

基于改进 ORB 特征的多姿态人脸识别

周凯汀¹⁾, 郑力新²⁾

¹⁾(华侨大学信息科学与工程学院 厦门 361021)

²⁾(华侨大学工学院 泉州 362021)

(1050920138@qq.com)

摘要: 为了克服通过全局特征以及每位个体单个模板样本进行多姿态人脸识别的不足, 提出基于改进的 ORB 局部特征以及每位个体多个模板样本的多姿态人脸识别方法。首先改进 ORB 算子的采样模式提高算子对人脸视角变化的鲁棒性, 并采用每位个体的多个训练样本建立模板库, 然后提取并匹配测试样本与模板样本的改进 ORB 特征。在特征提取阶段, 为避免关键点数目变化的干扰, 对全部样本提取一致数目的关键点; 在特征匹配阶段, 采用基于模型和基于方向的双重策略剔除误匹配点对, 使用匹配点对数目与平均距离评价测试样本与每个模板样本的吻合程度。对 CAS-PEAL-R1 和 XJTU 数据库的实验结果表明, 改进的 ORB 特征具有更好的识别性能; 与采用多个训练样本构建个体单个模板样本的方法相比, 在训练样本数目相同的条件下, 该方法能较好地避免姿态的干扰, 具有更好的识别效果。与 SIFT 算子相比, ORB 算子在特征提取与特征匹配 2 个阶段都具有明显的速度优势。

关键词: 人脸识别; 多姿态; 多视图; ORB; 特征匹配

中图法分类号: TP391

Multi-pose Face Recognition Based on Improved ORB Feature

Zhou Kaiting¹⁾ and Zheng Lixin²⁾

¹⁾(College of Information Science and Engineering, Huaqiao University, Xiamen 361021)

²⁾(College of Engineering, Huaqiao University, Quanzhou 362021)

Abstract: To overcome the weakness of multi-pose face recognition by global feature or by single template sample for each subject, a novel multi-pose face recognition method based on improved oriented FAST and rotated BRIEF (ORB) local feature and multiple template samples for each subject is proposed in this paper. The sampling pattern of ORB operator is first improved to enhance the robustness of the operator to the variation of viewpoint towards face, and template database is built with multiple training samples of each subject, the improved ORB features of the test sample are next extracted and matched with those of template samples. At the stage of feature extraction, consistent number of keypoints are extracted for all samples to avoid the disturbance of the variation of the number of keypoints. At the stage of feature matching, double strategies based on the model and orientation of matching-point pairs are adopted to eliminate outliers. The consistent degree of test sample and each template sample is evaluated by the number and average distance of the inliers. Experimental results on the CAS-PEAL-R1 and XJTU databases show that, the improved ORB operator has better recognition performance; compared with the methods of constructing single template sample from multiple training samples for each subject, the proposed method could better avoid the disturbance of pose variation, and obtain better recognition results under the condition of using the same number of training samples. Compared with scale invariant feature

收稿日期: 2013-11-27; 修回日期: 2014-09-02. 基金项目: 国家自然科学基金(61203369); 福建省科技创新平台基金(2013H2002); 福建省产业计划开发项目(25201071); 泉州市技术研究与开发项目(2011G74). 周凯汀(1968—), 女, 硕士, 副教授, 主要研究方向为计算机视觉、模式识别等; 郑力新(1967—), 男, 博士, 教授, 主要研究方向为机器视觉、模式识别等.

transform (SIFT) operator, the ORB operator has obvious advantages in speed at both stages of feature extraction and feature matching.

Key words: face recognition; multi-pose; multi-view; oriented fast and rotated BRIEF; feature matching

近年来, 人脸识别在人的身份识别与鉴别中得到了广泛的应用, 应用比例仅次于指纹识别而排在第 2 位^[1]。目前, 人脸识别研究主要集中在正面人脸识别。当人脸与摄像机之间存在一个变化的角度时, 所形成的为多姿态人脸。由于视角的变化, 不同姿态下人脸图像的相关性迅速下降, 给多姿态人脸识别造成困难。随着人脸识别技术和应用的发展, 多姿态人脸识别已经得到关注。

根据初始的训练样本数目, 多姿态人脸识别分为单视图与多视图方法。单视图方法根据模型建立单视图模板样本与其他姿态测试样本的映射^[2-5], 适用于缺乏多姿态模板样本的场合。多视图方法又可分为三维模型和二维的多视图方法: 三维模型法^[6]利用正面和侧面的 2 幅正交的人脸图像建立人脸的三维模型, 然后利用该模型的投影产生任意角度的多姿态人脸模板样本, 识别的效果依赖于建模精度; 二维的多视图方法^[7-8]采集不同姿态的人脸图像构建二维的模板样本, 由于样本充足, 可望取得较好的识别效果。

