

DOI: 10.19650/j.cnki.cjsi.J1904807

基于局部和全局信息的快速三维人耳识别*

钱昱来^{1,2}, 盖绍彦^{1,2}, 郑东亮³, 达飞鹏^{1,2}

(1. 东南大学自动化学院 南京 210096; 2. 东南大学 复杂工程系统测量与控制教育部重点实验室 南京 210096; 3. 南京理工大学电子工程与光电技术学院 南京 210014)

摘要:现有基于迭代最近点法(ICP)的三维人耳识别方法计算量大,配准时间长,且容易陷入局部最优,同时用于配准的人耳含有冗余信息,对配准造成干扰。基于此提出一种基于局部和全局信息的快速三维人耳识别方法,根据内部形状特征提取关键点并实现人耳归一化;提取低维度局部描述子实现关键点匹配并得到候选列表,之后先通过快速点特征直方图进行粗配准,最后用带法向量信息的改进ICP算法进行精配准得到识别结果。基于UND-J2数据库进行身份识别实验,并与一些经典方法的结果进行对比分析。实验表明关键点特征提取仅需0.026 s,关键点匹配仅需0.015 s,耗时很短。身份识别实验获得了98.55%的Rank-one识别率,证明该方法与其他现有算法相比,识别速度更快,识别率更高。

关键词: 三维人耳识别;关键点提取;人耳归一化;局部描述子;全局配准算法

中图分类号: TP391 TH164 **文献标识码:** A **国家标准学科分类代码:** 510.4050 520.2040

Fast 3D ear recognition based on local and global information

Qian Yulai^{1,2}, Gai Shaoyan^{1,2}, Zheng Dongliang³, Da Feipeng^{1,2}

(1. School of Automation, Southeast University, Nanjing 210096, China;

2. Key Laboratory of Measurement and Control of CSE, Ministry of Education, Southeast University, Nanjing 210096, China;

3. School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210014, China)

Abstract: Most existing 3D ear recognition methods are based on the iterative closest point (ICP) algorithm that requires large amount of calculation and long matching time. It is easy to fall into the local optimum problem. Meanwhile, the ear region used for registration contains much redundant information. To solve these problems, a fast 3D ear recognition method is proposed in this study, which is based on the local and global information. Firstly, key points are detected according to internal shape features. In this way, the normalization of the ear region is achieved. Secondly, the low-dimensional local features are extracted to realize key points match. The candidate list is obtained, and the registration process is realized by fast point feature histogram and improved ICP algorithm with the normal vector of the point cloud. The proposed algorithm is evaluated by using the UND-J2 database. Experimental results show that the key point feature extraction only takes 0.026 s, and it takes 0.015 s for key point match. The identification experiments show that Rank-one recognition rate can reach 98.55%, which shows that this method has faster recognition speed and higher recognition rate than other state-of-the-art algorithms.

Keywords: 3D ear recognition; key point extraction; ear normalization; local descriptor; global registration method

0 引 言

生物识别是机器视觉和模式识别领域的经典问题,而人耳识别技术是近年的一个研究热点^[1],相比人脸等

其他生物特征,人耳更近似刚性,且不受表情变化、年龄和化妆的影响;同时人耳表面沟回起伏变化丰富^[2],适合作为生物特征进行识别。

针对二维人耳的研究已经取得了比较好的识别效果^[3-4],但其易受姿态和光照的影响,这限制了二维人耳

识别技术的进一步发展。随着近几年信息技术的发展,三维视觉逐渐走进研究与应用,越来越多的学者开始致力于三维人耳识别技术的研究^[5]。三维人耳识别技术是指利用人耳在三维空间中的拓扑结构信息进行身份辨识的技术。目前,诸多学者提出了有效的三维人耳识别方法。

人耳检测是识别的前提和基础。首先要在具有复杂背景和侧脸的图像中检测到人耳区域。根据使用数据的不同,可以分为二维图像检测人耳和直接使用三维点云检测目标。

常见二维图像中检测人耳的方法有基于 Canny 边缘检测^[6]、Adaboost 级联分类器检测^[7-8]、蛇形算法^[9]等。近年来卷积神经网络推动着目标检测领域的发展;Zhang 等^[10]提出多尺度 Faster R-CNN 算法检测人耳;Emersic 等^[11]提出卷积编码-解码网络实现人耳检测,该方法可以精确地得到人耳所在区域的像素信息;Yan 等^[5]直接在三维人耳上基于耳洞和主动轮廓算法提取人耳;Prakash 等^[12]提出了具有旋转和尺度不变性的三维人耳检测方法;郭保青等^[13]采用区域生长及欧氏聚类分割并进行改进实现点云中单物体切割。

