

文章编号:1007-2780(2018)04-0306-11

一种多特征相结合的三维人脸关键点检测方法

冯 超¹, 陈清江^{2*}

(1. 陕西能源职业技术学院 资源与测绘工程学院, 陕西 咸阳 712000;

2. 西安建筑科技大学 理学院, 陕西 西安 710055)

摘要: 关键点检测是三维人脸识别过程中非常重要的一步, 为了提高关键点检测的准确度, 提出了一种多特征相结合的三维人脸关键点检测方法。首先对训练集中的三维人脸手工标记关键点, 计算三维人脸每一点的不同特征值, 得到每类关键点关于每个特征的均值和方差, 其次对关键点和非关键点上的特征进行线性判别分析, 得到与每个关键点相关的分值加权向量, 将前面得到的均值, 方差以及分值加权向量作为线下训练的结果输出。最后对于一个输入模型, 结合线下训练的结果得到每个关键点的候选点, 利用这些候选点构建人脸结构模型。再根据绝对距离约束, 相对位置约束, FLM 模型一致性分类, 自旋图描述符等方法确定最终的关键点。实验部分, 从 CASIA-3D FaceV1 和 FRGC V2.0 数据库中选不同姿态, 不同表情, 姿态与表情混合的三个数据集, 对其进行关键点检测。实验结果表明, 不同姿态的检测率为 94.5%, 不同表情的检测率为 94%, 和其他文献相比, 检测率平均提高了 20%, 并且有着较高的运算效率。

关键词: 有效能量; 鼻尖点检测; 姿态校正; 测地距离; 迭代最近点; 主成分分析; 分类器

中图分类号: TP391 **文献标识码:** A **doi:** 10.3788/YJYXS20183304.0306

A key point detection method of 3D face based on multi-feature

FENG Chao¹, CHEN Qing-jiang^{2*}(1. School of Resources & Surveying Engineering, Shaanxi Energy Institute,
Xianyang 712000, China;2. School of Science, Xi'an University of Architecture and Technology,
Xi'an 710055, China)

Abstract: Key point detection plays an important role in the process of 3D face recognition. In order to improve the accuracy of key points detection, a new method of 3D face key points detection based on multi feature is proposed. Firstly, the key points of the 3d face of the training set are manually marked, and the different eigenvalues of each point of 3D face are calculated, and the mean and variance of each feature are obtained for each key point. Secondly, the linear discriminant analysis is carried out on the characteristics of key points and non-critical points, thus get the score-weighted vector associated with each key point. The mean, variance, and score-weighted vectors of the previous ones are the output of the offline training. Finally, for an input model, we can get the candidate points of

收稿日期: 2017-12-11; 修订日期: 2018-02-28.

基金项目: 国家自然科学基金(No. 61403298); 陕西省自然科学基金(No. 2015JM1024)

Supported by National Natural Science Fund (No. 61403298); Shaanxi Natural Science Foundation (No. 2015JM1024)

* 通信联系人, E-mail: 280274242@qq.com

each key point combined the results of offline training, and construct the face structure model using these candidate points. According to the absolute distance constraint, relative position constraint, FLM model consistency classification, spin map and other methods we can determine the final point. In the experimental part, we select the three data sets of different postures, different expressions, gestures and expressions mixed from CASIA-3D FaceV1 and FRGC V2.0 database to detect the key points. Experiment results show that the detection rate of different gestures was 94.5%, and the detection rate of different expressions was 94%. Compared with other literatures, the detection rate increases by 20% on average. In addition, the algorithm has higher computing efficiency.

Key words: effective energy; nasal tip detection; posture correction; geodesic distance; iterative closest point; principal component analysis; classifier

1 引言

人脸识别^[1]是生物识别的一个重要分支,而关键点检测是人脸识别过程中至关重要的一个步骤。传统的二维人脸识别^[2]受光照,姿态和表情的影响,使得识别率到达一定的瓶颈期无法提高,而三维人脸上包含了更丰富,更有效的信息,并且可以找到一些光照,姿态和表情不变的特征描述符来提高人脸识别的精准度,因此研究三维人脸识别^[3-5]是有实际意义的。

三维人脸识别是指对输入的三维数据进行去噪和补洞,得到干净的,完整的三维人脸,然后对其进行姿态矫正,关键点检测,特征提取,特征匹配的过程。文献[6]提出了一种用分层机制来检测鼻子和鼻梁的位置,介绍了有效能量的概念,并用有效能量来表征鼻子附近的分布情况,最后采用SVM分类器来选择正确的鼻尖点,虽然采用了大量的数据来测试,但并没有得到准确的定位结果。文献[7]为了检测眼内角点,鼻尖点,鼻隔点,鼻翼点,结合深度图像上的曲率信息以及2D人脸分割技术,提出了一种人脸特征检测方法,尽管文章说在FRGC V2.0上有99.7%的检测成功率,但并没有给出检测成功的定义,另外,当人脸姿态偏转超过15°的时候鼻子和内眼角的检测失效。文献[8]提出了一种基于鼻梁轮廓几何分析的鼻尖点检测方法,主要应用在人脸识别过程中的姿态矫正阶段,文章没有给出明确的鼻尖点定位误差,并且此算法只适合于接近正面的人脸姿态。由此可见,当前的鼻尖点检测算法要么特征描述符单一,定位不够准确,要么姿态和表情受限,无法适用于更一般的,随机的人脸关键点

检测。

为了克服姿态,表情等的影响,文章提出了一种多特征相结合的三维人脸关键点检测方法。通过线下训练得到每个关键点的平均自旋图^[9],每个关键点关于每个特征的高斯分布以及与每个关键点相关的平均自旋图。实际应用中对于一个输入的三维数据,结合线下训练的结果得到每个关键点的候选点,通过对候选点进行筛选,组合,绝对距离与相对位置约束,FLM分类^[10],几何距离与自旋图相结和求相似度,得到最终关键点定位结果。实验结果表明,文章算法对不同姿态,不同表情的三维人脸数据都取得了较好的检测效果。和其它的文献相比较,检测率有所提高。