目前, 多姿态人脸识别大多延用了正面人脸的成功识别算法, 如基于主成分分析(principal component analysis, PCA)、线性判别分析、Gabor 小波特征等的人脸识别方法。这些方法首先从原始人脸图像中根据一定标准提取特征向量, 然后将待测人脸与模板人脸进行匹配。PCA 等方法提取的是整个人脸图像的特征, 由于模板人脸只包含有限个姿态, 而待测人脸的姿态是连续变化的, 当待测人脸与模板人脸姿态差别较大时 PCA 等方法难以取得较好的识别效果。近年来, 图像局部关键点(特征点)检测与描述算子的研究取得了突破性的进展, 已经提出多种特征算子, 如浮点或整数数据类型的 SIFT(scale invariant feature transform)^[9], SURF(speeded-up robust features)算子, 以及二进制数据类型的 ORB(oriented fast and rotated BRIEF)^[10], BRISK(binary robust invariant scalable keypoints)算子等, 它们对光线、旋转、尺度和视角变化都具有较强的鲁棒性。SIFT 已经应用于正面人脸识别^[11-12]以及多姿态人脸识别^[8]。SIFT 特征为

128 维的浮点或整数数据类型, 具有较高的计算复杂度。为了克服 SIFT 特征计算复杂度高的缺陷以及增强人脸识别算法的鲁棒性, 本文提出基于改进 ORB 特征的多姿态人脸识别方法, 通过对 ORB 描述子的采样模式提出改进措施, 提高特征对人脸视角变化的鲁棒性; 用经过选择的多个模板样本代替单个模板样本, 降低姿态间的相互干扰; 在特征提取阶段, 采用对不同个体与姿态的全部人脸图像提取一致数目关键点的策略, 克服了 SIFT 算法关键点数目不确定的瓶颈, 同时提高算法对个体、姿态等变化以及遮挡的鲁棒性; 在特征匹配阶段, 采用基于模型和基于方向的双重策略剔除误匹配点对, 提高匹配评价值的可靠性; 用 256 位的二进制 ORB 描述子代替 128 维的浮点或整数数据类型的 SIFT 描述子, 提高特征提取与特征匹配的速度。

为了验证本文方法, 我们在 CAS-PEAL-R1 和 XJTU 数据库上进行了实验: 比较了改进 ORB 算子与 ORB 算子的性能; 本文每位个体多个模板方法与文献[7-8]每位个体单个模板方法的识别效果; ORB(改进 ORB)与 SIFT 算子的速度。

1 ORB 算子及改进

ORB 是 Rublee 等^[10]在 2011 年提出的一种图像特征提取算法, 该算法首先建立尺度空间图像金字塔, 然后对尺度空间图像使用 FAST 角点检测子定位关键点, 最后在尺度空间图像提取不同尺度的关键点特征。步骤如下:

Step1. 尺度空间极值检测。使用 FAST 角点检测子在所有的尺度和图像位置搜索关键点, 并且用 Harris 角点滤波器除去沿着边缘的关键点。这些关键点对于平面尺度缩放具有鲁棒性。

Step2. 关键点方向确定。使用以关键点为中心的图像块的灰度重心为关键点确定方向。图像块的力矩定义为

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y);$$

图像块的重心 C 的坐标为

$$\mathbf{C} = \begin{pmatrix} \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \end{pmatrix};$$

从图像块中心 \mathbf{O} 指向重心 \mathbf{C} 的矢量 \mathbf{OC} 的方向

$$\theta = \arctan 2(m_{01}, m_{10})$$

即为关键点的方向。Step3 中对关键点特征的描述都相对于该主方向, 从而保证了特征对于旋转变化具有鲁棒性。

Step3. 关键点特征描述。在以关键点为中心的图像块内比较采样点对的灰度值, 得到一个 n 位二进制数, 该 n 位二进制数即为关键点的特征描述子, n 的典型值为 256。

ORB 采用 31×31 像素块内的测试点对信息描述关键点, 每个测试点是一个 5×5 的子窗口。考虑一个平滑的图像块 \mathbf{p} , 一个二进制测试 τ 定义为

$$\tau(\mathbf{p}; \mathbf{x}, \mathbf{y}) := \begin{cases} 1: \mathbf{p}(\mathbf{x}) < \mathbf{p}(\mathbf{y}) \\ 0: \mathbf{p}(\mathbf{x}) \geq \mathbf{p}(\mathbf{y}), \end{cases}$$

其中 $\mathbf{p}(\mathbf{x})$ 是点 \mathbf{x} 处 \mathbf{p} 的灰度值。特征定义为一个包含 n 个二进制测试的矢量

$$\mathbf{f}_n(\mathbf{p}) := \sum_{1 \leq i \leq n} 2^{i-1} \tau(\mathbf{p}; \mathbf{x}_i, \mathbf{y}_i).$$