检测出人耳之后,根据提取特征的种类,三维人耳识别方法可分为4类:基于空域直接配准的算法、基于局部特征匹配的算法、将两者融合的算法以及基于深度学习的算法。

1) 基于空域的直接配准,这类方法直接将目标人耳与库集人耳配准,计算配准误差得到识别结果。迭代最近点法^[14](iterative closet point, ICP)是最常用的配准方法,Yan 等^[5]就利用这种方法进行识别。文献^[15]在配准迭代时加入了每个点的特征,使得配准更精确;Zhang 等^[16]将人耳局部表面变化(local surface variation, LSV)特征加入迭代,取得了更高的识别率。由于人耳完全刚性,因此这类方法在精确获取人耳信息的前提下识别率较高,但计算量大。

2) 基于局部特征的匹配,这类方法抗干扰性强,匹配速度快。Ganapathi 等^[17]对每个关键点提取3种不同的特征向量进行匹配;Zhang 等^[18]分块提取人耳特征,基于多字典的协同表示框架进行分类;Sun 等^[19]构造并计算出了一个432维向量来表示局部特征,该方法特征的计算需15s左右。这类方法是为了精确描述关键点特征,特征描述子的维度通常很高,提取特征耗时较长。

3) 基于局部和全局融合的方法。Ganapathi 等^[20]融合全局描述符和4个局部三维特征描述子进行识别;吴从中等^[21]和汤兰兰等^[22]基于网格局部二值模式,融合多种特征,在三维人脸识别上取得较好效果;Zhu 等^[23]先基于曲面直方图提取局部特征进行初始关键点匹配,再利用改进体素化方法计算整体配准误差,最后用决策级融

合算法生成最终匹配分数,该方法对带遮挡的人耳效果很好;此外 Prakash 等^[24]利用基于关键点和全局信息的两步 ICP 配准方法进行识别。这类方法融合了局部和全局信息,在前两种方法的基础上进行改进与拓展,显示出较强的鲁棒性。本文方法就属于这种类型。

4) 基于深度学习的方法。Klokv 等^[25]对点云结构使用乘法变换,并进一步分享点云上变换的参数;Feng 等^[26]提出分层校正建模的分组视图(group view convolutional neural network, GVCNN)框架。基于深度学习方法和比基于特征方法的鲁棒性更强,识别率更高,但依赖各种具体场景下的大量样本训练,而在实际场景中,很多场合难以满足深度学习所需样本规模,基于此局限,本文方法未使用深度学习。

本文提出一种基于局部和全局信息的快速三维人耳识别算法,对现有方法做出了有效改进。针对现有局部特征描述子维度高、提取时间长的问题,本文局部特征仅为9维,特征提取仅耗时0.026s,与文献^[16, 18]中的提取耗时很长的高维特征相比,在速度上提高约97%,但由于特征维数低,对局部特征描述不够全面,Rank-1 识别率为92.77%,因此本文用此方法快速筛选出具有高相似度的人耳作为候选列表。针对全局配准中现有 ICP 方法易受冗余信息干扰且配准慢的问题,本文对人耳进行归一化来去除冗余信息,之后先用快速点特征直方图进行粗配准,再利用带法向量信息的 ICP 方法精配准,识别率与现有最好结果相当。相比现有算法,本文方法特征提取与匹配都很快,全局配准只需和候选列表里的人耳进行,在保证识别率的同时大大提高了识别速度。

1 人耳区域提取

本文算法首先需要在包含侧脸和背景信息的二维图像中找到人耳区域,再利用二维人耳信息提取对应的三维区域。

1.1 二维人耳检测

本文利用基于 Faster R-CNN^[27]框架的目标检测算法从侧脸图像中粗略地提取方形人耳区域。此方法在实时性和检测成功率上相比其他人耳检测方法都更好^[10]。本文成功检测出了 UND-J2^[5]数据库里的所有人耳,检测结果如图1(a)所示。

1.2 三维人耳提取

UND-J2 数据库中的所有数据都由 Minolta Vivid 910 激光扫描仪获得,其中每幅二维彩色图像都有对应的三维点云数据。检测到二维人耳区域后,可以把对应的三维点云区域提取出来,提取效果如图1(b)所示。

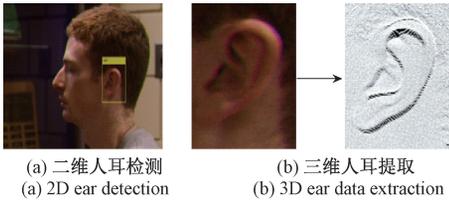


图 1 人耳检测及提取

Fig.1 Ear detection and extraction

2 关键点匹配

在关键点匹配部分,首先需要提取出关键点并去除皮肤和头发等冗余区域,接着利用局部特征实现关键点匹配。冗余信息会对后续关键点匹配以及全局配准过程造成干扰,因此需要去除。

2.1 关键点检测

为了去除冗余信息,首先需要选择稳定且具有显著表征性的关键点。文献[28]中针对三维点云提出了内部形状签名(intrinsic shape signatures, ISS)特征描述符。本文基于该特征描述符来检测关键点,设点云 P 内任意一点 p_i ,关键点的计算过程如下。

