



基于个性化最小二乘法的软特征集成实验方法

张 擎, 张洋洋*

(山东大学艺术学院, 济南 250100)

摘要: 在生物特征识别技术的研究中, 软特征的集成实验一直采用主特征之间的集成实验模型。软特征在适用范围、识别能力等方面与主特征有很大差别, 直接使用主特征之间的集成机制无法有效地挖掘和利用软特征中的区分性信息, 造成实验结果不够理想和准确。基于此, 该文深入分析软特征的特性, 提出“有效的互补性”和“个性化集成”两点集成要求, 并在量化有效的互补性基础上, 结合“最小错误率”目标, 利用最小二乘法为每一个用户建立局部集成模型。将集成模型用于人脸识别和指纹识别两个实验场景, 验证了该文对问题分析的合理性以及所提方法在提高识别准确率方面的有效性。

关键词: 软特征; 集成实验; 生物特征识别; 个性化; 最小二乘法

中图分类号: TP391

文献标志码: A

DOI: 10.12179/1672-4550.20220630

Fusion Experimental Method of Soft Biometrics Based on Personalized Least Squares

ZHANG Qing, ZHANG Yangyang*

(Fine Arts School of Shandong University, Jinan 250100, China)

Abstract: In the study of biometrics recognition, the fusion experiment of soft biometrics has always adopted the fusion model between the main features. Soft biometrics are different from the main features in terms of application scope, recognition ability, etc. The fusion mechanism between the main features cannot effectively mine and utilize the discrimination information in the soft biometrics, resulting in less ideal and accurate experimental results. Based on this, this paper deeply analyzes the characteristics of soft biometrics, and puts forward the requirements of ‘effective complementarity’ and ‘personalized fusion’. Based on the quantified effective complementarity, and combined with the goal of ‘minimum error rate’, this paper establishes a local fusion model for each user using the least squares method. The proposed fusion model is applied in face and fingerprint recognition experimental scenarios to verify the rationality of the analysis and the effectiveness of the proposed method in promoting the recognition performance.

Key words: soft biometrics; fusion experiment; biometrics recognition; personalized; least squares

1 问题的提出

随着信息技术的发展, 层出不穷的网络诈骗、电信诈骗给社会和人民财产安全带来巨大挑战。生物特征识别技术利用一种或多种人类的生理特征(如指纹、人脸、静脉等)或行为特征(如步态、签名等)进行身份识别^[1], 因具有不会忘记、不会丢失、安全性高等诸多优势, 得到了国内外

学术界和企业界的高度关注, 也成为各高校实验教学 and 科学研究的重要内容。

通过研究和实验, 人们发现利用两种或两种以上生理或者行为特征, 或利用同一生物特征的多个不同方面(多台设备上采取的指纹)进行识别, 可以克服单一生物特征识别的固有缺陷和局限性, 获得更高的识别性能^[2]。在这一研究分支中, 文献 [3] 提出将某些辅助信息, 如性别、民

收稿日期: 2022-11-03; 修回日期: 2023-10-18

基金项目: 国家自然科学基金(62176139)。

作者简介: 张擎(1982-), 女, 博士, 高级实验师, 主要从事计算机应用、实验室安全技术、生物特征识别等方面的研究。

* 通信作者: 张洋洋(1989-), 女, 硕士, 实验师, 主要从事数据分析、实验室安全技术等方面的研究。E-mail: zhangyangyang@sdu.edu.cn

族、身高与主生物特征进行集成, 进一步提高生物特征识别系统的性能, 这些辅助信息被称为“软特征”。目前, 针对集成软特征的验证实验已有很多, 主要集中在验证软特征提取的可行性和利用软特征提高性能的有效性等几个方面^[4-6]。软特征被多次证明比利用多生物特征识别具有更好的灵活性和方便性, 可以同时提高系统的可用性和识别准确率。

在软特征(如性别、年龄)与主特征(如指纹、人脸)之间的集成机制上, 则主要借鉴主特征之间的集成机制, 通常采用简单的线性模型, 模型参数大多利用交叉验证等实验性或经验性方法确定。目前尚未见到在充分分析软特征特性的基础上, 有针对性地设计软特征集成实验方法的工作。基于软特征的独有特性, 软特征与主特征之间的集成问题与主特征之间的集成问题具有完全不同的特点, 直接使用主特征之间的集成机制无法有效地挖掘和利用软特征中的区分性信息, 造成软特征集成实验的结果不够理想和准确。因此, 针对软特征的独有特性, 设计适用于软特征的集成模型进行实验是必要的。

