖点,同时利用多模态方法提取出眼角和嘴角点.该方法能扩展到任意姿态,但是姿态的量化使得算法所耗费时间以指数级增长,严重影响了实际应用. Colby 等人^[14]首先检测出一系列满足特定要求的待选点,并采用迭代最近点匹配方法(Iterative Closest Point, ICP)来判断最终适合的特征点.该方法对正面姿态人脸数据达到了 99% 的正确率,对有较大姿态和表情变化的人脸数据达到了 82% 的正确率,但是在该算法中,任意姿态特征点检测需要大约 15 s 才能完成,算法复杂度大、耗时长. Zhang 等人^[15]提出了一种多姿态三维人脸特征点提取方法,首先采用曲率特征对人脸表面进行粗分割,其次采用统计形状模型对分割后的区域进行特征点定位,从而实现姿态不变特征点的提取,但是该算法对多姿态人脸模型的特征点定位精度不高. Nair 等人^[16]提出了一种基于三维点分布模型拟合的面部特征点提取方法,无需对姿态进行估计和补偿,但是该方法需要基于预先确定的基准点来进行. Szeptycki 等人^[17]提出了一种基于面部曲率分析和几何分布模型的面部特征点定位方法,并通过由粗到精的筛选策略来完成面部特征点的提取.在

FRGC 数据库下与手动选取的特征点进行对比,鼻尖点定位误差小于 8 mm,内眼角的定位误差小于 12 mm.

由以上描述可以看出大多数已有的方法或假设姿态是正面的,或以某种方法对姿态进行估计并由此确定平均分布位置下的特征点,或仅对正面姿态人脸模型适用.然而,当三维模型姿态未知或姿态估计过于耗费时间时,以上方法均不适合实际应用.

为了克服姿态变化对定位精度的影响,五官标志点定位算法必须对人脸的姿态不敏感,同时能够在非均匀分布三角网格数据中进行.三维人脸数据获取时大都包含相应的二维纹理图像,同时二维纹理图像数据与三维网格数据存在一一对应的映射关系,因此可以利用这种对应关系结合二维图像和三维曲面在人脸五官描述的优势来完成人脸面部五官标志点的定位.根据以上分析,本文提出一种多姿态三维人脸五官标志点定位方法,该方法采用多信息结合的思路,借助二维图像中五官结构轮廓清晰的优势,在多姿态二维人脸图像中提取特征点,并将其映射到相应的三维空间中进行表示;对三维人脸数据采用曲面上局部突变描述的迭代约束进行特征点筛选,最终准确定位出人脸面部五官标志点.

2 基于二维纹理图像的特征点检测

二维人脸图像中的五官结构轮廓相比三维曲面较为清晰,因此可以首先在二维纹理图像中进行人脸特征点检测.而对于二维图像的特征点检测也已有很多有效的方法, Lowe^[18]提出的 SIFT (Scale Invariant Feature Transform) 就是其中最具有代表性的方法. SIFT 描述子具有对旋转、尺度缩放、平移、亮度变化的不变特性,同时对视角变化、噪声也保持一定程度的稳定性,因此适合多姿态人脸图像特征点的检测.但是在实际应用中,不同相机获取到的人脸图像由于相机光轴方向变化往往产生仿射畸变失真, SIFT 虽然对仿射变换具有一定的稳定性,但仍不能克服其影响.为了解决 SIFT 对仿射变换不鲁棒的缺点, Morel 等人^[19]提出了 Affine-SIFT (ASIFT),该方法能够对尺度、相机经度角和相机纬度角等三个参数进行仿真,同时对旋转和平移参数进行归一化,因此在保留 SIFT 方法优良性能的基础上,对仿射变换具有不变性.

ASIFT 方法进行人脸二维纹理图像特征点匹配的具体实现步骤如下:

1. 将每幅人脸二维图像对由前视位置绕相机光轴方向可能引起的仿射形变进行仿真,该形变由 2 个参数决定:相机经度角 ϕ 和纬度角 θ . 图像经由依照转换倾斜度相关参数 $t = \left| \frac{1}{\cos\theta} \right|$ 进行 ϕ 旋转. 文中,一个由 t 描述的在 x 方向上的转换倾斜度可以表示为 $u(x, y) \rightarrow u(tx, y)$. 对于数字图像而言,转换倾斜度由一个定向的 t 子采样来完成,该过程需要首先在 x 方向上使用一个有着标准偏差 $c\sqrt{t^2-1}$ 的高斯函数来对图像进行卷积操作. 其中 $c=0.8$,该值已被证实能够保证有着非常小的重叠误差^[20].

2. 旋转和转换倾斜度变量通过一系列有限的纬度和经度角来调节,这些参数的采样步骤保证仿真图像与其它由相机经度角 ϕ 和纬度角 θ 产生的视图相似性较强.

3. 所有仿真图像用相似度不变匹配方法 SIFT 来进行比较.

由图 1、图 2 可以看出 ASIFT 对仿射变换具有不变性,同时由于其保留了 SIFT 本身对尺度、旋转、平移的不变特性,因此 ASIFT 能够较为鲁棒地检测同一个体不同姿态人脸图像对间相匹配的特征点对. 可以看出,ASIFT 提取的特征点对并非正确对应,大多数特征点在人脸五官描述方面并没有实际意义,对于人脸识别所需要的最为重要的五官描述标志点,绝大多数特征点为冗余特征. 为了有效地进行人脸分析,必须从提取出的特征点中抽取出最具有代表性的五官特征描述点.