1) 搜索每个点邻域 r 范围内的所有点并对每个点 p_j 计算权重 ω_j :

$$\omega_j = 1/|p_j - p_i|, |p_j - p_i| < r \quad (1)$$

2) 利用邻域 r 范围所有点计算点 p_i 的协方差矩阵:

$$\text{COV}(p_i) =$$

$$\sum_{|p_j - p_i| < r} \omega_j (p_j - p_i) (p_j - p_i)^T / \sum_{|p_j - p_i| < r} \omega_j \quad (2)$$

3) 将协方差矩阵的特征值 $\{\lambda_1^i, \lambda_2^i, \lambda_3^i\}$ 按递减顺序排列并将满足下式的点定义为关键点:

$$\lambda_1^i / \lambda_2^i \leq \varepsilon, \lambda_2^i / \lambda_3^i \leq \varepsilon \quad (3)$$

选取更小的阈值 ε 可以获得更稳定的关键点,但点数会更少,而点数太少不利于人耳和标准模板的准确匹配来调整姿态。经过对多个样本的测试,最终将阈值 ε 设置为 0.8。得到的关键点主要集中在人耳以及右方和上方的头发处,这部分也是表面变化较为剧烈的区域,左方和下方平坦的皮肤部分几乎没有关键点,检测到的关键点如图 2 所示。

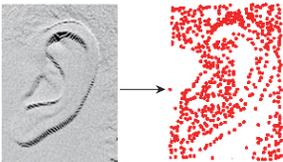


图 2 人耳关键点检测

Fig.2 Ear key point detection

2.2 人耳归一化

把 2.1 节中得到的人耳关键点和一个标准姿态人耳的关键点利用 ICP 方法进行对齐,经过几步 ICP 之后,就可以把这些特征点的姿态粗略调整到标准姿态。之后将该 ICP 变换最终矩阵作用于原人耳区域,就可以使整个对齐之后的人耳都具有竖直向上姿态,并且坐标系原点 $(0,0)$ 都位于人耳左下角。接着用长方体模板 $(0 \leq x \leq 40, 0 \leq y \leq 65, -5 \leq z \leq 5)$ 对标准姿态的人耳进行切割,得到归一化后的人耳。经过对库中大量人耳点云的测试,此模板在保留人耳点云信息的同时,可有效去除绝大部分皮肤及头发信息,同时此方法相比 Zhang 等^[15]的方法速度更快,并且保留了人耳内部的结构信息。归一化后的人耳点云约含 4 500~6 000 个点,其中关键点个数约为 250~330 个。归一化后的三维人耳如图 3 所示。



图 3 归一化后人耳

Fig.3 Normalized ear data

2.3 局部特征提取

Salti 等^[29]针对三维点云提出了一种改进的协方差矩阵描述子,根据关键点邻域信息建立局部参考坐标系,再对关键点邻域按不同方向计算直方图信息,提取出 352 维的向量用于特征描述。本文受其启发并从三维人耳特点进出发,采取这种方式构造描述子并对关键点重置三维坐标,接着从每个矩阵描述子的特征向量出发进行匹配。

1) 去中心化

在构造协方差矩阵之前,首先对关键点进行去中心化,该过程如下:对每个关键点计算其邻域半径 r 内点集的质心,并将这个点集内所有点坐标减去质心坐标。此步骤目的是消除点云原有坐标的干扰,只关注局部形状信息,使得局部形状越相似的关键点在式(5)转换至新坐标后更接近,为后续关键点匹配环节建立良好基础。

2) 提取局部特征

在去中心化的基础上,对每个关键点 p 利用其邻域内点集 p_m 构造一个改进的协方差矩阵 C ,此矩阵中,若某点距离 p 越远,则其被赋予权重越小,文献[29]表明这种方法对噪声具有更强的鲁棒性, C 的构造如下:

$$C = \frac{1}{\sum_{m: d_m \leq r} (r - d_m)} \sum_{m: d_m \leq r} (r - d_m) (p_m - p) (p_m - p)^T \quad (4)$$

式中: $d_m = \|p_m - p\|_2$, r 是构造邻域的半径。

计算矩阵 C 的特征向量, 得到 e_1, e_2, e_3 , 令 $\alpha_p = e_1$, $\beta_p = e_3 \times e_1$, $\gamma_p = e_3$, 同时构造特征描述矩阵 $[RF]_{3 \times 3} = [\alpha_p^T \beta_p^T \gamma_p^T]^T$, 并将关键点 p 转为新三维空间中的点 p_{new} :