2 特征描述符

三维模型上关键点检测的准确度往往依赖于特征描述符的选取,特征描述符可以反映一点周围的凹凸性或者曲率特征,对于比较明显的关键点,这些点上的特征值是区别于非关键点的,因此可以通过某些特征描述符检测三维模型上的关键点。下面说明文章用到的一些特征描述符。

2.1 局部平面距离

对于三维人脸上一点 p ,取距离 p 点10 mm内的点为 p 的邻域点(假设有 N 个),设此邻域的重心为 $\bar{x}$,按照公式(1)计算协方差矩阵:

$$\text{cov} = \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})', \quad (1)$$

其中: x_i 为邻域点,对上面的协方差矩阵采用奇异值分解(SVD),得到3个特征值和对应的特征向量,取最小特征值对应的向量 n 为局部曲面法向量,采用公式(2)计算局部平面距离(Local

Plane Distance, DLP):

$$d_i = \mathbf{n} \cdot (\mathbf{x}_i - \bar{\mathbf{x}}). \quad (2)$$

一般情况下 d_i 越大, d_i 的邻域所在的曲面越凸, 因此可以用 DLP 筛选三维人脸上比较凸的区域, 一副测试图如图 1 所示(深灰色的地方代表检测到的凸的区域)。

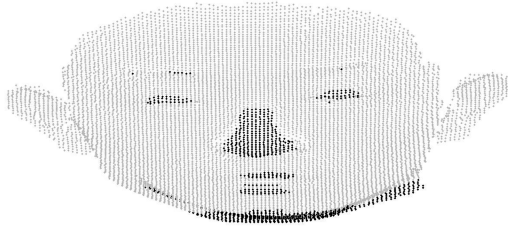


图 1 采用 DLP 的凸区域检测结果
Fig.1 Results of the convex region of DLP

2.2 自旋图描述符(Spin Image decriptor)

自旋图可以表征三维人脸上一局部形状, 对于三维人脸上一局部点 p , 设 p 点的法向量为 $\mathbf{n}$, 取距离 p 点 10 mm 内的点为 p 的邻域点(假设有 N 个), 按照下面的公式将邻域三维点 x 转化成二维点; 鼻尖点邻域的三维点投影到二维的自旋图如图 2 所示。

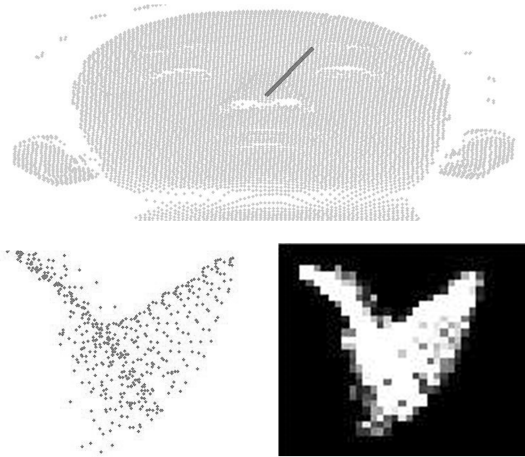


图 2 原始三维点云
Fig.2 Original 3D point cloud

$$\alpha = \sqrt{|\mathbf{x} - \mathbf{p}|^2 - (\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}))^2}, \quad (3)$$

$$\beta = \mathbf{n} \cdot (\mathbf{x} - \mathbf{p}), \quad (4)$$

其中: α 表示点 p 与点 x 的连线在 $\mathbf{n}$ 上的投影, β 表示点 x 到 $\mathbf{n}$ 的距离。

2.3 与曲率相关的特征描述符

曲率可以反映三维人脸点云上一点在邻域上

的变化快慢, 像眼睛, 嘴, 鼻轮廓上的点变化都比较大, 因此可以通过曲率检测三维人脸上的这些关键位置, 下面主要说明文章中用到的与曲率相关的一些特征描述符。

最大主曲率 k_1 和最小主曲率 k_2 : 三维曲面上经过一点有无数条曲线, 每条曲线在这点都有一个曲率, 其中曲率最大的称为最大主曲率, 曲率最小的称为最小主曲率。

LC(Log Curvedness):

$$LC = \frac{2}{\pi} \times \log \sqrt{\frac{k_1^2 + k_2^2}{2}}, \quad (5)$$

W(Willmore Energy):

$$W = \frac{(k_1 - k_2)^2}{4}, \quad (6)$$

SC(Shape Curvedness):

$$SC = \frac{4}{\pi^2} \log \sqrt{\frac{k_1^2 + k_2^2}{2}} \arctan \frac{k_1 + k_2}{k_1 - k_2}, \quad (7)$$

LD(Log Difference map)

$$LD = \ln(k_1 k_2 - \frac{k_1 + k_2}{2} + 1), \quad (8)$$

SI(Shape Index)

$$SI = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{k_1 + k_2}{k_1 - k_2}. \quad (9)$$

上述公式(5)~(9)中的 k_1 代表曲面上一点的最大主曲率, k_2 代表最小主曲率。上面的特征描述符是通过结合最大主曲率和最小主曲率得到的, 这些特征描述符可以从不同角度反映三维上一局部的局部特征。

3 平均标记模型与对齐

手工标记三维人脸点云上 8 个关键点(左外眼角, 左内眼角, 右外眼角, 右内眼角, 鼻尖点, 左嘴唇, 右嘴唇, 额前点), 由文献[11-12]可知, 8 个关键点构成的三维人脸标记模型可以由下面的向量形式表示:

$$\mathbf{x} = [p_{x,1}, \dots, p_{x,i}, p_{y,1}, \dots, p_{y,i}, p_{z,1}, \dots, p_{z,i}]^T$$