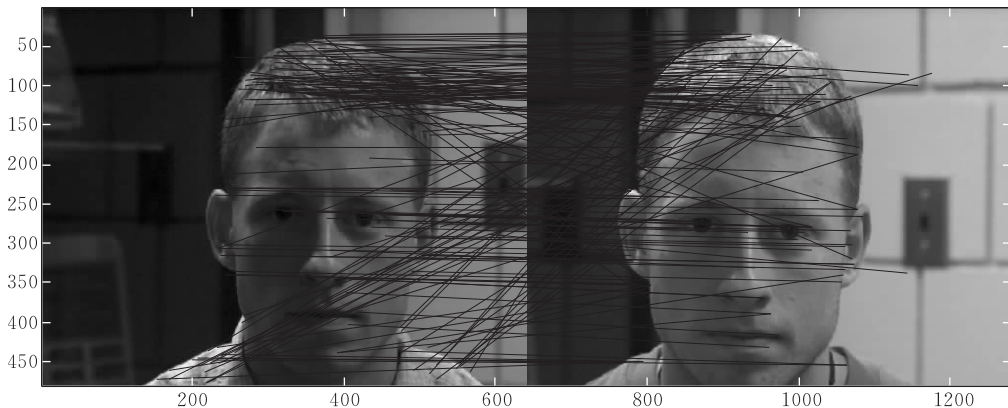


图 1 ASIFT 方法正面姿态人脸图像特征点提取结果(FRGC2.0 库)

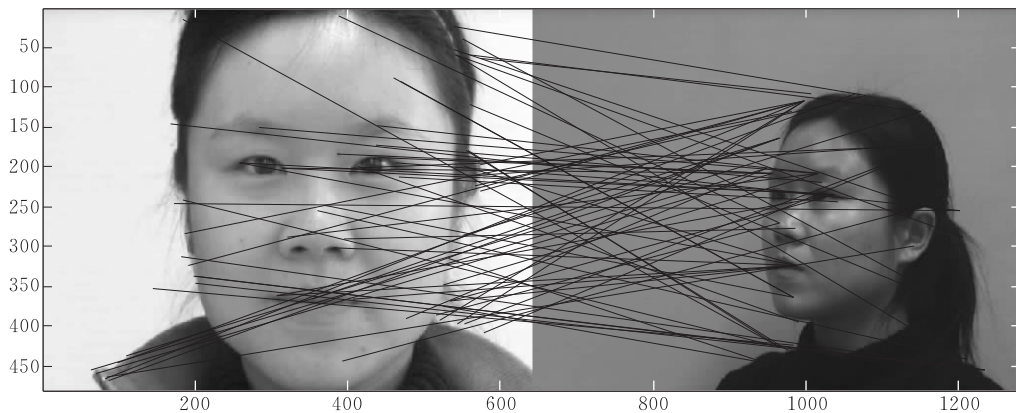


图 2 ASIFT 方法仿射畸变人脸图像特征点提取结果(NPU3D 库)

3 三维空间标志点定位

三维人脸库中的二维纹理图像数据 $U = \{u_i \in R^2; i=1, 2, \dots, N\}$ 与三维网格数据 $V = \{v_j = (x_j, y_j, z_j) \in R^3; j=1, 2, \dots, M\}$ 存在映射关系,记为 $\phi: u_i \rightarrow v_j, i=1, 2, \dots, N | v_j \notin 0$. 将第 2 节得到的二维图像的特征点向三维空间进行映射,二维图像中非人脸区域内检测出的特征点在三维人脸数据中对

应点的特征值为 0,即非人脸区域,由此可以将非面部区域点从三维人脸数据中自动剔除,但在面部区域仍存在较多无特定意义的非五官描述点,需要进一步筛选.

人脸区别于其它自然图像最为明显的不同在于有其特定的描述规范,即人脸的面貌特征,而在面貌特征中,又以五官特征为主导,反应了不同人脸之间最为明显的区别. 在五官区域的描述中,曲率是较为简单准确的特征,因为五官区域的曲率都有着较为

明显的变化. 以五官中的鼻子为例, 鼻尖点的曲率变化是其邻域中最为明显的. 针对该特征, 本文提出一种局部邻域突变描述和迭代约束优化相结合的人脸面部五官标志点精确定位方法, 首先采用局部邻域曲率变化最大的规则来对第二部分提取出的特征点进行筛选; 然后, 利用迭代约束的优化方法精确定位五官标志点.

3.1 五官局部突变描述

以 $P\{p_i, i=1, 2, \dots, N\}$ 表示三维人脸数据集, 对该数据集中的每一个顶点 p_i , 其最大和最小曲率分别表示为 $k_{1_{p_i}}$ 和 $k_{2_{p_i}}$, 能够描述曲面结构的重要信息, 而由最大最小曲率计算出的三维表面上的脊线和谷线能够较好地描述人脸五官区域的凹凸变化^[21], 分别以 l_{ridge} 和 l_{valley} 表示, 其计算原则为

$$l_{\text{ridge}} = \{p_i, \text{ if } (k_{1_{p_i}} > k_{\text{thresh1}}) \mid p_i \in P\} \quad (1)$$

$$l_{\text{valley}} = \{p_i, \text{ if } (k_{2_{p_i}} < k_{\text{thresh2}}) \mid p_i \in P\} \quad (2)$$

其中 k_{thresh1} 、 k_{thresh2} 分别为局部区域的门限曲率, 其值可以由遗传算法计算得出. 由以上定义脊线和谷线的计算原则, 能分别得到人脸五官区域位置的脊线和谷线.