为了获得对旋转变化的鲁棒性, 需要对特征描述子定向。对于位置 $(\mathbf{x}_i, \mathbf{y}_i)$ 处的包含 n 个二进制测试的任何特征集合, 定义 $2 \times n$ 矩阵

$$\mathbf{S} = \begin{pmatrix} \mathbf{x}_1, \dots, \mathbf{x}_n \\ \mathbf{y}_1, \dots, \mathbf{y}_n \end{pmatrix},$$

令 \mathbf{R}_θ 为与关键点方向 θ 对应的旋转矩阵, 与方向 θ 对应的特征点对矩阵为

$$\mathbf{S}_\theta = \mathbf{R}_\theta \mathbf{S},$$

定向后的特征为

$$\mathbf{g}_n(\mathbf{p}, \theta) := \mathbf{f}_n(\mathbf{p}) | (\mathbf{x}_i, \mathbf{y}_i) \in \mathbf{S}_\theta.$$

为了提高特征的判别性能, ORB 采用学习算法从全部可能的二进制测试中选择了一个子集, 该子集的每一位特征既有高方差(均值接近 0.5), 又不互相关联。算法步骤如下:

Step1. 使用 PASCAL 2006 集合中的图像建立一个大约 300 k 个关键点的训练集合。枚举训练块的全部可能的二进制测试, 对全部训练块运行每个测试。

Step2. 根据均值与 0.5 的距离对测试排序, 形成矢量 \mathbf{T} 。

Step3. 贪婪搜索:

Step3.1. 将第一个测试放进结果矢量 \mathbf{R} 中并将它从 \mathbf{T} 中移去。

Step3.2. 从 \mathbf{T} 中取下一个测试, 并将它与 \mathbf{R} 中

的全部测试比较。如果它的绝对关联度大于一个阈值, 则删除它; 否则, 将它加入 \mathbf{R} 。

Step3.3. 重复 Step3.2 直到 \mathbf{R} 中有 256 个测试。如果测试数目少于 256, 则提高阈值并再次尝试。

与现有特征算子相比, ORB 算子引入了特征描述子的寻优过程, 算子的性能得到加强。然而不难发现, ORB 算子的寻优以整个图像块为基础, 远近距离测试的机会均等。考虑到近距离测试具有更好的局部性质, 本文提出基于分块模式的改进的 ORB 算子: 图像块的尺寸取为 32×32 像素块, 仿照 SIFT 描述子方法将每一个 32×32 像素块分成 8×8 子块, 对每一 8×8 子块枚举可能的二进制测试。全部子块的二进制测试构成寻优过程中候选的二进制测试集合。

采用分块模式的优势如下: 首先, 多姿态人脸识别中, 人脸的姿态变化主要不是平面内的旋转, 而是人脸曲面与摄像头之间的视角变化。由于分块模式中的子块所对应的人脸区域更容易用一个平面模型描述, 与整个图像块模式相比, 分块模式更能适应视角的变化, 因此这一改进在多姿态人脸识别中具有更为重要的意义; 其次, 与整个图像块模式相比, 分块模式的测试点对被限制在各个子块之中, 自动排除了子块之外的远距离点对, 具有更好的局部性质。

2 基于改进 ORB 特征的人脸识别

对于多视图方法采集的多姿态人脸图像, 文献[7-8]采用由多姿态人脸图像构建表征个体的单个模板样本的方法。文献[7]采用图像拼接的方法由多姿态人脸图像构建个体的单个模板样本, 文献[8]则将多姿态人脸图像关联程度高的 SIFT 特征集合作为个体人脸的单个模板样本特征。在识别阶段, 测试人脸特征必须与单个模板所包含的来自不同姿态的所有特征进行比较。实际上, 当姿态变化超过一定角度时, 同一个体姿态间的类内距离将大于不同个体的类间距离。为避免大角度姿态变化对识别结果的干扰, 本文采用为每位个体配置多个模板样本的方法, 从多姿态图像中选择几个不同视角的图像作为模板, 每个模板覆盖一定范围的视角扰动, 同时采用测试样本与单个模板样本的匹配评价值衡量测试人脸与模板人脸的吻合程度。

基于关键点特征的人脸识别需要匹配测试人脸

与模板人脸的关键点特征。SIFT 算法中，关键点的检测是基于图像的边缘响应和阈值参数，因此关键点的数目随着个体、姿态等的变化而变化。为克服 SIFT 算法关键点数目不确定的瓶颈，本文采用对图像提取一致数目关键点的策略，首先通过将阈值设置得足够小获得足够数量的关键点，然后根据 Harris 角点测量值对关键点排序并挑选测量值大的前 N 个点。提取一致数目关键点有助于减少由于个体、姿态等变化所引起的关键点数目的变化对识别结果的干扰，同时还提高了算法对遮挡的鲁棒性。具体的人脸识别算法如下：