其中: $(p_{x,i}, p_{y,i}, p_{z,i})$ 为第 i 个关键点的坐标。如果训练集中有 N 个三维模型, 平均标记模型可以通过公式(10)计算:

$$\mathbf{x}_m = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i. \quad (10)$$

通过上面的计算可以得到完整人脸的平均标

记模型(FLM8),也可以得到人脸左半边脸平均标记模型(FLML5)以及右半边脸的平均标记模型(FLMR5)。上面的计算需要将所有的标记模型归一化到同一状态下进行,即以一个为标记模型为标准标记模型,其它的标记模型为样本模型,目的是要将样本模型和标准模型对齐。普氏分析提供了一种归一化方法,即计算样本模型与标准模型的普氏距离,通过不断的迭代,完成样本模型与标准模型的对齐过程,具体过程如下:设 x_0 为标准模型, x_1 为样本模型,在将样本模型和标准模型对齐之前要进行3种基本变换:平移变换,尺度变换和旋转变换。

平移变换:对任意一个标记模型 r ,利用公式(11)计算 r 的中心:

$$r_c = [\frac{1}{8} \sum_{j=1}^8 x_j, \frac{1}{8} \sum_{j=1}^8 y_j, \frac{1}{8} \sum_{j=1}^8 z_j]. \quad (11)$$

有了标记模型的中心,设 r_j 为标记模型 r 的第 j 个标记点,用公式(12)将 r 平移到原点:

$$r_j' = r_j - r_c. \quad (12)$$

尺度变换:对标记模型的每个标记点乘以一个系数,完成对标记模型的单位化:

$$r_j' = \alpha r_j, \quad (13)$$

其中: $\alpha = 1/s(x)$, $s^2(x)$ 是标记模型上说有标记点到中心点的欧氏距离的平方和:

$$s(x) = \sqrt{\sum_{j=1}^8 |r_j - r_c|^2}. \quad (14)$$

旋转变换:结合普氏分析方法,通过最小化旋转后的 x_1 与 x_0 的普氏距离来实现 x_1 与 x_0 的对齐。设 R 为旋转矩阵,则 R 可以表示成3个旋转矩阵的乘积,这3个旋转矩阵分别是与3个主方向相关的:

$$R = R_{z,\psi} \times R_{y,\varphi} \times R_{x,\theta}, \quad (15)$$

其中:

$$\begin{aligned} R_{x,\theta} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \\ R_{y,\varphi} &= \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix} \\ R_{z,\psi} &= \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

有了旋转矩阵之后,对齐的问题可以转化为

一个求解下面的最优化问题: $\min |R(x_1) - x_0|^2$, 这里 $R(x_1)$ 是 x_1 经过旋转后的模型 $R'(x_1)$ 的向量表示形式,即:

$$\begin{aligned} R'(x_1) &= R \times \begin{pmatrix} x_1^x(1) & x_1^x(2) & \cdots & x_1^x(8) \\ x_1^y(1) & x_1^y(2) & \cdots & x_1^y(8) \\ x_1^z(1) & x_1^z(2) & \cdots & x_1^z(8) \end{pmatrix} = \\ &= \begin{pmatrix} x^{x'}_1(1) & x^{x'}_1(2) & \cdots & x^{x'}_1(8) \\ x^{y'}_1(1) & x^{y'}_1(2) & \cdots & x^{y'}_1(8) \\ x^{z'}_1(1) & x^{z'}_1(2) & \cdots & x^{z'}_1(8) \end{pmatrix} \\ R(x_1) &= [x^{x'}_1(1), \cdots, x^{x'}_1(8), \\ & x^{y'}_1(1), \cdots, x^{y'}_1(8), x^{z'}_1(1), \cdots, x^{z'}_1(8)] \end{aligned}$$

由此可知, $|R(x_1) - x_0|^2$ 是与 θ, φ, ψ 相关的函数,为求 $|R(x_1) - x_0|^2$ 的最小值,对 $|R(x_1) - x_0|^2$ 分别关于 θ, φ, ψ 求偏导并使导函数等于0,通过求解可得:

$$\theta = \arctan\left(\frac{S_{x_0,y_1} - S_{y_0,z_1}}{S_{y_0,y_1} + S_{x_0,z_1}}\right), \quad (16)$$

$$\varphi = \arctan\left(\frac{S_{x_0,z_1} - S_{z_0,x_1}}{S_{z_0,z_1} + S_{x_0,x_1}}\right), \quad (17)$$

$$\psi = \arctan\left(\frac{S_{y_0,x_1} - S_{x_0,y_1}}{S_{x_0,x_1} + S_{y_0,y_1}}\right), \quad (18)$$

其中:

$$S_{x_0,x_1} = \sum_{j=1}^8 x_{0j}x_{1j}, S_{x_0,y_1} = \sum_{j=1}^8 x_{0j}y_{1j},$$

$$S_{x_0,z_1} = \sum_{j=1}^8 x_{0j}z_{1j}.$$

$$S_{y_0,x_1} = \sum_{j=1}^8 y_{0j}x_{1j}, S_{y_0,y_1} = \sum_{j=1}^8 y_{0j}y_{1j},$$

$$S_{y_0,z_1} = \sum_{j=1}^8 y_{0j}z_{1j}.$$

$$S_{z_0,x_1} = \sum_{j=1}^8 z_{0j}x_{1j}, S_{z_0,y_1} = \sum_{j=1}^8 z_{0j}y_{1j},$$

$$S_{z_0,z_1} = \sum_{j=1}^8 z_{0j}z_{1j}.$$

这样就得到了一次旋转绕 x 轴, y 轴, z 轴的角度,从而求得了旋转矩阵。而模型 x_1 与 x_0 的对齐过程,实际上是通过不断的迭代旋转,使得旋转后的 x_1 与 x_0 的普氏距离不断接近的过程。具体方法如下:

Step1: 平移 x_0 使得 x_0 的质心在坐标原点,对 x_0 进行尺度变换使得 x_0 的形状大小为1;