假设映射到三维空间的面部区域特征点集由 $P_F\{p_{F_j}, j=1, 2, \dots, M; M < N\}$ 表示, 在该集合中按照式(1)和(2)计算满足脊线和谷线的点, 并保留相应的脊线和谷线. 令特征点集 P_F 内的脊线和谷线分别标记为 $P_{F|\text{ridge}}$, $P_{F|\text{valley}}$, 并将两集合合并为 $P_{F|\text{mark}}$, 表示人脸五官的局部突变点集.

3.2 迭代约束的五官标志点定位

为了从五官局部突变点集合 $P_{F|\text{mark}}$ 中确定最终的标志点, 需要建立不同姿态人脸数据集中五官标志点间的约束关系, 并采用迭代优化的方式进行修正, 达到最优匹配. 算法思想为: 采用 spin image 算法^[22] 建立不同姿态人脸特征点对间的约束关系, 然后用最小二乘法进行匹配优化, 对匹配结果进行统计分析, 得出某一偏差值 θ , 使绝大多数点均在该偏差值以内, 以此偏差值作为阈值滤除偏差大的点, 建立新的集合 $P'_{F|\text{mark}}$, 认为被滤出的点是非五官标志点, 直到相邻两次阈值之差小于给定的误差值, 将此时集合 $P'_{F|\text{mark}}$ 中的点确定为人脸五官标志点.

算法过程如下:

1. 假设不同姿态三维人脸数据集 M_f, M_{n_f} , 分别对 M_f 和 M_{n_f} 的突变点集合 $P'_{F|\text{mark}}$ 和 $P'_{F|\text{mark}}$ 中的每一个标志点, 建立 spin image 形状编码, 并比较这些 spin images 建立 M_f 和 M_{n_f} 特征点对间的约束关系 $R = \{p'_{i,F|\text{mark}} \leftrightarrow p'_{j,F|\text{mark}} \mid i, j \in N\}$;

2. 对 R 建立的点对约束, 计算 $P'_{F|\text{mark}}$ 和 $P'_{F|\text{mark}}$ 各点的

匹配误差 $D(p'_{i,F|\text{mark}}, p'_{j,F|\text{mark}})$, 统计得出偏差阈值 ξ ;

3. 去除 $P'_{F|\text{mark}}$ 和 $P'_{F|\text{mark}}$ 中误差值大于阈值 ξ 的点, 建立新的集合 $P''_{F|\text{mark}}$ 和 $P''_{F|\text{mark}}$, 转步 2, 若相邻两次阈值差小于给定误差值, 迭代结束;

4. 记此时的 $P''_{F|\text{mark}}$ 和 $P''_{F|\text{mark}}$ 即为相应人脸数据模型中的五官标志点.

经过迭代优化, 眉峰点、内、外眼角点、鼻尖点、嘴角点等共 9 个人脸五官标志点被确定, 对于正面姿态、中性表情人脸模型, 以上 9 个标志点均能精确定位; 对于姿态变化较大的中性表情人脸模型, 如绕 Y 轴左、右旋转 60° 或 90° 的人脸模型, 由于采集数据中标志点本身有所缺失, 因此能确定的标志点个数仅为 5~6 个.

图 3 分别给出了不同姿态、表情人脸模型五官标志点定位结果(见下页). 由图 3 可以看出, 对于正面姿态或者平面内旋转姿态变化较小的中性表情人脸模型, 9 个五官标志点均能较好的定位; 但是对于平面外旋转角度较大的中性表情人脸模型(如图 3 中第 2 行, 第 3 列), 由于采集数据标志点所在区域存在数据缺失, 因此仅能定位 6 个标志点, 同时标志点定位精度也有所下降. 而对于表情变化的人脸模型, 由于表情变化造成人脸表面的严重变形, 对五官局部突变影响较大, 且由此建立的迭代约束关系 R 与实际误差亦较大, 因此虽然标志点定位个数在 7 个以上, 但是定位出的标志点精度明显比中性表情人脸模型标志点定位精度低(如图 3 中第 3 行表情变化人脸模型).

4 实验结果与分析

4.1 实验数据

采用国际基准三维人脸库 FRGC2.0^① 以及所在实验室建立的 NPU3D 数据库进行对比实验. FRGC2.0 三维人脸库中包含 4003 个不同个体在 4 种不同光照下的三维人脸扫描数据(不仅有表面形状数据, 还包括相应的纹理图像), 但每个个体仅采集正面姿态的三维人脸数据. 作者所在实验室自建的 NPU3D 人脸库是目前国际上包含人脸变换因素最多的东方三维人脸库, 共采集 300 人的数据, 每个个体采集 35 幅包含姿态、表情、遮挡等变化的人脸数据, 共包含 10 500 幅人脸数据. 姿态变化考虑了三维空间的旋转, 每个个体采集 14 种不同姿态变