训练阶段

Step1. 假设模板样本集合为 $\Omega = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,n_1}\} \cup \{\omega_{2,1}, \omega_{2,2}, \dots, \omega_{2,n_2}\} \cup \dots \cup \{\omega_{j,1}, \omega_{j,2}, \dots, \omega_{j,n_j}\} \cup \dots \cup \{\omega_{m,1}, \omega_{m,2}, \dots, \omega_{m,n_m}\}$ ，其中第 j 类(个体)分为 n_j 个子类(姿态)，本文中 $n_1 = n_2 = \dots = n_m$ ，且每位个体的每个姿态只有一个模板样本。

Step2. 根据本节介绍的方法提取每张模板人脸的改进 ORB 特征。

决策阶段

Step1. 根据本节介绍的方法提取测试人脸的改进 ORB 特征。

Step2. 分别匹配测试人脸的改进 ORB 特征与各类型姿态模板人脸的改进 ORB 特征。由改进 ORB 特征的匹配结果得出测试人脸与各模板人脸的匹配评价值。

Step3. 在同类模板样本集合中，将其中最好的评价值作为该类的评价值。

Step4. 比较不同类的评价值，确定测试人脸属于模板样本集合中的哪一类。

3 实验评价

3.1 人脸数据库和实验设置

为验证本文方法的效果，我们使用 CAS-PEAL-R1 数据库^[13-14]和 XJTU 人脸库^[15]进行实验。CAS-PEAL-R1 数据库包含有人脸图像的表情、饰物、光照、时间、背景、距离、仰视、平视和俯视子集，数据库中的全部图像已经被转化为灰度图像，并且裁剪到 360×480 ，去掉了大部分背景。本文选择的是平视子集，其中包括 1 040 幅人脸的 7 种姿态图像，姿态类别与人脸图像例子如表 1 所示。XJTU 数据库中的图像为彩色图像，实验中已经被转化为灰度图像，图像尺寸为 500×480 ；数据库包含有人脸图像的光照、视点子集。本文选择的是视点(平视)子集，其中包括 1 247 幅人脸的 19 种姿态

图像，姿态类别与灰度人脸图像例子如表 2 所示。在人脸识别实验中对数据库的图像进行裁剪，去除了人脸之外的其余部分，允许裁剪后的图像尺寸随个体及姿态的变化而变化。

考虑到人脸的对称性，在数据库中均选择包含正面图像的左右对称的多姿态样本构建多模板样本库。为了降低姿态变化对识别结果的影响，尽量使模板样本的视角均匀分布。CAS-PEAL-R1 数据库中样本的视角分别为 $0^\circ, \pm 22^\circ, \pm 45^\circ$ 和 $\pm 67^\circ$ ，因此每位个体用 0° 及 $\pm 45^\circ$ 的样本作为模板样本， $\pm 22^\circ$ 及 $\pm 67^\circ$ 的样本作为测试样本；实验中测试样本与邻近模板样本的最大视角间隔为 23° 。XJTU 数据库中样本的视角分别为 $0^\circ, \pm 10^\circ, \pm 20^\circ, \pm 30^\circ, \pm 40^\circ, \pm 50^\circ, \pm 60^\circ, \pm 70^\circ, \pm 80^\circ$ 和 $\pm 90^\circ$ ，因此每位个体用 $0^\circ, \pm 50^\circ$ 以及 $\pm 90^\circ$ 的样本作为模板样本， $\pm 10^\circ, \pm 20^\circ, \pm 30^\circ, \pm 40^\circ, \pm 60^\circ, \pm 70^\circ, \pm 80^\circ$ 的样本作

表 1 CAS-PEAL-R1 数据库的人脸例子

姿态角度/(°)	视图
-67	
-45	
-22	
0	
22	
45	
67	

表 2 XJTU 数据库的人脸例子

姿态角度/(°)	视图	姿态角度/(°)	视图	姿态角度/(°)	视图
-90		-20		50	
-80		-10		60	
-70		0		70	
-60		10		80	
-50		20		90	
-40		30			
-30		40			

为测试样本; 实验中测试样本与邻近模板样本的最大视角间隔为 20°。对于这 2 个数据库, 每次均随机选取其中的 50 位个体的样本进行实验, 实验结果为 20 次实验的平均值。