Step2: 平移 x_1 使得 x_1 的质心在坐标原点,对 x_1 进行尺度变换使得 x_1 的形状大小为1;

Step3: 初始化 $R = I$;

Step4: 计算 $R_{x,\theta}$, 对 x_1 进行 $R_{x,\theta}$ 变换得到 x'_1 , 令 $R = R_{x,\theta} \times R$;

Step5: 计算 $R_{y,\varphi}$, 对 x'_1 进行 $R_{y,\varphi}$ 变换得到 x''_1 , 令 $R = R_{y,\varphi} \times R$;

Step6: 计算 $R_{z,\psi}$, 对 x''_1 进行 $R_{z,\psi}$ 变换得到 x'''_1 , 令 $R = R_{z,\psi} \times R$;

Step7: 计算变换后的模型 x'''_1 与 x_0 的普氏距离, 如果 $|x'''_1 - x_0| < \epsilon$, 停止计算, 输出变换矩阵与变换后的模型 x'''_1 ; 否则, 令 $x_1 = x'''_1$, 重复 Step4~Step6, 直到满足条件为止。

4 FLM 变形及分类

将训练集中的所有对齐的三维人脸标记模型表示成向量形式(24 维), 每个向量可以看作 24 维空间中的点, 这些点构成了 24 维空间中特殊的分布, 对这些点采用主成分分析(PCA), 取最大的 p 个特征值对应的特征向量构成一个 $24 \times p$ 的矩阵 A , 根据文献^[13-14], 集合 $\{x_m, A_i, \lambda_i\}$ 构成了一个 FLM, 这里 $i \in \{1, 2, \dots, p\}$ 。有了 FLM, 平均人脸模型的形状变形就可以表示成下面的形式:

$$x' = x_m + Ab, \quad (19)$$

其中: b 为 $p \times 1$ 的形变参数向量, b 中的每个分量表示平均模型在某个方向上的形变参数, 如脸部大小变形, 鼻子扁平变形, 下巴移动变形等。 x' 表示变形后的模型。一般情况下, 关键点检测方法只能得到一些候选点, 并不能知道候选点在人脸上的具体位置, 往往需要通过分类筛选得到更可靠的候选点, 再通过相似度比较来识别和标记候选点。文章利用 FLM 变形参数对关键点候选点进行分类和筛选, 对于一个候选标记模型组合 y , 利用前面的对齐方法将 y 和 x_m 对齐, 设对齐后的模型为 y' , 将 y' 表示为向量形式, 利用公式(20)计算模型形变参数向量 b :

$$b = A^T(y - x_m), \quad (20)$$

对于 b 中的每个分量 b_i , 计算 $|b_i| < 3\sqrt{\lambda_i}$, 拿出所有满足条件的 b_i , 计算候选标记模型属于平均模型变形集合的概率:

$$p(y) = \frac{\sum \lambda_i}{v_p}, \quad (21)$$

其中: λ_i 是满足约束条件的特征值, v_p 是所有构

成 FLM 的特征值之和。给定一个阈值 T (文章中 T 取 0.9), 如果 $p(y) \geq T$, 保留候选标记模型, 否则, 弃除候选标记模型。

5 算法流程

文章的关键点检测算法主要分为两个部分。第一部分为线下训练, 这部分通过计算训练集上每一点的不同特征, 得到每个关键点关于不同特征的均值和方差, 再通过相似分值计算以及 LDA 分析, 得到不同关键点的分值加权向量。第二部分为线上测试, 对于一个输入模型, 利用线上训练得到每个关键点关于不同特征的均值和方差以及不同关键点的分值加权向量, 得到与每个关键点相关的一些候选点, 利用这些候选点构建人脸结构模型, 再根据绝对距离约束, 相对位置约束, FLM 模型一致性分类, 自旋图描述符等方法确定最终的关键点。具体流程如下:

5.1 线下训练

Step1: 将 CASIA-3D FaceV1 数据库分为训练集和测试集, 训练集中包含 300 个不同表情和姿态的三维人脸, 剩下的为测试集, 对训练集中的每个模型手工标记关键点(左外眼角, 左内眼角, 右外眼角, 右内眼角, 鼻尖点, 左嘴唇, 右嘴唇, 额前点);

Step2: 对于训练集中每个三维人脸 Ω_i , 计算 8 个关键点中每个关键点的自旋图 ($A_1^i, A_2^i, \dots, A_8^i$), 将训练集中每一类关键点的自旋图求和取平均, 得到 8 个关键点的平均自旋图 ($\bar{A}_1, \bar{A}_2, \dots, \bar{A}_8$);

Step3: 对于训练集中的每个三维人脸 Ω_i , 计算 Ω_i 上每一点的 LC, W, SC, LD, SI, DLP, 关于第 d 个特征就得到 300 个特征描述图 DL-Feature Maps, 计算 DL-Feature Maps 中所有第 λ 个关键点第 d 个特征的特征值的均值 $\mu_{d,\lambda}$ 和方差 $\delta_{d,\lambda}$, 就得到了每个关键点关于每类特征的均值和方差, 作为训练结果的第一部分;

Step4: 利用公式(22)计算三维模型上任意一点 v_i 和第 λ 个关键点关于第 d 个特征的相似度, 得到 $300 \times 6 \times 8$ 个相似分值描述图, 记为 DL-Score Maps;

$$s_\lambda^d(i) = \exp - \left[(x_d(i) - \mu_{d,\lambda})^2 / 2\delta_{d,\lambda}^2 \right], \quad (22)$$

Step5:将 $300 \times 6 \times 8$ 个相似分值描述图转化为 300×8 个相似分值向量描述图,即原先描述图上一点只有一个值,表示三维人脸上一点和某个关键点在同一特征上的相似度,经过转化后描述图上一点是一个分值向量,向量中的一个元素表示三维人脸上一点和某个关键点在同一特征上的相似度;