① <http://frvt.org./FRGC/>

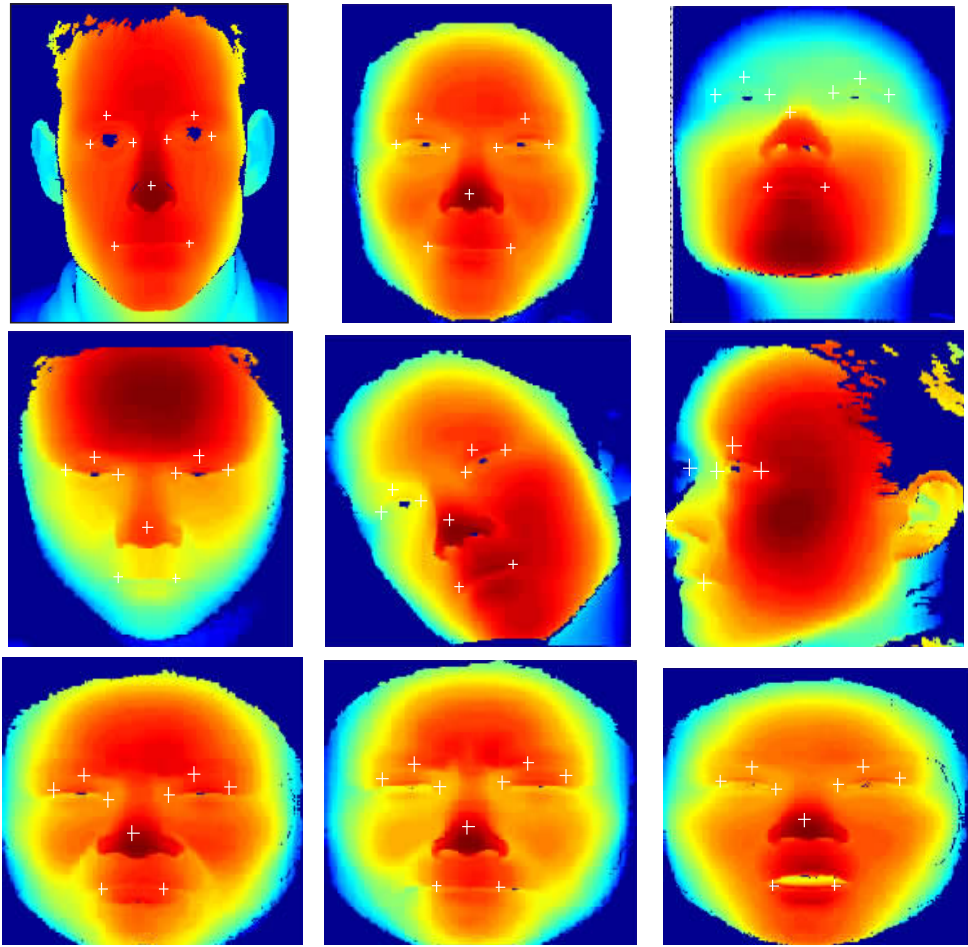


图3 FRGC2.0数据库正面姿态模型和NPU3D数据库多姿态、多表情模型标志点定位结果(左上为FRGC2.0人脸数据结果,其它为NPU3D多姿态、多表情人脸数据结果)

化的人脸数据,分别为4幅正面姿态,绕Y轴左、右分别旋转 30° 、 60° 、 90° ,俯视,仰视,绕Z轴旋转左、右旋转 45° ,较为全面地反映了人脸的三维空间姿态变化。

为了较为系统地对本文提出的五官标志点定位方法进行评价,将包含姿态变化因素较多的NPU3D数据库按照姿态旋转轴和角度大小的不同划分为若干子库,划分规则如表1所示,由于FRGC2.0库中仅包含正面姿态数据,因此将该数据库整体作为正面姿态(NF)子库。

表1 NPU3D姿态变化子库

姿态变化子库名称	FRGC2.0	NPU3D
正面姿态(NF)	$4 \times 4003 = 16012$	$4 \times 300 = 1200$
平面内绕X轴姿态变化(PIX) (俯视,仰视)	NULL	$2 \times 300 = 600$
平面内绕Z轴姿态变化(PIZ) (绕Z轴左、右旋转 45°)	NULL	$2 \times 300 = 600$
平面外姿态较小变化(POS) (绕Y轴左、右旋转 30°)	NULL	$2 \times 300 = 600$
平面外姿态较大变化(POL) (绕Y轴左、右旋转 60°)	NULL	$4 \times 300 = 1200$

4.2 实验结果与分析

为了检测本文所提方法的有效性,在FRGC2.0数据库和实验室自建NPU3D数据库中进行对比实验,分别将个数据库划分为一个基准集合和一个测试集合,对于FRGC2.0数据库,由于数据库中每个个体仅包含4个正面三维人脸数据,因此选取200个人的三维人脸数据,每个个体第1个人脸数据作为基准集合,采用手动标记点的方法对五官标志点进行提取,后3个人脸数据作为测试集合;对于包含多种姿态变化的NPU3D数据库,基准集合包含了50个正面姿态人脸数据,测试集合包含50个个体的450个9种不同姿态的人脸扫描数据,姿态分别为:正面、俯视、仰视、绕Z轴左、右旋转 45° 、绕Y轴左、右旋转 30° 、绕Y轴左、右旋转 60° 。对以上数据集合,分别采用定位误差、定位准确率、基于定位特征点的人脸识别正确率、算法复杂度以及算法适用性等5种评价准则来对本文提出的算法与现有算法^[12-17]进行对比评估。所有的实验都运行在2.94

GHz CPU 和 1.90 GB 内存的 PC 机环境上,采用 Matlab7.8.0 编程语言实现.