$$p_{new} = RF \cdot p \quad (5)$$

3) 关键点匹配

利用 2.2 节的方法将待检测人耳 s_1 和库集人耳 s_2 的每个关键点都转换到新坐标中。经过上述转化, 局部形状类似的关键点在新的坐标中位置会非常接近。因此对于 s_1 中的一个关键点, 可以通过一个半径为 T_d 的最近邻搜索方法在 s_2 新坐标中找到候选匹配点集, 这将大大减少候选匹配点的搜索范围。根据本文人耳实验结果, 设定 $T_d = 3$ mm。在新坐标中, 对 s_1 中的一个新坐标关键点 p_1 , 找到 s_2 中距离在 T_d 范围内的所有关键点, 构成点集 $p_2^i (i = 1 \cdots n)$, 对 p_1 和每个 p_2^i 利用上一节得到的特征向量 $\alpha \beta \gamma$ 计算角度差值, 计算过程如下:

$$\theta_\alpha = \cos^{-1} \left(\frac{\alpha_{p_1} \cdot \alpha_{p_2^i}}{\|\alpha_{p_1}\| \cdot \|\alpha_{p_2^i}\|} \right) \quad (6)$$

$$\theta_\beta = \cos^{-1} \left(\frac{\beta_{p_1} \cdot \beta_{p_2^i}}{\|\beta_{p_1}\| \cdot \|\beta_{p_2^i}\|} \right) \quad (7)$$

$$\theta_\gamma = \cos^{-1} \left(\frac{\gamma_{p_1} \cdot \gamma_{p_2^i}}{\|\gamma_{p_1}\| \cdot \|\gamma_{p_2^i}\|} \right) \quad (8)$$

按照上述角度提取方法, 若两个关键点局部形状很相似, 则 3 个角度都很小。若局部形状差异大, 则很可能 3 个角度中会有一个或多个角度较大。因此本文选取最大的角度作为判别的出发点, 将这个最大的角度定义为这两个关键点的差异度, 对 p_2^i 中的每个点都提取出这样一个差异度。本文在 $p_2^i (i = 1, \dots, n)$ 中找到与 p_1 差异度最小的点, 作为初步对应点。采用这种方法遍历 s_1 中的每个关键点, 得到其在 s_2 中的初步对应点集。

至此获得的对应点集中, 所有对应点都具有非常相似的局部表面变化信息, 但实验中这两个点所在的区域可能相差较大, 因此初步对应点集中包含了很多误匹配点, 本文采用随机采样一致性算法 (random sample consensus, RANSAC) [30] 来剔除误匹配点对, 剔除误匹配点对前后的效果如图 4 所示。

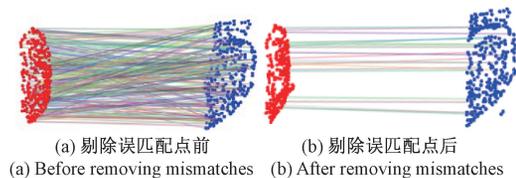


图 4 剔除误匹配点前后效果

Fig.4 Before and after removing mismatches

将库集人耳按照与待测人耳成功匹配的点对数量进行排序, 按照本文算法原理, 匹配点对数越多则是同一个人耳的概率越大。本文通过分析匹配效率与准确度, 将后续匹配的库集人耳数设定为 8, 即选择前 8 张库集人耳作为候选列表进行全局信息配准。本文只使用 $\alpha \beta \gamma$ 3 个向量共 9 维来表示关键点的形状信息, 其维数低, 特征提取时间相比文献 [19] 减少了约 97%, 同时在一定程度上保证了正确率。尽管维数低会导致对关键点的形状描述不够精确, 识别率相比高维度特征的文献 [17, 19] 等有所欠缺, 但经过实验与分析, 前 8 张库集人耳足以包含待测人耳。同时这一步对人耳进行了归一化, 去除了冗余信息, 快速筛选出具有高相似度的人耳放入候选列表, 为提高后续配准环节的精度和速度提供了良好基础。

3 全局信息配准

由于人耳完全刚性, 因此基于 ICP 的方法 [5] 取得了很高的识别率, 而文献 [15] 把点云坐标和特征信息融合, 每个向量包含三维坐标和 k 维特征, 实验表明这种改进的 ICP 方法能比原始方法取得更好的结果。这类方法需要两个点云有较好的初始姿态, 否则容易陷入局部最优。受此启发, 本文选取法向量 [31] 为不变特征, 先利用快速点特征直方图进行粗配准, 最后利用带法向量的改进 ICP 算法进行精配准并获得识别结果。

3.1 基于快速点特征直方图的粗配准

快速点特征直方图 [32] 对点特征直方图 [33] 进行改进, 时间复杂度上从 $O(nk^2)$ 减少到 $O(nk)$, 同时还具有和点特征直方图相当的区分能力。快速点特征直方图将某点表面法线和该点邻近点表面法线间的差异参数化为一个 33 维的特征向量, 并用这个特征向量来完整描述该点的几何特性。而人耳表面凹凸变化相当剧烈, 法向量的分布差异也可以很好地体现这种凹凸变化。本文采用文献 [34] 的方法进行法线估计, 先利用移动最小二乘法 (moving least squares, MLS) 将某点及其 k 个近邻点拟合在一个局部平面上, 再求出平面的法向量作为该点的法线。快速点特征直方图计算过程如下。