Step6:为了提高计算速度,采用如下一种降维的方法,对于每个关键点 λ ,取距离这个关键点 5 mm 以内的点为这个关键点的邻域点,取距离这个关键点在 15~45 mm 之间的点为这个关键点的非邻域点,与每个关键点相关的有 300 个分值描述图,将这些分值描述图中的邻域点和非邻域点对应的分值向量集合在一起,构成一个大的集合,对此集合采用 LDA 分析,得到与每个关键点相关的权重向量,具体流程图如图 3 所示。

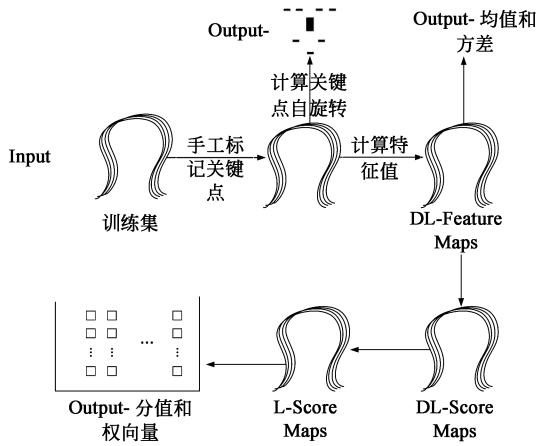


图3 线下训练过程

Fig.2 Offline training process

5.2 线上阶段

Step1: 对于输入的三维模型,分别计算模型上所有点的不同的特征值,构成 6 个特征描述图;

Step2: 对于每一个特征描述图,结合线下训练得到的每个关键点关于每类特征的均值和方差,利用公式(22)计算模型上一点和某个关键点在这个特征上的相似分值,共得到 6×8 个分值描述图;

Step3: 上面计算得到的分值描述图中,关于每一个关键点有 6 个分值描述图,将 6 个分值描述图可以融合为一个分值向量描述图,其中每个点是一个分值向量,代表三维模型上一点和某个关键点关于不同特征的特征向量,通过融合计算

总共可以得到 8 个分值向量描述图,记为 DL-Score Maps;

Step4: 结合线下训练得到的分值加权向量,利用下面的公式将 DL-Score Maps 转化为 L-Score Maps,即为了降低复杂度,提高计算效率,将三维模型上一点和某个关键点的相似分值向量转化为一个总分值

$$S_{\lambda i}(u_j^i) = \omega'_i u_j^i, \quad (23)$$

其中: u_j^i 为三维数据第 j 个顶点和第 i 个关键点的相似分值向量, ω'_i 为与第 i 个关键点相关的权重向量;

Step5: 上面的计算得到了输入模型上所有点分别与每个关键点的 8 个相似图,如第一个相似图表示输入模型上所有点与左外眼角的相似分值。从每个相似图中筛选分值最大的 30 个点作为每类关键点的候选点,从左外眼角候选点,左内眼角候选点,鼻尖点候选点,左嘴唇候选点,颏前点候选点中各任意拿出一个作为人脸左半边脸的候选形状结构模型,总共有 30^5 种可能;从右外眼角候选点,右内眼角候选点,鼻尖点候选点,右嘴唇候选点,颏前点候选点中各任意拿出一个作为人脸右半边脸的候选形状结构模型,总共有 30^5 种可能;

Step6: 根据绝对距离约束^[10]和相对位置约束^[10]对左半边脸候选形状结构模型和右半边脸

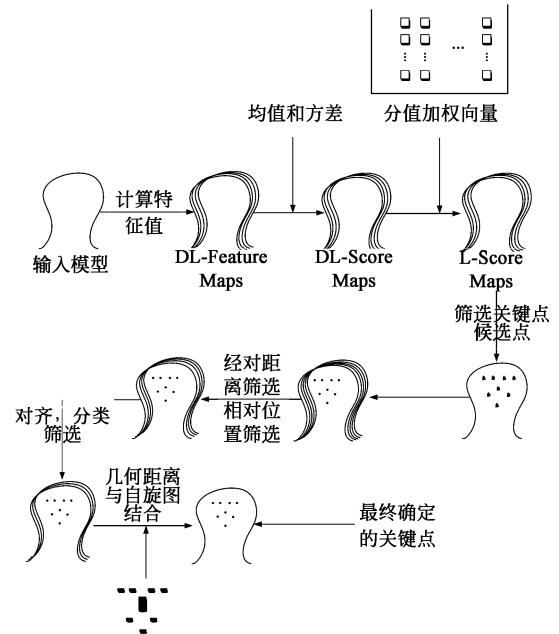


图4 线上测试阶段

Fig.4 Online test phase

候选形状结构模型进行筛选,筛选出尽量符合人脸结构的左半边脸形状结构模型和右半边脸形状结构模型;采用第二节提到的对齐方式对筛选出来的左右半边脸结构模型进行对齐;采用第三节提到的分类方法去除掉与 FLM5L 和 FLM5R 不一致的结构模型;

Step7: 经过上面的筛选,得到较少的左半边脸候选形状结构模型和右半边脸候选形状结构模型,将左半边脸候选形状结构模型和右半边脸候选形状结构模型结合成多个包含 8 个关键点的完整人脸结构模型。采用第二节提到的对齐方法对完整的人脸结构模型进行对齐,采用第三节提到的方法去除掉与 FLM8 不一致的结构模型;

Step8: 采用几何距离与自旋图相结合的方式计算剩余完整的人脸结构模型的 D_{ps} 值并进行排序,选取最优的人脸结构模型对应的点为最终检测到的关键点,具体流程图如图 4 所示。