4.2.1 定位误差

对于 FRGC2.0 和 NPU3D 数据库,分别在基准集中手动确定五官标志点作为基准参照点,将每一个由算法定位出的特征点与基准参照点进行比对,并计算两者的欧式距离作为定位误差,用于评价算法对姿态的鲁棒性.

图 4 和图 5 分别为 FRGC2.0 库和 NPU3D 中 9 个五官标志点的定位误差直方图.由图 4 可以看出,本文算法对于全部为正面姿态人脸数据的 FRGC2.0 库,92% 以上的模型标记点平均定位误差均在 10 mm 以内.相比同样使用 FRGC 数据库的文献[17]中的方法(鼻尖点定位误差小于 8 mm,内眼角定位误差小于 12 mm),本文方法能够精确定位更多的五官标志点(9 个).值得一提的是,文献[17]方法仅对仿真姿态变化的人脸模型进行了实验验证,对于多姿态人脸模型的有效性无从考证.由图 5 可以看出,对于包含平面内及平面外多角度旋转的多姿态变化 NPU3D 人脸数据库,本文方法能够使得大约 90% 的标记点定位误差都在 20 mm 以内,因此

本文方法对于姿态变化具有较强的鲁棒性.同时,可以看出鼻尖点和内眼角点的定位精度要高于其它标志点,这是因为鼻尖点和内眼角点的曲率要比外眼角点受表情和噪声的影响要小,同时鼻尖点和内眼角对姿态变化造成的数据缺失影响亦小于其它标志点.