1) 设某点 p 在半径 r 内有 k 个邻近点, 选择任意两点 p_s 和 p_t 之间欧氏距离 d 以及对应法线 n_s 和 n_t (选择法线和两点连线夹角较小的为 p_s), 在 p_s 上定义一个局部坐标系 $UVW (U = n_s; V = (p_t - p_s) \times U; W = U \times V)$, 再计算此坐标系和 p_t 之间的角度关系, 局部坐标系如图 5 所示。

计算出的特征可以表示为:

$$\begin{cases} \alpha = V \cdot n_t \\ \varphi = U \cdot \frac{p_t - p_s}{d} \\ \theta = \tan^{-1}(W \cdot n_s, U \cdot n_t) \end{cases} \quad (9)$$

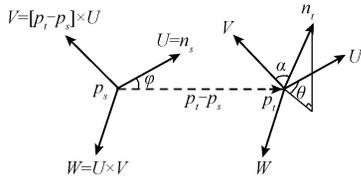


图 5 局部坐标系及特征构造

Fig.5 Local coordinate system and feature extractor

2) 计算邻域内所有点对的 $(\alpha, \varphi, \theta)$ 三组值, 构造出简化的点特征直方图 (SPF), 利用点 p 的 k 个最近邻点的 SPF 值来计算 p 的最终直方图 (FPFH):

$$FPFH(p) = SPF(p) + \frac{1}{k} \sum_{i=1}^k \frac{1}{\omega_i} \cdot SPF(p_i) \quad (10)$$

式中: ω_i 为点 p 和邻域点 p_i 的距离。

对于待测人耳点集 k_1 和库集人耳点集 k_2 , 粗配准过程如下。

1) 从 k_1 的所有点中选出 n 个作为采样点, 为了增大差异性, 这些采样点间的最小距离必须大于设定的阈值 d_i ;

2) 对于 k_1 中的每个采样点 q_i , 在 k_2 中找到直方图与 q_i 直方图最相似的 l 个点, 接这些相似点集存入一个临时列表;

3) 在临时列表中随机选取 n 个匹配点对, 根据这 n 个匹配点对计算出最优旋转和平移变换矩阵 (R, T) ;

4) 假设 m_i 为 q_i 的初步对应点, 如果 $((Rq_i + T) - m_i)^2 < \varepsilon$ (ε 为误差阈值), 则接受这一对匹配点, 将所有被接受的匹配点对作为当前一致集;

重复步骤 3)、4), 选取具有最多匹配点个数的一致集作为最优一致集, 并将其对应的最佳旋转平移变换 (R, T) 作用于待测人耳所有点云, 至此完成粗配准。

3.2 基于迭代最近点的精配准

本文最后采用法向量作为不变特征的 NICP (ICP with normals) 方法^[15]对待测人耳点云和库集人耳点云进行精配准, 和传统 ICP 方法相比, 距离度量函数除了三维坐标还加入了法向量信息。为了加快配准过程, 利用 K-D Tree 加快近邻搜索。

精配准完成后, 选取与待测人耳均方根 (root mean square, RMS) 误差最小的库集人耳作为最终识别结果。

4 实 验

本文所有实验都基于 UND-J2 数据库, 该数据库是目前公开的最大的三维人耳数据库, 库中共 415 个人, 1 801 幅二维彩色图像 (及对应的三维点云数据), 每人有至少 2 幅彩色图像和对应的三维点云。在 i5 处理器的 Windows 操作系统上通过 Visual Studio 2015 和

MATLAB R2017a 平台进行三维人耳识别实验, 实验流程如图 6 所示。

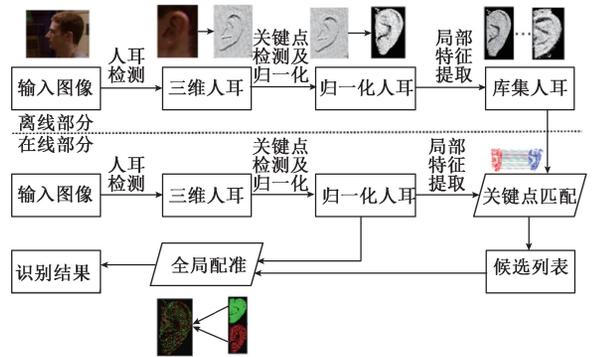


图 6 三维人耳识别流程

Fig.6 Flowchart of 3D ear recognition

4.1 人耳检测

本文使用基于 Faster R-CNN 的检测算法在二维图像侧脸中检测人耳, 从网络上和 UND-J2 数据库中共选取 350 张图片作为训练集, 训练出的检测器成功检测出了 UND-J2 库中的所有人耳, 部分检测结果如图 7 所示。