6 实验结果与分析

文章的实验机器为联想 y50,处理器为 i7 4710HQ,仿真软件为 Matlab 2014a。实验数据采用 CASIA-3D FaceV1 数据库和 FRGC V2.0 数据库,CASIA-3D FaceV1 数据库包含了 123 个人的 4 674 个三维数据,其中每个人有大概 38 个不同表情,姿态,光照的三维数据,FRGC V2.0 数据库包含了 477 个人的 4007 个三维数据,其中每个人的数据都接近正面但带有表情变化。为了更好的说明文章算法的有效性,从 FRGC V2.0 数据库中选取 500 个作为表情变化数据集(EXVDATA),从 CASIA-3D FaceV1 数据库中选取 500 个作为偏移变化数据集(YAVDATA),从两个数据库混合选取 500 个作为一般数据集(GENDATA),并且对 EXVDATA, YAVDATA, GENDATA 手工标记关键点(计算检测误差)。这里总共分为 4 个部分:第一部分从主观视觉观察文章的算法的关键点检测效果,第二部分从客观指标来展示文章算法的关键点检测的检测误差,第三部分通过比较不同文献的检测误差来说明文章算法的可取性,第四部分说明文章算法的计算效率。

6.1 主观视觉效果

从 CASIA-3D FaceV1 数据库或者 FRGC V2.0 数据库中任意拿出一个作为输入模型,结合训练得到的每个关键点关于每个特征的均值和方差以及分值加权向量,通过组合,约束,FLM 分类之后得到候选标记模型,最终根据自旋图相似度确定最终关键点。由于数据库庞大,无法展示全部检测结果,这里仅展示两个人的不同姿态和表情的检测结果。图 5 按照从左往右从上往下的顺序前 4 个为第一个人的开心,伤心,右侧 45°,左侧 45°的原始三维数据,后 4 个为第二个人的开心,伤心,右侧 45°,左侧 45°的原始三维数据。图 6 为检测到的关键点候选点模型,图 7 为最终确定的关键点模型。

图 6 是每个关键点筛选 10 个候选点的筛选结果,图 7 是从 10 个候选点中确定的最终关键点,从视觉效果来看,的确检测到了人脸上预期的关键位置,至于偏离误差要通过下面的客观指标来说明。



图 5 原始三维数据

Fig.5 Original 3D data

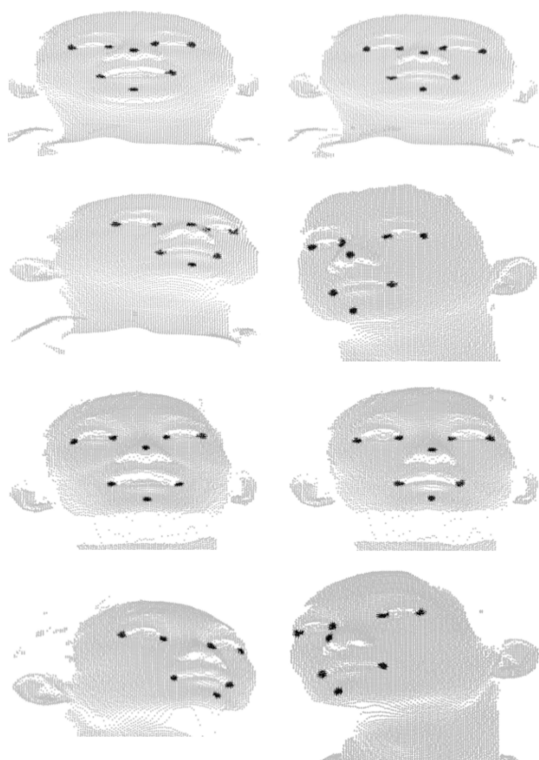


图6 候选点模型

Fig.6 Candidate model

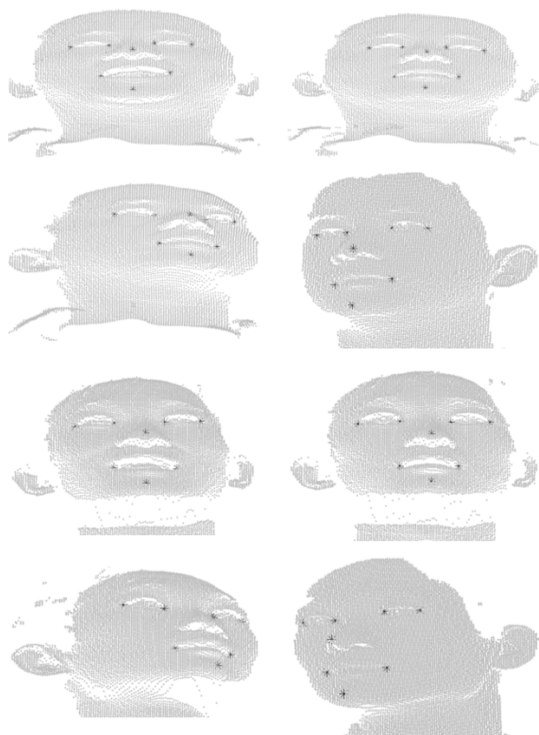


图7 最终检测到的关键点

Fig.7 Final key points detected

6.2 客观指标

为了定量分析文章算法的检测效果,分别对 EXVDATA 和 YAVDATA 进行检测实验,并计算检测位置与标记位位置的偏差,以及每类关键点偏差的均值和方差。表 1 是 YAVDATA 中左偏 22.5° 和 45° 的三维数据关于每一个关键点的检测偏差的均值和方差,表 2 是 YAVDATA 中右偏 22.5° 和 45° 的三维数据关于每一个关键点的检测偏差的均值和方差,表 3 是 EXVDATA 中中性表情的三维数据关于每一个关键点的检测偏差的均值和方差,表 4 是 EXVDATA 中极端表情的三维数据关于每一个关键点的检测偏差的均值和方差。

从表 1,表 2,表 3,表 4 可以看出,文章算法的关键点检测误差在 5 mm 左右,误差偏差不超过 3 mm,将检测误差在 3 mm 以内看作检测成功,除了极端表情下的检测率略低外,其它姿态和表情下的检测率接近 95%,说明了文章算法对姿态和表情有较好的鲁棒性。