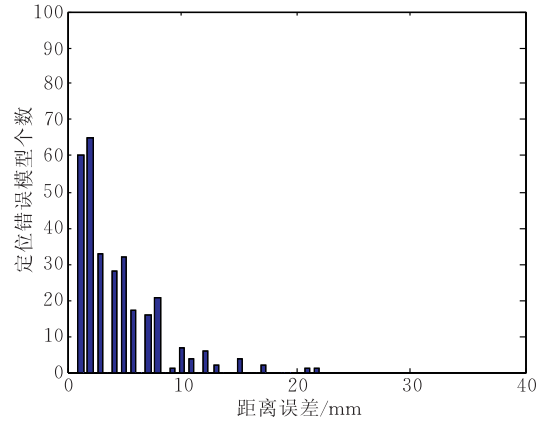


图 4 FRGC2.0 数据库人脸五官标志点平均定位误差直方图

4.2.2 定位准确率

由文献[14]可知,ICP 方法对三维模型进行配准时,容差为 20 mm,也就是说对于标记点定位误差

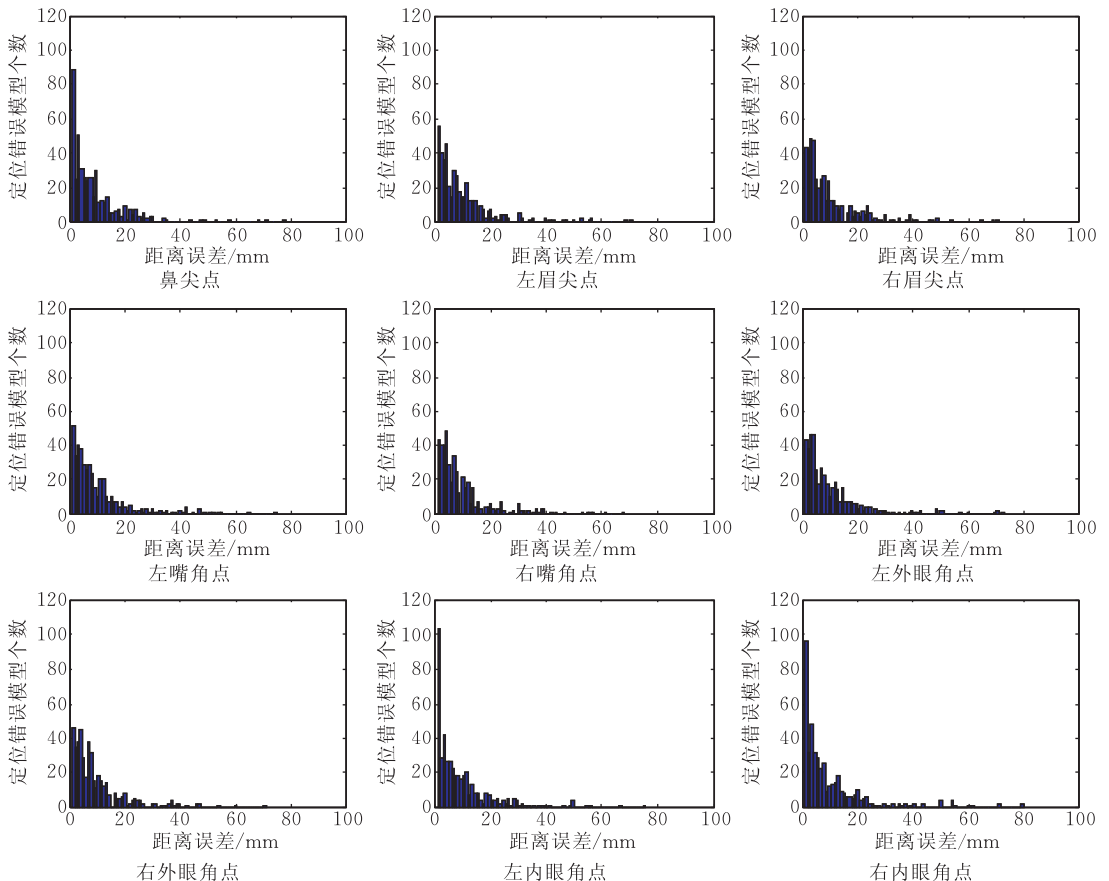


图 5 NPU3D 数据库人脸五官标志点定位误差直方图

计算来讲,如果定位误差在 20 mm 以内,可以认为定位准确.依据上述原则,对本文提出方法进行定量分析与比较,对于每一个人脸数据模型来说,如果有 6 个以上五官标志点定位误差在 20 mm 以内,即 6 个以上五官标志点定位准确,则认为该模型标志

点定位成功.由此分别计算 FRGC2.0 和 NPU3D 数据库中各个姿态子库人脸数据模型的五官标志点定位准确率,并与文献[14-15]中的方法进行对比,如表 2 所示.

表 2 不同姿态子库人脸数据的五官标志点定位准确率

姿态变化子库	本文方法 9 标志点定位正确率 FRGC2.0 数据库	本文方法 9 标志点定位正确率 NPU3D 数据库	5 标志点定位正确率 文献[15]方法	5 标志点定位正确率 文献[14]方法
NF	98.8%	98.5%	98.1%	99%
PIX	N/A	91.6%	88.3%	N/A
PIZ	N/A	89.7%	82.5%	N/A
POS	N/A	73.5%	57.7%	82%
POL	N/A	68.2%	N/A	N/A
平均定位正确率	98.8%	84.3%	84.2%	90.5%

由表 2 可知,本文提出的方法能够定位面部五官的 9 个标志点,其中眉峰点和嘴角点是文献[14-15]方法所没有考虑的 3 个标志点.3 种方法对正面姿态人脸模型的五官标志点定位准确率很高,均接近 100%.本文方法对 FRGC2.0 数据库的正面姿态人脸数据(NF)的五官标志点定位准确率达到 98.8%的正确率.对于姿态变化较小的人脸模型,如仰视、俯视人脸模型(PIX),本文方法的五官标志点定位准确率亦达到了 91.6%,高于文献[15]对 5 个标志点的定位精度 88.3%.而对于姿态变化较大的人脸模型,本文方法具有显著的优越性,对于绕 Y 轴旋转 60°的姿态变化模型(POL),达到了 68.2%的定位正确率,远远高于文献[15]方法对于绕 Y 轴旋转 30°姿态变化模型(POS)的五官标志点定位正确率 57.7%.本文方法虽然相比文献[14]方法对任意姿态 82%的定位精度,略显不足,但是值得注意的是本文方法的定位精度是在正确定位 6 个以上标志点的前提下计算的,其测评难度本身就远大于文献[14].

虽然本文提出的五官标志点定位方法对姿态变化有着较好的鲁棒性,但是,对于平面外旋转姿态变化下的定位准确率相对于其它姿态变化仍然较低,这是因为平面外旋转会引起较大的变形和数据缺失,特别是对于外眼角点、眉尖点以及嘴角点,数据缺失的可能性非常大,因此本文定义的定位正确的条件本身就无法满足,由此造成定位准确率偏低.

4.2.3 基于定位特征点的人脸识别正确率

本文提出的五官标志点定位方法的直接应用为人脸识别,即将本文定位出的五官标志点作为特征点用于人脸识别.为了验证本文方法对于人脸识别应用的有效性,将本文定位出的五官标志点作为特

征点采用多姿态三维人脸识别的基准方法——ICP^[23]进行配准识别.由此分别计算 FRGC2.0 和 NPU3D 数据库中各个姿态子库人脸数据模型的识别准确率,并与未进行特征点初始化的 ICP 方法进行比较,对比结果如表 3 所示.

表 3 不同姿态子库人脸数据的识别准确率

姿态变化子库	本文方法 FRGC2.0 库	ICP 方法 FRGC2.0 库	本文方法 NPU3D 库	ICP 方法 NPU3D 库
NF	98.2%	97.0%	98.2%	96.6%
PIX	N/A	N/A	91.1%	88.4%
PIZ	N/A	N/A	87.5%	81.2%
POS	N/A	N/A	71.0%	63.5%
POL	N/A	N/A	65.1%	58.2%
平均识别正确率	98.2%	97.0%	82.6%	77.6%

由表 3 可以看出,采用本文方法定位出的五官标志点作为特征点用 ICP 方法进行配准识别的识别率总体要比未能提供特征点计算初始迭代参数的 ICP 方法识别率高.对于姿态变化较大的人脸数据特别是平面外姿态变化数据(POS 和 POL 子库),以上结果尤为明显,采用本文方法定位的标志点作为特征点的 ICP 方法在 POS 和 POL 子库下的人脸识别率为 71.0%和 65.1%,远远高于传统 ICP 方法的 63.5%和 58.2%,可见本文方法定位出的五官标志点可以有效地用于人脸识别.