3.2 匹配评价值

本文中, 测试人脸与模板人脸的吻合程度取决于 2 幅图像中 ORB 关键点的特征匹配情况。特征匹配实际上要解决特征向量的最近邻搜索问题, 即在 D 维的度量空间 \mathcal{B}^D 中, 给定一个包含 n 个点的数据集合 $\mathcal{S} \subset \mathcal{B}^D$, 对于一个给定的查询点 q , $q \in \mathcal{B}^D$, 从 \mathcal{S} 中寻找与 q 的度量距离最近的点。这一过程不可避免地产生误匹配点对。剔除误匹配点对有多种方法, 例如最近次近距离比方法、基于模型的方法等, 采用单一的准则难以取得理想的

效果。考虑到人脸识别中测试人脸与模板人脸的取向基本一致, 本文分别根据匹配点对的模型信息与匹配点对连线的方向信息剔除误匹配点对: 首先采用基于模型的 RANSAC 算法剔除误匹配点对, 然后进一步利用匹配点对连线的方向信息剔除连线方向异常的误匹配点对。

RANSAC 算法通过多次随机抽取一定量的样本对模型参数进行估计, 每次随机抽取多个数据确定模型参数。再根据估计参数对所有数据分类, 一部分数据在允许的误差范围之内, 称之为内点; 反之则为外点。

令图像间的投影变换模型为

$$\mathbf{A}' = \mathbf{H}\mathbf{A},$$

其中 $\mathbf{A}' = [x' \ y' \ 1]^T$, $\mathbf{A} = [x \ y \ 1]^T$,

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & 1 \end{bmatrix},$$

\mathbf{A}' 和 \mathbf{A} 分别是两幅图像上匹配点对的齐次坐标, \mathbf{H} 为两图像间的投影变换矩阵.

匹配点对双向距离函数为

$$d_2(\mathbf{H}, \mathbf{A}, \mathbf{A}') = d_1(\mathbf{A}, \mathbf{H}^{-1}\mathbf{A}) + d_1(\mathbf{A}', \mathbf{H}\mathbf{A}),$$

模型误差函数为

$$e_M(\mathbf{H}, \{\mathbf{A}, \mathbf{A}'\}) = \sum_i [d_1(\mathbf{A}_i, \mathbf{H}^{-1}\mathbf{A}'_i) + d_1(\mathbf{A}'_i, \mathbf{H}\mathbf{A}_i)],$$

其中 d_1 为两点之间的单向距离函数, i 为参与计算的匹配点对序号.

算法步骤如下:

Step1. 设置迭代次数 N , 阈值 T , 记第 i 次迭代的内点数为 N_i .

Step2. 从 I 对匹配点对中随机选取 4 对初始匹配点对, 计算两平面间投影变换矩阵 \mathbf{H} 的参数.

Step3. 按匹配点对双向距离函数 $d_2(\mathbf{H}, \mathbf{A}, \mathbf{A}')$ 计算其余特征点经过变换矩阵的坐标值与它的匹配点之间的距离. 若 $d_2(\mathbf{H}, \mathbf{A}, \mathbf{A}')$ 小于阈值 T , 则认为该匹配点对为内点; 否则为外点. 更新 N_i 值.

Step4. 用新的内点集合去重新估计模型, 并按模型误差函数 $e_M(\mathbf{H}, \{\mathbf{A}, \mathbf{A}'\})$ 计算与模型对应的误差.

Step5. 如果本次所得的内点数 N_i 大于之前的内点数, 则将当前的模型保存为最优模型; 如果本次所得的内点数 N_i 等于之前的内点数, 则将误差较小的模型保存为最优模型; 否则, 转 Step2 继续下一轮迭代.

经过指定的迭代次数后, 在去除图像间误匹配点对的同时又估算出了图像间的投影变换矩阵参数. 采用 RANSAC 算法初步剔除偏离模型的误匹配点对之后, 进一步利用方向信息剔除连线方向异常的误匹配点对.

令 \mathbf{G} 为匹配点对连线单位方向向量构成的矩阵,

$$\mathbf{G} = [\mathbf{m}^1 \quad \mathbf{m}^2 \quad \cdots \quad \mathbf{m}^N]^T,$$

其中 \mathbf{m}^n 为第 n 个方向向量, $n = \{1, 2, \dots, N\}$. 方向向量 \mathbf{m}^j 与 \mathbf{m}^i 的相似性用内积度量, 即

$$d^{ji} = \langle \mathbf{m}^j, \mathbf{m}^i \rangle.$$

如果 d^{ji} 大于某一阈值 T , 则 2 个匹配点对连线的取向一致. 将第 j 个方向向量 $\mathbf{m}_{(1 \times 2)}^j$ 与 $\mathbf{G}_{(2 \times N)}^T$ 的乘积记为

$$\mathbf{D}_{(1 \times N)} = \mathbf{m}^j \mathbf{G}^T.$$

矩阵 \mathbf{D} 的元素为方向向量 $\mathbf{m}_{(1 \times 2)}^j$ 与 \mathbf{G} 中各个

方向向量的相似度. 因此对 \mathbf{D} 中的元素从高到低排序并计算大于某一阈值的元素的个数 $w^j = \sum_{i=1}^N (d^{ji} > T) - 1$, 可以得到方向向量 $\mathbf{m}_{(1 \times 2)}^j$ 与全体方向向量的吻合程度, 由此得出与 \mathbf{G} 对应的矩阵

$$\mathbf{w} = [w^1 \quad w^2 \quad \cdots \quad w^N]^T.$$

若 w^j 小于某一阈值 T , 说明所对应的匹配点对的连线方向偏离大部分匹配点对的连线方向, 该匹配点对被当作误匹配点对而被剔除.