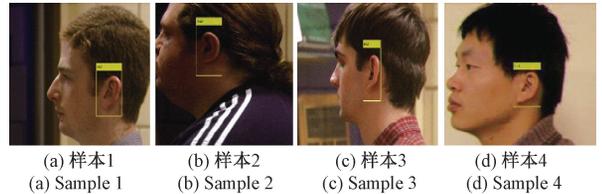


图 7 人耳检测

Fig.7 Ear detection

4.2 关键点匹配

在关键点检测时设置邻域 $r=2$ mm, 得到关键点后对人耳进行归一化, 部分归一化前后的三维人耳如图 8 所示。

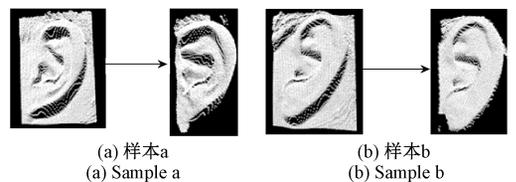


图 8 人耳归一化

Fig.8 Ear data normalization

提取出局部特征后, 先对关键点进行初步匹配, 再剔除误匹配点对, 剔除误匹配点对时的接受误差阈值设置为 2 mm。部分关键点匹配结果如图 9 所示 (图 9 (a) 和 (b) 为同一人耳匹配结果, 图 9 (c) 为不同人耳匹配结果)。

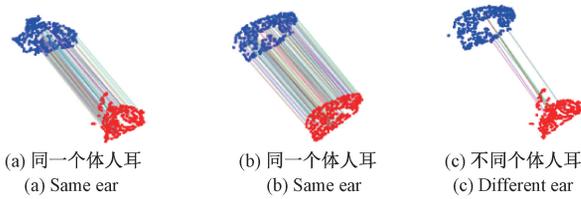


图9 人耳关键点匹配结果

Fig.9 Ear keypoints matching result

4.3 全局配准

将库集人耳按照与待测人耳成功匹配的点对个数进行排序,选取前8个作为候选列表。采用快速点特征直方图进行粗配准并利用带法向量信息的ICP算法精配准,同一人耳的配准结果如图10所示。

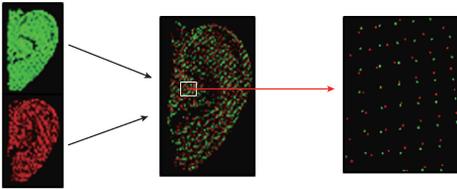


图10 人耳配准

Fig.10 Ear registration

4.4 实验结果与分析

本文方法的 Rank- r 测试结果如表1所示。为了更加充分地验证本文所提算法的有效性,使用累计匹配特性曲线(Cumulative Match Characteristic, CMC)作为识别性能指标,CMC曲线反映了系统返回最相似模型的数目。图11所示为本文识别算法和分步识别算法的性能,其中带▶曲线为只利用本文关键点匹配算法的识别性能,带◆曲线为关键点匹配筛选出候选列表后使用ICP算法识别的性能。从图中曲线的分布来看,本文算法R1RR(rank-1 recognition rate)为98.55%,相比其他方法识别率更高,这个结果证实了本文提出的算法可以很好地完成人耳识别。

表1 UND-J2库的 Rank- r Table 1 Rank- r on the UND-J2 database

Rank- r	结果/%
Rank-1	98.55
Rank-2	98.80
Rank-3	99.28
Rank-4	99.28
Rank-5	99.76

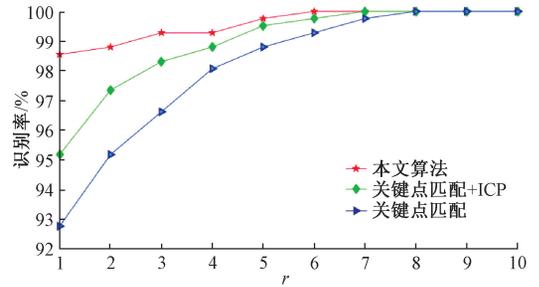


图11 UND-J2库的 CMC 曲线

Fig.11 CMC curves on UND-J2 database

4.5 与其他算法对比实验

将本文算法耗时以及该算法与现有三维人耳识别方法进行对比,结果如表2和3所示,本文实验及所有对比实验都是基于UND-J2数据库。本文算法在局部特征提取和关键点匹配环节耗时都很短,用此方法可以快速筛选出具有高相似度的候选人耳列表进行配准。Sun等^[19]为关键点匹配方法的典型代表,没有使用ICP进行配准,特征提取需15s左右,而本文只需0.026s。Islam等^[7]和Prakash等^[24]都是基于ICP方法,预处理时都一定程度上依赖二维图像的特征,同时需和所有库集人耳配准,耗时较长,而本文方法只需和候选列表里的库集人耳进行配准,节省了大量时间。Zhu等^[23]先对人耳点云进行体素化,后续非合作博弈算法(noncooperative game theory, NGT)算法需要预先对库集进行训练,对带遮挡人耳取得较好识别效果,而本文算法省去了训练的耗时。因此本文提出的基于局部和全局信息结合的算法在保证识别率的基础上,大大提高了整个识别系统的速度。