表 1 左偏 22.5° 和 45° 的三维数据关于每一个关键点的检测偏差的均值和方差

Tab.1 Mean and variance of the detection deviation about every key point with respect to left 22.5° and 45° 3D data

偏转角度	左偏 22.5° 和 45°		
定位误差	均值(mm)	方差(mm)	<7mm
左外眼角	5.81	3.82	93.69%
左内眼角	6.00	3.83	93.40%
右外眼角	4.36	2.38	95.32%
右内眼角	6.01	3.84	95.47%
鼻尖点	5.42	3.82	92.43%
左嘴唇	4.30	2.97	93.94%
右嘴唇	4.68	3.54	93.72%
颏前点	5.24	2.35	94.73%
平均值	5.23	3.32	94.09%

表 2 右偏 22.5°和 45°的三维数据关于每一个关键点的检测偏差的均值和方差

Tab.2 Mean and variance of the detection deviation about every key point with respect to right 22.5° and 45° 3D data

偏转角度		左偏 22.5°和 45°	
定位误差	均值(mm)	方差(mm)	<7mm
左外眼角	5.49	4.01	96.90%
左内眼角	5.55	2.78	97.59%
右外眼角	4.88	3.27	93.47%
右内眼角	5.45	2.54	89.38%
鼻尖点	5.41	3.60	89.49%
左嘴唇	4.72	2.61	90.57%
右嘴唇	4.66	3.11	96.40%
颞前点	5.19	3.49	90.54%
平均值	5.17	3.18	93.04%

表 3 中性表情的三维数据关于每一个关键点的检测偏差的均值和方差

Tab.3 Mean and variance of the detection deviation about every key point with respect to neutral expression

表情		中性	
定位误差	均值(mm)	方差(mm)	<7mm
左外眼角	5.18	2.52	92.63%
左内眼角	4.21	3.24	97.77%
右外眼角	5.37	2.87	92.02%
右内眼角	4.39	2.82	96.64%
鼻尖点	4.13	3.37	96.90%
左嘴唇	4.22	2.42	97.21%
右嘴唇	4.84	3.13	92.50%
颞前点	4.60	3.13	94.39%
平均值	4.62	2.94	95.01%

表 4 极端表情的三维数据关于每一个关键点的检测偏差的均值和方差

Tab.4 Mean and variance of the detection deviation about every key point with respect to extreme expression

表情	极端		
定位误差	均值(mm)	方差(mm)	<7mm
左外眼角	5.38	5.09	89.80%
左内眼角	5.27	5.42	86.28%
右外眼角	6.39	5.09	87.57%
右内眼角	5.18	4.30	89.37%
鼻尖点	6.05	4.68	90.46%
左嘴唇	6.06	3.68	85.56%
右嘴唇	6.72	5.47	91.50%
颞前点	5.96	3.25	90.43%
平均值	5.88	4.62	88.87%

6.3 不同文献的算法结果比较

为了说明文章算法的优越性,采用文献[15],文献 [16] , 文 献 [17] 以 及 本 文 算 法 对 GENDATA 进行检测实验,并计算每个关键的检测误差以及检测成功率,如表 5 所示。

表 5 不同文献的关键点平均检测误差

Tab.5 Average detection error of key points in different literatures

比较文献	文献[16]	文献[17]	文献[18]	本文
实验数据	BU-3DFE (2350)	FRGC v1 (946)	FRGC v1 (953)	CASIA-3D FRGC V2.0
所用方法	wPDM	3D+2D	3D	3D
定位误差	均值(mm)			
左外眼角	16.44	9.38	9.01	3.75
左内眼角	13.78	9.86	9.55	4.28
右外眼角	18.11	6.35	7.10	5.90
右内眼角	15.32	4.22	6.58	3.58
鼻尖点	13.50	5.17	7.48	6.28
左嘴唇	19.39	6.15	7.59	5.69
右嘴唇	18.75	7.36	8.12	4.96
颞前点	15.50	5.20	8.53	5.23
平均值	16.35	6.71	8.00	4.96

表6 不同文献的关键点检测成功率

Tab. 6 Detection success rate of key points in different literatures

比较文献	文献[16]	文献[17]	文献[18]	本文
实验数据	BU-3DFE (2350)	FRGC v1 (946)	FRGC v1 (953)	CASIA-3D FRGC V2.0
所用方法	wPDM	3D+2D	3D	3D
定位误差	<7mm			
左外眼角	0%	81.79%	70.36%	96.89%
左内眼角	0%	78.95%	78.92%	96.83%本文
右外眼角	0%	78.67%	76.00%	92.79%
右内眼角	0%	84.87%	75.75%	95.29%
鼻尖点	0%	75.37%	80.85%	95.07%
左嘴唇	0%	83.85%	77.31%	94.78%
右嘴唇	0%	84.13%	77.41%	97.08%
颞前点	0%	82.96%	80.31%	92.85%
平均值	0%	81.32%	77.11%	95.20%
平均时间	7.11s	8.02s	6.96s	7.23s

从表5结果可以看出,文献[15]的每个关键点检测误差都在13 mm以上,平均检测误差达到了16 mm,误差较大,文献[16]的平均检测误差接近于7 mm,有所改进,文献[17]的平均检测误差为8 mm,与文献[16]的结果较为接近,而本文的平均检测误差还不到5 mm,比文献[15][16][17]的结果要好很多,并且从表6就可以看出来,在保持运行效率相对稳定的前提下,较大的提高了关键点检测率。