4.2.4 算法复杂度

算法的复杂度是制约算法进行实际应用的关键因素之一,对于三维人脸五官标志点定位方法来讲,五官标志点的定位结果是为了对三维人脸模型的进一步处理提供较好的初始值,因此五官定位算法的实时性成为衡量算法优良的一个重要指标.本节对不同定位算法对正面及多姿态人脸数据的平均定位

时间进行对比评估,结果如表 4 所示.

表 4 不同算法效率比较

不同方法	正面姿态数据 平均定位时间/s	姿态变化数据 平均定位时间/s
文献[14]方法	0.30	15.0
文献[15]方法	1.91	2.68
本文方法	1.83	2.15

由对比结果可知,文献[14]方法对计算量较小的正面人脸模型,特征点定位平均耗时很小,仅有 0.3 s,而对任意姿态模型,算法复杂度极高,平均耗时达 15 s,说明该方法对姿态估计的耗时较长.文献[15]方法建立的统计形状模型对姿态有着较好的鲁棒性,因此该算法复杂度对姿态相对不敏感,正面姿态与姿态变化下定位时间相差不大.本文方法对正面姿态数据的平均定位时间为 1.83 s,优于文献[15]的方法;对于姿态变化数据,本文方法有着更为明显的优势,平均定位时间仅为 2.15 s,远远小于文献[14].因此本文方法在算法复杂度和时间效率方面有着较明显的优势.

4.2.5 算法适用性

本文所采用的 NPU3D 数据库与先前的其它方法^[12-14,16-17]使用的数据库相比,包含了多种姿态的变化,同时本文提出的方法无需对姿态和三维数据的格式进行预先估计和定义.虽然文献[16]中的方法也无需对姿态进行估计和补偿,但是该方法需要基于预先确定的基准点来进行.文献[14]方法定位出了鼻尖点和内、外眼角点 5 个特征点,实验数据库中包含了与本文实验相同的正面姿态 UND 数据库(其数据亦被收录在 FRGC2.0 库中),和有着姿态和表情变化的数据库(对每个个体包含 6 幅姿态和表情变化人脸数据).对 UND 数据库人脸模型特征点的定位正确率达到 99%,平均耗时 0.3 s,对任意姿态模型定位正确率达到 86%,平均耗时达到 15 s,算法复杂度较高.同时该算法在定位嘴角和眼角点时需要预先通过对 2000 个人脸数据特征点位置的训练建立特征点的统计分布模型.文献[17]中的方法使用了与本文相同的 FRGC 数据库,其特征点定位准确率亦采用与手动选取的特征点进行对比的方法来评估,该方法的评估结果为鼻尖点定位误差小于 8 mm,内眼角的定位误差小于 12 mm,由于 FRGC 数据库中仅包含正面姿态人脸模型,因此该方法对正面姿态人脸模型进行人工姿态变化仿真,构建了多姿态变化数据,并进行了实验验证,结果达到了惊人的 95.6% 的定位准确率.但是该方法的姿

态变化数据为仿真所得,可信度较低,而本文方法在 FRGC 数据库中的 9 个五官标志点平均定位准确率达到 98.8%,与文献[17]方法的结果相当,同时也在 NPU3D 的多姿态三维人脸子库中进行了实验验证,得到了满意的定位准确率.由以上分析可知,本文提出的方法相比已有方法有着更高的五官标志点定位精度,同时适用于各种姿态变化,可以为人脸模型的下一步处理提供较为准确的标记点,算法复杂度较低,具有很强的实用性.

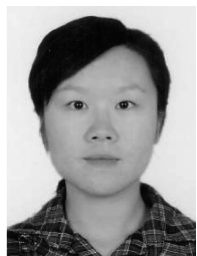
5 结 论

本文提出了一种多姿态三维人脸五官标志点定位方法,该方法采用多信息结合的思路,借助二维图像中五官结构轮廓清晰的优势,在多姿态二维人脸图像中提取特征点,并将其映射到相应的三维空间中进行表示;对三维人脸数据采用曲面上局部突变描述的迭代约束进行特征点筛选,最终准确定位出人脸面部五官标志点.在国际公共人脸库 FRGC2.0 和实验室自建东方三维人脸库 NPU3D 的实验结果表明,本文方法对 FRGC2.0 正面三维人脸模型的 9 个典型五官标志点平均定位准确率达到 98.8%,对 NPU3D 多姿态三维人脸模型库中姿态变化较小的仰视、俯视人脸模型,定位准确率达到 91.6%,高于文献[15]方法的 88.3%,对于姿态变化较大的人脸模型,本文方法具有显著的优越性,远远高于现有五官标志点定位方法的定位准确率,姿态变化数据的单次定位时间为 2.15 s.因此本文提出的多信息融合的三维人脸面部五官标志点定位方法对姿态变化有着较强的鲁棒性,时间效率满足实际应用需要,能够为三维人脸模型的进一步处理提供良好的基础.