表2 本文算法耗时列表

Table 2 Time-consuming list of algorithms in this paper

步骤	耗时/s
人耳检测	0.25
关键点提取	0.1
局部特征提取	0.026
关键点匹配	0.015
全局配准	2.0
总时间	2.39

5 结论

本文提出了一种基于局部和全局信息相结合的快速三维人耳识别算法。此方法首先将待测样本和所有库

表 3 不同算法对比试验(UND-J2 库)

Table 3 Comparison results on the UND-J2 database

算法	库集图像,测试图像	预处理方法	识别方法	特征提取与匹配/s	识别率/%
文献[19]方法	404,1 376	手动切割+归一化	关键点匹配	15.23	95.10
文献[7]方法	415,415	目标检测	局部特征+ICP	0.06(局部)+2(全局)	94.00
文献[24]方法	404,1 376	边缘连接	SURF+GPA-ICP	3~4	98.30
文献[23]方法	415,415	目标检测+体素化	SPHIS+two-step NGT	0.02	98.80
本文方法	415,415	目标检测+归一化	关键点匹配+NICP	0.04(局部)+2(全局)	98.55

集样本进行关键点匹配并筛选出候选列表,这一步中特征提取时间仅为 0.026 s,相比现有高维度特征提取速度上加快了约 97%,主要优势在于本文使用的特征向量仅为 9 维,计算快。为了进一步提高识别率,将匹配点数最多的前 8 个库集人耳放入全局配准环节。由于配准需要计算法向量点对之间的夹角,因此配准时间稍长,但待测样本只需和候选列表中的 8 个库集样本进行配准,而现有的其他配准方法大多是和所有库集样本配准,因此本文全局配准环节所需时间也很短。

同时本文通过归一化获得标准姿态人耳,去除无效区域对人耳配准的干扰,所以配准更精确,计算效率更高。基于 UND-J2 数据库的实验取得了 98.55% 的 Rank-1 识别率,实验结果表明本文算法和现有算法相比在保证高识别率的同时大大提高了识别速度。在未来工作中,将进一步解决遮挡等具有挑战性问题的处理,期望获得更灵活的识别方案。

参考文献

- [1] EMERŠIĆ Ž, ŠTRUC V, PEER P. Ear recognition: More than a survey [J]. *Neurocomputing*, 2017, 255(13): 26-39.
- [2] 王凯, 穆志纯. 基于沟回结构特征的三维人耳识别方法[J]. *仪器仪表学报*, 2014, 35(2): 313-319.
WANG K, MU ZH CH. 3D human ear recognition method based on auricle structural feature [J]. *Chinese Journal of Scientific Instrument*, 2014, 35(2): 313-319.
- [3] ZHANG B Q, MU ZH CH, LI CH, et al. Robust classification for occluded ear via Gabor scale feature-based non-negative sparse representation [J]. *Optical Engineering*, 2014, 53(6): 061702.
- [4] TIAN L, MU ZH CH. Ear recognition based on deep convolutional network [C]. *IEEE International Congress on Image & Signal Processing*, 2017, doi: 10.1109/CISP-BMEI.2016.7852751.
- [5] YAN P, BOWYER K W. Biometric recognition using 3D ear shape [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(8): 1297-1308.
- [6] ANSARI S, GUPTA P. Localization of ear using outer helix curve of the ear [C]. In *Proceedings of the International Conference on Computing: Theory and Applications*, 2007: 688-692.
- [7] ISLAM S M S, DAVIES R, BENNAMOUN M, et al. Efficient detection and recognition of 3D ears [J]. *International Journal of Computer Vision*, 2011, 95(1): 52-73.
- [8] ABAZA A, HEBERT C, HARRISON M A F. Fast learning ear detection for real-time surveillance [C]. *Fourth IEEE International Conference on Biometrics: Theory Applications & Systems*, 2010: 1-6.
- [9] DEEPAK R, NAYAK A V, MANIKANTAN K. Ear detection using active contour model [C]. *IEEE International Conference on Emerging Trends in Engineering*, 2016: 1-7.
- [10] ZHANG Y, MU ZH CH. Ear detection under uncontrolled conditions with multiple scale faster region-based convolutional neural networks [J]. *Symmetry*, 2017, 9(4): 53.
- [11] EMERŠIĆ Ž, GABRIEL L L, ŠTRUC V, et al. Convolutional encoder-decoder networks for pixel-wise ear detection and segmentation [J]. *IET Biometrics*, 2018, 7(3): 175-184.
- [12] PRAKASH S, GUPTA P. A rotation and scale invariant technique for ear detection in 3D [J]. *Pattern Recognition Letters*, 2012, 33(14): 1924-1931.
- [13] 郭保青, 余祖俊, 张楠, 等. 铁路场景三维点云分割与分类识别算法 [J]. *仪器仪表学报*, 2017, 38(9): 19-27.
GUO B Q, YU Z J, ZHANG N, et al. 3D point cloud segmentation, classification and recognition algorithm of railway scene [J]. *Chinese Journal of Scientific Instrument*, 2017, 38(9): 19-27.
- [14] BESL P J, MCKAY N D. Method for registration of 3-D shapes [C]. *Sensor Fusion IV: Control Paradigms and Data Structures*. International Society for Optics and Photonics, 1992: 586-607.
- [15] SHARP G C, LEE S W, WEHE D K. ICP registration using invariant features [J]. *IEEE Transactions on Pattern*