通过如上 2 个步骤, 匹配点对被划分为最终的内点(群内匹配点对)和外点(离群匹配点对)两大类. 图 1 所示为 cch5 与 cch6 这 2 幅图像的关键点匹配及图像拼接结果, 其中 cch5(正面)与 cch6(-22° 侧面)来自 CAS-PEAL-R1 数据库平视子集, 并且经过裁剪去除了人脸之外的其余部分, cch5 的尺寸为 214×236 像素, cch6 的尺寸为 220×244 像素. 可以看出, 经过 RANSAC 以及基于方向信息的两级算法, 原始匹配点对中的误匹配点对被有效剔除. 图 1d 所示为最终的内点对应的投影变换模型得到的图像拼接结果, 2 幅图像的重叠区域采用平均算法融合. 由于误匹配点对被有效剔除, 投影变换模型误差较小, 尽管重叠的脸部区域并非一个严格的平面, 从图 1d 可以看出, 2 幅图像的重叠区域仍然取得了较好的融合效果.

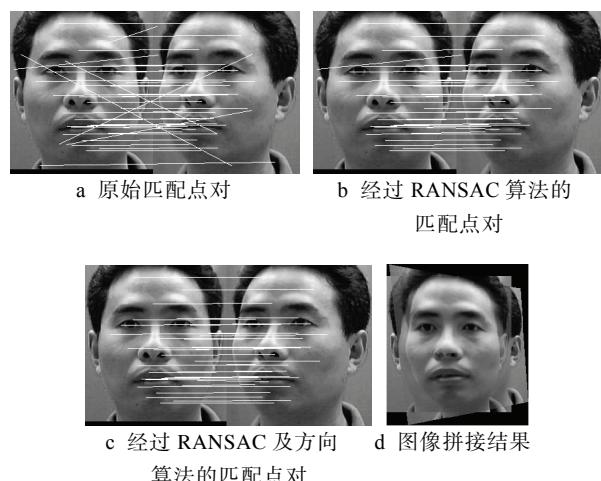


图 1 关键点匹配及图像拼接

考虑到 2 幅图像的匹配程度不仅与内点数目有关, 还与匹配点对的距离有关, 因此测试人脸与模板人脸的吻合程度用 2 个匹配评价值分层次评价. 首先采用内点数目作为第一层次匹配评价值; 对于内点数目相同的情况, 则进一步采用内点的

平均距离作为第二层次匹配评价值.

3.3 实验结果与分析

为了评价本文方法, 我们进行了 3 个实验.

实验 1. 验证本文方法的识别效果并比较改进 ORB 算子与 ORB 算子的性能. 实验结果如表 3 所示, 可以看出, 无论是改进 ORB 算子还是 ORB 算子, 对 2 个数据库都具有良好的识别效果, 平均正确识别率均高于 97.5%, 反映了本实验对多模板样本的视角设置是有效的, 能较好地克服姿态的干扰. 对于 2 个数据库, 改进 ORB 算子的正确识别率分别高于 ORB 算子 0.48% 和 0.36%, 显示改进 ORB 算子能更好适应姿态的变化, 具有更好的判别性能.

表 3 改进 ORB 与 ORB 对 2 个数据库的正确识别率 %

数据库	ORB	改进 ORB
CAS-PEAL-R1	97.62	98.10
XJTU	98.23	98.59

实验 2. 不同方法的比较实验. 实验在 2 个数据库上同时进行. 对于 CAS-PEAL-R1 数据库, 不同方法均由 0° 和 $\pm 45^\circ$ 的训练样本生成模板样本. 由于过多的训练样本反而使文献[7]方法性能下降, 因此对于 XJTU 数据库, 文献[7]方法仍然由最佳

组合的 0° 和 $\pm 50^\circ$ 训练样本生成模板样本, 而文献[8]与本文方法由 0° , $\pm 50^\circ$ 和 $\pm 90^\circ$ 的训练样本生成模板样本, 实验结果分别如表 4, 5 所示. 对于 CAS-PEAL-R1 数据库, 测试样本的最大视角只有 $\pm 67^\circ$. 由表 4 可以看出, 由于视角范围相对较小, 3 种方法均取得较稳定的识别结果, 本文方法的平均正确识别率高于文献[7]方法 0.74%, 高于文献[8]方法 1.56%. 对于 XJTU 数据库, 测试样本的最大视角为 $\pm 80^\circ$, 视角范围大. 由表 5 可以看出, 当测试样本视角大于 $\pm 60^\circ$ 时, 文献[7]方法的正确识别率急剧下降. 尽管文献[8]方法能较好地适应大视角的测试样本, 然而对于视角小于 $\pm 60^\circ$ 的测试样本, 识别效果并不具备优势. 本文方法在测试样本的整个视角范围内均取得了较稳定的较好识别效果, 对于各个视角的测试样本, 正确识别率均大于 98%.