6.4 计算效率

文章算法耗时的部分主要体现在前期的线下训练部分,一旦有了训练结果,实际中只需要考虑

线上应用,而文章线上部分的复杂度为 $N \times o(n)$, N 为特征描述符的个数。统计了文章算法的运行时间,从读入到检测出关键点的平均总时间为7.23 s。这里面每一步骤的时间如下:计算DL-Feature Maps用时4.47 s,结合线下结果筛选关键点候选点用时1.09 s,绝对距离与相对位置筛选用时1.21 s,对齐,分类筛选用时0.97 s,几何距离与自旋图结合求相似度用时0.80 s,这些是用Matlab仿真时所花费的时间,众所周知,Matlab是一个编程效率高,但运行效率慢的实现工具,因此文章算法可通过c,c++或者并行运算快速实现,有一定的实际意义。

7 结 论

文章采用了一种多特征相结合的三维人脸关键点检测方法。通过线下训练得到每个关键点的平均自旋图,每个关键点关于每个特征的均值和方差,以及与每个关键点相关的分值加权向量。实际中输入一个模型,结合线下训练的结果得到每个关键点的多个候选点,通过对候选点进行筛选,组合,绝对距离与相对位置约束,FLM分类,最后采用几何距离与自旋图相结合得到最终关键点。实验部分采用CASIA-3D FaceV1和FRGC V2.0数据库,分别选取500个不同姿态,不同表情以及一般的三维人脸数据集,从主观视觉,客观指标以及比较不同文献的检测结果来说明文章算法的有效性。实验结果表明,不同姿态的检测率为94.5%,不同表情的检测率为94%,和其他文献相比,检测率平均提高了20%,结合运行时间可以看出,文章算法在保持运行效率相对稳定的前提下,较大的提高了关键点检测率。

参 考 文 献:

- [1] 何春.人脸识别技术综述[J].智能计算机与应用,2016,6(5):112-114.
HE C. Survey of face recognition technology [J]. *Intelligent Computer and Applications*, 2016, 6(5): 112-114. (in Chinese)
- [2] 王晓华,孙小姣.联合Gabor降维特征与奇异值特征的人脸识别[J].光学精密工程,2015,23(10):553-558.
WANG X H, SUN X J. Face recognition based on Gabor reduction dimensionality features and singular value decomposition features [J]. *Optics and Precision Engineering*, 2015, 23(10): 553-558. (in Chinese)
- [3] 刘明晶.3D人脸识别中的支持向量机研究[D].沈阳:东北大学,2009.

- LIU M J. Support vector machine in 3D face recognition [D]. Sheyang: Northeastern University, 2009. (in Chinese)
- [4] 郭哲, 樊养余, 刘妹, 等. 三维到二维: 人脸本征形状描述图[J]. 光学 精密工程, 2014, 22(12): 3391-3400.
GUO Z, FAN Y Y, LIU S, *et al.* 3D to 2D: Facial intrinsic shape description maps [J]. *Optics and Precision Engineering*, 2014, 22(12): 3391-3400. (in Chinese)
- [5] DRIRA H, BEN AMOR B, SRIVASTAVA A, *et al.* 3D face recognition under expressions, occlusions, and pose variations [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(9): 2270-2283.
- [6] XU C H, TAN T N, WANG Y H, *et al.* Combining local features for robust nose location in 3D facial data [J]. *Pattern Recognition Letters*, 2006, 27(13): 1487-1494.
- [7] SEGUNDO M P, QUEIROLO C, BELLON O R P, *et al.* Automatic 3D facial segmentation and landmark detection [C]//*Proceedings of the 14th International Conference on Image Analysis and Processing*. Modena, Italy: IEEE, 2007: 431-436.
- [8] MIAN A, BENNAMOUN M, OWENS R. An efficient multimodal 2D-3D hybrid approach to automatic face recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(11): 1927-1943.
- [9] 刘瑶, 马杰, 赵季, 等. 基于自旋图的三维自动目标识别[J]. 红外与激光工程, 2012, 41(2): 543-548.
LIU Y, MA J, ZHAO J, *et al.* Three dimensional automatic target recognition based on spin-images [J]. *Infrared and Laser Engineering*, 2012, 41(2): 543-548. (in Chinese)
- [10] PERAKIS P, PASSALIS G, THEOHARIS T, *et al.* 3D facial landmark detection under large yaw and expression variations [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(7): 1552-1564.
- [11] STEGMANN M B, GOMEZ D D. A brief introduction to statistical shape analysis [R]. Denmark: Technical University of Denmark, 2002.
- [12] COOTES T F, TAYLOR C J. Statistical models of appearance for computer vision [J]. *Proceedings of SPIE-The International Society for Optical Engineering*, 2000, 4322(1): 236-248.
- [13] PERAKIS P, THEOHARIS T, PASSALIS G, *et al.* Automatic 3D facial region retrieval from multi-pose facial datasets [C]//*Proceedings of the 2nd Eurographics Conference on 3D Object Retrieval*. Munich, Germany: ACM, 2009: 37-44.
- [14] PERAKIS P, PASSALIS G, THEOHARIS T, *et al.* Partial matching of interpose 3D facial data for face recognition [C]//*Proceedings of the 3rd International Conference on Biometrics: Theory, Applications, and Systems*. Washington, DC, USA: IEEE, 2009: 1-8.
- [15] NAIR P, CAVALLARO A. 3-D face detection, landmark localization, and registration using a point distribution model [J]. *IEEE Transactions on Multimedia*, 2009, 11(4): 611-623.
- [16] LU X G, JAIN A K. Multimodal facial feature extraction for automatic 3D face recognition [R]. East Lansing, Michigan: Department of Computer Science, Michigan State University, 2005.
- [17] LU X G, JAIN A K. Automatic feature extraction for multiview 3D face recognition [C]//*Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*. Southampton, UK: IEEE, 2006: 585-590.

作者简介: 冯超(1986—), 男, 陕西周至人, 硕士, 讲师, 主要研究领域为矿业工程、遥感技术与图像处理。E-mail: 546160401@qq.com.

陈清江(1966—), 男, 河南信阳人, 博士, 教授, 主要研究领域为计算数学、小波分析理论与图像处理。E-mail: 280274242@qq.com.