- Analysis and Machine Intelligence, 2002, 24 (1): 90-102.
- [16] ZHANG Y, MU ZH CH, YUAN L, et al. 3D ear normalization and recognition based on local surface variation[J]. Applied Sciences, 2017, 7(1): 104.
- [17] GANAPATHI I I, ALI S S, PRAKASH S. Geometric statistics-based descriptor for 3D ear recognition[J]. The Visual Computer, 2018: 1-13.
- [18] ZHANG L, ZHU X N, LI L D. Multi-dictionary based collaborative representation: With applications to 3D ear and 3D palmprint identification [J]. Mathematical Problems in Engineering, 2019:1-13.
- [19] SUN X P, WANG G, WANG L, et al. 3D ear recognition using local salience and principal manifold [J]. Graphical Models, 2014, 76(5): 402-412.
- [20] GANAPATHI I I, PRAKASH S. 3D ear recognition using global and local features [J]. IET Biometrics, 2018, 7(3): 232-241.
- [21] 吴从中, 王浩宇, 詹曙. 融合曲面形状和纹理特征的三维人脸识别 [J]. 电子测量与仪器学报, 2018, 32(9): 150-156.
- WU C ZH, WANG H Y, ZHAN SH. 3D face recognition by fusing surface shape and texture features[J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(9): 150-156.
- [22] 汤兰兰, 盖绍彦, 达飞鹏, 等. 基于网格纵横局部二值模式的三维人脸识别 [J]. 仪器仪表学报, 2016, 37(6): 1413-1420.
- TANG L L, GAI SH Y, DA F P, et al. A 3D face recognition method based on the local binary pattern from vertical and horizontal on the mesh [J]. Chinese Journal of Scientific Instrument, 2016, 37(6): 1413-1420.
- [23] ZHU Q P, MU ZH CH. Local and holistic feature fusion for occlusion-robust 3D ear recognition [J]. Symmetry, 2018, 10(11): 565.
- [24] PRAKASH S, GUPTA P. Human recognition using 3D ear images [J]. Neurocomputing, 2014, 140 (22): 317-325.
- [25] KLOKOV R, LEMPITSKY V. Escape from cells: Deep kd-networks for the recognition of 3D point cloud models [C]. IEEE International Conference on Computer Vision (ICCV), 2017: 863-872.
- [26] FENG Y F, ZHANG Z ZH, ZHAO X B, et al. GVCNN: Group-view convolutional neural networks for 3D shape recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 264-272.
- [27] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]. Advances in Neural Information Processing Systems, 2015: 91-99.
- [28] ZHONG Y. Intrinsic shape signatures: A shape descriptor for 3D object recognition [C]. IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), 2009: 689-696.
- [29] SALT I S, TOMBARI F, DI STEFANO L. SHOT: Unique signatures of histograms for surface and texture description [J]. Computer Vision and Image Understanding, 2014, 125(8): 251-264.
- [30] FISCHLER M A, BOLLES R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24(6): 381-395.
- [31] MITRA N J, NGUYEN A. Estimating surface normals in noisy point cloud data [C]. Proceedings of the Nineteenth Annual Symposium on Computational Geometry, 2003: 322-328.
- [32] RUSU R B, BLODOW N, BEETZ M. Fast point feature histograms (FPFH) for 3D registration [C]. IEEE International Conference on Robotics and Automation, 2009: 3212-3217.
- [33] RUSU R B, BLODOW N, MARTON Z C, et al. Aligning point cloud views using persistent feature histograms [C]. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008: 3384-3391.
- [34] ALEXA M, BEHR J, COHEN-OR D, et al. Computing and rendering point set surfaces [J]. IEEE Transactions on Visualization and Computer Graphics, 2003, 9(1): 3-15.

作者简介



钱昱来, 2017 年于东南大学获得学士学位, 现为东南大学硕士研究生, 主要研究方向为三维人耳识别。

E-mail: 220171590@seu.edu.cn

Qian Yulai received his B. Sc. degree from Southeast University in 2017. He is currently a M. Sc. candidate at Southeast University. His main research interest is 3D ear recognition.



盖绍彦 (通信作者), 2008 年于东南大学获得博士学位, 现为东南大学副教授、博士生导师, 主要研究方向为三维测量及三维人脸识别。

E-mail: qxyymm@163.com

Gai Shaoyan (Corresponding author) received his Ph. D. degree from Southeast University in 2008. He is currently an associate professor and a Ph. D. advisor at Southeast University. His main research interests include 3D measurement and 3D face recognition.