表 4 3 种方法对 CAS-PEAL-R1 的正确识别率 %

	测试视角			平均
	$\pm 22^\circ$	$\pm 67^\circ$		
文献[7]方法	97.55	97.17	97.36	
文献[8]方法	96.80	96.28	96.54	
本文方法	98.55	97.64	98.10	

表 5 3 种方法对 XJTU 的正确识别率

	测试视角							平均
	$\pm 10^\circ$	$\pm 20^\circ$	$\pm 30^\circ$	$\pm 40^\circ$	$\pm 60^\circ$	$\pm 70^\circ$	$\pm 80^\circ$	
文献[7]方法	98.24	97.91	97.50	97.13	96.88	82.07	43.46	87.60
文献[8]方法	97.78	95.56	96.05	97.34	96.89	93.78	95.83	96.18
本文方法	99.25	98.71	98.33	98.73	98.52	98.19	98.37	98.59

实验 3. SIFT 与 ORB(改进 ORB)算子的速度比较. 实验在 Laptop 上进行, CPU 为 Core i5 1.7GHz, 内存为 4.00GB; 操作系统为 Windows 7; 编程环境为 Microsoft Visual Studio 2010 与 Intel OpenCV2.3.1. 表 6 所示为 cch6 的关键点检测与描述耗时结果, 表 7 所示为 cch6 与 cch5 的匹配耗时结果, 其中的数据为 20 次运行的平均值. 可以看出, ORB(改进 ORB)算子的关键点检测及描述耗时仅为 SIFT 算子的 9.03%(26/288), 匹配耗时仅为 SIFT 算子的 34.48%(10/29), 检测、描述及匹配总耗时仅为 SIFT 算子的 11.36%(36/317). ORB 特征描述子为 256 位二进制数, 由关键点邻域的二进制测试矢量形成. 在特征匹配时, 特征向量的最近邻搜索以汉明距

离作为指标, 用 2 个二值向量异或运算, 然后统计结果向量中 “1” 的个数的算法实现. 如果使用新式 CPUs 的 SSE 指令集提供的执行异或操作及统计

表 6 cch6(220×244 像素)的关键点检测与描述耗时

	关键点数目	检测与描述耗时/ms
SIFT	50	288
ORB(改进 ORB)	50	26

表 7 cch6 与 cch5(214×236 像素)的匹配耗时

	关键点数目		匹配耗时/ms
	cch5	cch6	
SIFT	50	50	29
ORB(改进 ORB)	50	50	10

二值向量中“1”的个数的特定指令，还能够高速地计算汉明距离^[16]，进一步减少匹配耗时。SIFT 特征描述子为 128 维特征向量，由关键点邻域图像子块的梯度直方图形成，可以转化为整数数据类型。在特征匹配时，用欧氏距离等指标度量匹配点对的距离远近。与 SIFT 算子相比，无论是关键点检测与描述阶段还是特征匹配阶段，ORB 算子均具有很好的效率。

在存储空间方面，一个 SIFT 特征向量需要 128 字节存储空间，而一个 ORB 特征向量仅需要 32 字节存储空间，为 SIFT 特征的 25%。

4 结 论

本文提出基于改进的 ORB 局部特征以及每位个体多个模板样本的多姿态人脸识别方法。首先用分块采样模式改进原 ORB 算子的整图像块采样模式，提高算子对人脸视角变化的鲁棒性，并为个体设置多个模板样本克服姿态的干扰。在特征提取阶段，为克服 SIFT 算法关键点数目不确定的瓶颈，减少关键点数目变化对识别结果的干扰，同时提高算法对遮挡的鲁棒性，采用对关键点按 Harris 角点测量值排序并提取固定数目关键点的策略。在特征匹配阶段，采用基于模型和基于方向的双重策略剔除误匹配点对，提高群内匹配点对的可靠程度，并使用群内匹配点对数目与匹配点对平均距离评价测试样本与模板样本的吻合程度。对 CAS-PEAL-R1 和 XJTU 数据库的实验结果表明，对于±67°的视角变化，每位个体用 0°、±45°的样本作为模板样本即可取得很好的识别效果；对于±90°的视角变化，每位个体用 0°、±50°、±90°的样本作为模板样本即可取得很好的识别效果，即对于与模板样本之间的视角变化在±23°范围内的测试样本，本文方法能够得到很好的正确识别率。与采用多个训练样本构建个体单个模板样本的方法相比，本文的多模板方法即使对于±90°的视角变化，仍然能够有效减少姿态变化对识别结果的干扰，具有较好的识别效果。ORB 不仅比 SIFT 算子具有明显的速度优势，而且所需要的存储空间也仅为后者的 25%。需要指出的是，尽管本文方法允许脸部姿态存在±90°的视角变化，但是并不允许脸部存在过大的变形(例如闭眼)，过大的变形将使关键点的相关性下降，降低算法的性能甚至使算法失去识别能力。如果允许脸部存在过大的变形，