参 考 文 献

- [1] Li S Z, Jain A K. Handbook of Face Recognition. New York: Springer, 2005
- [2] Bowyer K W, Chang K, Flynn P. A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. Computer Vision and Image Understanding, 2006, 101(1): 1-15
- [3] Abate A F, Nappi M, Riccio D, Sabatino G. 2D and 3D face recognition: A survey. Pattern Recognition Letters, 2007, 28(1): 1885-1906
- [4] Feng Z, De la Torre F, Cohn J F. Unsupervised discovery of facial events//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, USA, 2010: 2574-2581

- [5] Samir C, Srivastava A, Daoudi M. Three-dimensional face recognition using shapes of facial curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(11): 1858-1863
- [6] Cootes T F, Taylor C J, Cooper D H, Graham J. Active shape models-their training and their application. *Computer Vision and Image Understanding*, 1995, 61(1): 38-59
- [7] Cootes T F, Edwards G, Taylor C J. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(6): 681-685
- [8] Matthews I, Baker S. Active appearance models revisited. *International Journal of Computer Vision*, 2004, 60(2): 135-164
- [9] Wang Li-Ting, Ding Xiao-Qing, Fang Chi. A novel method for robust and automatic facial feature localization. *Acta Automatica Sinica*, 2009, 35(1): 9-16(in Chinese)
(王丽婷, 丁晓青, 方驰. 一种鲁棒的全自动人脸特征点定位方法. *自动化学报*, 2009, 35(1): 9-16)
- [10] Sung J, Kim D J. Pose-robust facial expression recognition using view-based 2D+3D AAM. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 2008, 38(4): 852-866
- [11] Igual L, De la Torre F. Continuous procrustes analysis to learn 2D shape models from 3D objects//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco, USA, 2010; 17-22
- [12] Gordon G G. Face recognition from depth and curvature [Ph. D. dissertation]. Division of Applied Sciences, Harvard University, USA, 1991
- [13] Lu X, Jain A K. Automatic feature extraction for multi-view 3D face recognition//*Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*. Southsampton, UK, 2006; 585-590
- [14] Colbry D, Stockman G, Jain A K. Detection of anchor points for 3D face verification//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition-Workshops*. San Diego, CA, USA, 2005; 118-125
- [15] Zhang G, Wang Y. A 3D facial feature point localization method based on statistical shape model//*Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*. Honolulu, HI, USA, 2007; 249-252
- [16] Nair P, Cavallaro A. 3D face detection, landmark localization and registration using a point distribution model. *IEEE Transactions on Multimedia*, 2009, 11(4): 611-623
- [17] Szeptycki P, Ardabilian M, Chen L. A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking//*Proceedings of the IEEE International Conference on Biometrics: Theory, Applications, and Systems*. Washington DC, USA, 2009; 1-6
- [18] Lowe D G. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision*, 2004, 60(2): 91-110
- [19] Morel J M, Yu G. ASIFT: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2009, 2(2): 438-469
- [20] Morel J, Yu G. On the consistency of the SIFT method. CMLA, ENS Cachan, France; Technical Report CMLA 2008-26, 2008
- [21] Gordon G G, Vincent L. Application of morphology to feature extraction for face recognition//*Proceedings of the SPIE/SPSE Vol. 1658, "Nonlinear Image Processing III"*. San Jose CA, USA, 1992; 151-164
- [22] Johnson A, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 1999, 21(5): 433-449
- [23] Besl P J, MacKay N D. A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(2): 239-256



GUO Zhe, born in 1984, Ph. D. candidate. Her current research interests focus on pattern recognition.

ZHANG Yan-Ning, born in 1967, Ph. D., professor, Ph. D. supervisor. Her research interests include digital signal processing, image information processing, and pattern recognition.

LIN Zeng-Gang, born in 1977, Ph. D., lecturer. His current research interests focus on pattern recognition.

Background

This paper focuses on the issue of multi-pose 3D facial landmark localization which can be widely applied in computer vision applications, such as face modeling, facial expression analysis and face recognition. Pose change is one of the

inevitable factors in 3D facial landmark localization as the practical face analysis system should be effective without any pose restrictions. Different from the 2D regular image, 3D face data present as the irregular mesh, therefore, popular

literatures based on 2D face data such as active shape model, active appearance model, and a series of the improved methods are hardly applied to 3D face data which lack of the obvious facial contour as these methods all need well-located face landmarks manually. Recent literatures on 3D facial landmark localization either suppose the posture of the face data is frontal, or estimate the face posture and obtain the feature points in the mean contributed orientation, or only fit for the frontal face data, consequently, not suitable for practical application. The proposed method does not need estimate or define the posture of the face model and the format of 3D face mesh, as based on the advantages of obviousness of 2D facial contour and concavity of 3D facial surface, therefore, is more suitable for practical application. The proposed method can achieve higher localization accuracy for face pose change compared with the existing methods. Moreover, the time complexity is much lower, and it can be used for any 3D format data, hence has wider applications. The research work and

the result of this paper is a part of the project supported by the National Natural Science Foundation of China (No. 60872145), the National High Technology Research and Development Program (863 Program) of China (grant No. 2009AA01Z315), and the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (No. 708085). The NSFC project is to solve the issues of feature correspondence and image restoration without baseline images. The 863 project aims at multi-object tracking based on multiple cameras. The project aims to provide users an efficient way to detect multiple interesting objects at one time with different postures, trajectory, and occlusions. The first task of the projects is to locate the facial landmarks of the interesting objects with multi-pose, which is the key and important step of the whole system. This paper is to address such issue. The proposed method can locate facial landmarks with high accuracy and robust to face pose change, and also has low time complexity.