则需要增加额外的模板样本。本文方法属于人脸辨识(Identification)系统，将来的工作包括基于本文方法的人脸确认(Verification)系统的研究。

致谢

本文是作者在美国罗切斯特理工大学计算机工程系访学期间的工作。感谢 Andreas Savakis 教授对本文的指导！感谢 Raymond Ptucha 助理教授对本文有益的建议和意见！

参考文献(References):

- [1] Abate A F, Nappi M, Riccio D, et al. 2D and 3D face recognition: a survey [J]. Pattern Recognition Letters, 2007, 28(14): 1885-1906
- [2] Chai X J, Shan S G, Chen X L, et al. Locally linear regression for pose-invariant face recognition [J]. IEEE Transactions on Image Processing, 2007, 16(7): 1716-1725
- [3] Prince S J D, Elder J H, Warrell J, et al. Tied factor analysis for face recognition across large pose differences [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(6): 970-984
- [4] Sharma A, Dubey A, Jagannatha A N, et al. Pose invariant face recognition based on hybrid-global linear regression [J]. Neural Computing and Applications, 2010, 19(8): 1227-1235
- [5] Zhu Changren, Wang Runsheng. Multi-pose face recognition based on a single view [J]. Chinese Journal of Computers, 2003, 26(1): 104-109 (in Chinese)
(朱长仁, 王润生. 基于单视图的多姿态人脸识别算法[J]. 计算机学报, 2003, 26(1): 104-109)
- [6] Liu Zhijing, Li Xiazhong, Wu Mang. Pose-varied face recognition based on the 3D model [J]. Journal of Xidian University: Natural Science, 2004, 31(2): 218-222 (in Chinese)
(刘志镜, 李夏忠, 武芒. 基于三维模型的多姿态人脸识别[J]. 西安电子科技大学学报: 自然科学版, 2004, 31(2): 218-222)
- [7] Singh R, Vatsa M, Ross A, et al. A mosaicing scheme for pose-invariant face recognition [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2007, 37(5): 1212-1225
- [8] Xian X B, Wu H J, Zhang M X, et al. A novel multi-pose face recognition via robust SIFT feature [C] // Proceedings of International Conference on Wavelet Analysis and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2013: 32-37
- [9] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110
- [10] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF [C] // Proceedings of IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2011: 2564-2571

- [11] Bicego M, Lagorio A, Grosso E, *et al.* On the use of SIFT features for face authentication [C] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2006: 35-41
- [12] Majumdar A, Ward R K. Discriminative SIFT features for face recognition [C] // Proceedings of Canadian Conference on Electrical and Computer Engineering. Los Alamitos: IEEE Computer Society Press, 2009: 27-30
- [13] Gao W, Cao B, Shan S G, *et al.* The CAS-PEAL large-scale Chinese face database and baseline evaluations [J]. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, 2008, 38(1): 149-161
- [14] Zhang Xiaohua, Shan Shiguang, Cao Bo, *et al.* CAS-PEAL: a large-scale Chinese face database and some primary evaluations [J]. Journal of Computer-Aided Design & Computer Graphics, 2005, 17(1): 9-17 (in Chinese)
(张晓华, 山世光, 曹波, 等. CAS-PEAL 大规模中国人脸图像数据库及其基本评测介绍[J]. 计算机辅助设计与图形学报, 2005, 17(1): 9-17)
- [15] The Institute of Artificial Intelligence and Robotics of Xi'an Jiaotong University. The oriental face database [OL]. [2013-11-27]. <http://www.aiar.xjtu.edu.cn/dfrlsjk3.htm> (in Chinese)
(西安交通大学人工智能与机器人研究所. 东方人脸数据库 [OL]. [2013-11-27]. <http://www.aiar.xjtu.edu.cn/dfrlsjk3.htm>)
- [16] Calonder M, Lepetit V, Ozuysal M, *et al.* BRIEF: computing a local binary descriptor very fast [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1281-1298