2 问题分析

不同于主特征, 软特征区分性较差, 不足以单独用来进行身份识别。这就意味着软特征具有高噪声的特点。因此, 不是所有对软特征的集成都能获得提高识别性能的实验结果, 甚至有些集成会影响主特征的准确识别。以“性别”特征为例, 我们在公开库 FaceMLA 中选择 3 个符合情况的用户人脸图像进行说明, 如图 1 所示, 在人脸识别中, 假设用户 A 和 B 都是男性, 但属于不同个体, 即使通过主特征“人脸”可以很好地区分 A 和 B, 但如果识别中扩大“性别”特征的作用, 则会加大 A 和 B 属于同一个体的证据, 这里, “性别”显然就是噪声信息。那么“性别”特征将如何利于识别呢? 同样在图 1 中, 如果 A 是男性而 C 是女性, 在通过主特征“人脸”认定 A 和 C 比较相似的情况下, “性别”信息将起到区分 A 和 C 的关键辅助作用。

通过上面的例子可以看到, 实验中, 对软特征的集成并不是在区分所有用户时都能发挥作用, 且一定要与主特征有一定的互补性才能有助于识别性能的提高。但具有互补性, 并不是软特

征能起到辅助识别作用的充分条件。上述例子中, 在区分用户 A 和用户 B 时, 主特征与软特征也具有互补性, 但此时软特征就是噪声信息。我们分析, 只有在主特征区分性弱, 并且软特征与主特征在识别中具有互补性的情况下, 软特征才能起到辅助识别的作用。我们可以称这时软特征与主特征之间的互补性为“有效的互补性”。此外, 软特征在区分不同用户时往往有不同表现。因此, 需要针对每个用户设计不同的集成模型进行实验。已有文献中将这类针对每个用户设计集成模型的方法叫作“依赖于用户”的方法或“个性化集成”实验方法^[7-8]。

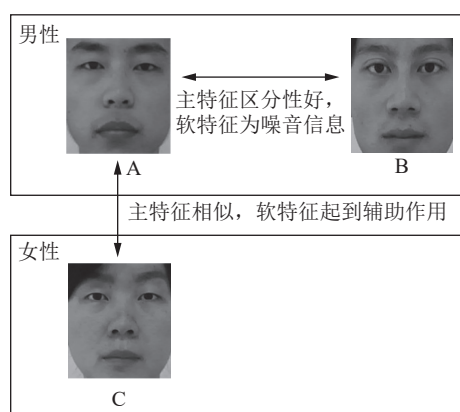


图 1 人脸识别中“性别”特征的作用

总之, 软特征集成实验模型的设计需要满足“有效的互补性”和“个性化集成”两点要求。

3 基于个性化最小二乘的集成实验方法

3.1 量化“有效的互补性”

要满足“有效的互补性”和“个性化集成”两个要求: 首先要对互补性进行量化, 判断其有效性; 其次, 要针对每个系统用户设计个性化实验模型。因此, 本文将两个目标结合起来, 分别利用属于每个用户的样本, 量化软特征与主特征的互补性。

量化软、主特征之间的互补性, 首先要分别量化软特征与主特征在识别个体上的区分性。不失一般性, 假设系统中有 N 个注册用户, 且利用用户的某一个主特征 p (如人脸) 以及与 p 相关的某些软特征(如性别、雀斑、皱纹)进行识别。针对用户 $u(u=1,2,\dots,N)$, 如果可以分别得到其主特征 p 和其某个软特征 s 在区分各个用户方面的区分性, 分别记为 d_p^u 和 d_s^u 。在此基础上, 就可以通过计

算 d_p^u 和 d_s^u 的差值获得软特征 s 与主特征的互补性,但有效的互补性还要将主特征的区分性考虑在内。因此,用户 u 的软特征 s 与主特征的有效互补性可计算得:

$$\delta_s^u = \frac{|d_p^u - d_s^u|}{d_p^u} \quad (1)$$

这里,软特征与主特征的互补性越大且主特征的区分性越小,有效的互补性越大,符合我们对问题的分析。

软特征和主特征的区分性,可以借鉴文献[9]中的“区分性判定准则”,基于样本的类内聚合度和类间离散度进行计算。

对于用户 $u(u=1,2,\dots,N)$,首先介绍软特征区分性 d_s^u 的计算方法。不失一般性,设用户 u 的某一软特征 s 有多个注册模板,记模板集合为 T_s^u 。基于此,计算用户 u 的软特征 s 在区分 u 与其他用户时的区分度,记为 d_s^{uv} ($v=1,2,\dots,N,v \neq u$)。然后可以简单地通过求平均的方式得到 d_s^u 。具体方法如下:

将 T_s^u 中模板两两之间相似度归为一类,用 ω_1 表示。将 T_s^u 中模板与 T_s^v ($v=1,2,\dots,N,v \neq u$)中模板的相似度归为一类,用 ω_2 表示。基于此,计算 ω_1 的类内聚合度及 ω_1 和 ω_2 的类间离散度分别为:

$$\sigma_W^2 = \frac{1}{|\omega_1|} \sum_{x \in \omega_1} (x - m_1)^2 \quad (2)$$

$$\sigma_B^2 = \frac{1}{2} \sum_{i=1}^2 (m_i - m)^2 \quad (3)$$

式中: m_1 和 m_2 分别代表 ω_1 和 ω_2 两类元素的均值, m 是 m_1 和 m_2 的均值。

类间离散度 σ_B^2 和类内聚合度 σ_W^2 的比值为:

$$d_s^{uv} = \frac{\sigma_B^2}{\sigma_W^2} \quad (4)$$

得到 d_s^{uv} 后,可通过求平均的方式得到:

$$d_s^u = \frac{\sum_{v=1, v \neq u}^N (d_s^{uv})}{N-1} \quad (5)$$

d_s^u 越大,表示针对用户 u ,软特征 s 的区分度越高。同理,可以得到主特征 p 对于个体 u 的区分度 d_p^u 。

由于 d_p^u 和 d_s^u 均大于0,且一般主特征区分性大于软特征,也就是 $d_p^u > d_s^u$,结合式(1)可见,有效

的互补性 $\delta_s^u \in (0,1)$ 。

3.2 基于最小二乘法的个性化集成实验模型

获得软特征与主特征“有效的互补性”这一量化指标后,不失一般性,可以从基于转换的模型^[10-11]出发对实验模型进行设计。基于转换的模型是将各特征匹配得分转换到统一区间,再通过加权相加的方式进行集成的一类实验模型。在这类集成实验模型中,加权权重的设置是核心问题。以往工作大都采用经验或交叉验证等实验手段确定软、主特征的集成权重。为将“有效的互补性”运用到集成模型中去,且兼顾“个性化”集成模型的要求,本文基于最小二乘法(least squared method, LSM),在最小化样本错误的目标下,结合“有效的互补性”,为每个用户分别确定各模态的集成权重。实验模型的形式化表达如下:

假设共有 M 个模态(1个主特征, $M-1$ 个软特征),针对每个用户,利用注册模板计算得到 G 个匹配得分向量及其标记,记为 $S = \{(s_1, y_1), (s_2, y_2), \dots, (s_G, y_G)\}$,其中,第 i ($i=1,2,\dots,G$)个匹配得分向量 $s_i = (s_{i1}, s_{i2}, \dots, s_{iM})^T$ 是在所有模态上获得的匹配得分组成的向量,标量 y 为标记。如果为同源匹配得分(两个来自同一用户的特征进行匹配所得的得分),则 y 为1;否则, y 为0。在软特征集成实验模型中,以最小化样本匹配错误为目标,LSM可以通过优化下式得到用于集成 M 个模态的权重向量 $w = (w_1, w_2, \dots, w_M)^T$ 。

$$\operatorname{argmin}_w \sum_{i=1}^G \left(\sum_{j=1}^M w_j s_{ij} - y_i \right)^2 \quad (6)$$

$$\text{s.t.} \quad \sum_{j=1}^M w_j = 1$$

$$w_1 > w_j, w_1 - w_j > \beta_j, (j=2,3,\dots,M)$$

式中: w_1 为主特征的集成权重, w_2, w_3, \dots, w_M 为软特征的集成权重。由于主特征的区分性大于软特征,一般要求主特征权重大于软特征权重。 $\beta_2, \beta_3, \dots, \beta_M$ 作为调节各软特征和主特征重要性的参数,至关重要:其值越大,主特征的权重越大,而相应的软特征所起作用越小;其值越小,相应的软特征的权重越大,作用越明显。根据分析,对于某用户 u 的某一软特征 j ($j=2,3,\dots,M$),有效的互补性 δ_j^u 越大,特征 j 的权重应越大。这

里, 为表达统一, 不再标记用户 u 。可见, 可以利用 δ_j 确定 β_j , 且 δ_j 和 β_j 应是负相关关系。通过分析, $\delta_j \in (0, 1)$, 由式 (6) 中的约束, 同样可以得到 $\beta_j \in (0, 1)$, 且有研究指出, 主特征的权重应该远大于软特征^[3]。由此, 这里将 β_j 的范围控制在 0~0.5 之间。综上, 需要找到合适的函数, 将 δ_j 由负相关关系函数映射到 β_j , 也就是, 找到映射函数, 将 (0,1) 之间的数以负相关关系映射到 (0,0.5) 之间。考虑函数的连续可导等性质, 我们采用下面函数进行映射:

$$\beta_j = \frac{1}{1 + e^{5\delta_j}} \quad (7)$$

式(7)是 Sigmoid 函数的变形:

$$y = \frac{1}{1 + e^{-x}} \quad (8)$$

其具有与 Sigmoid 函数同样的连续可导性质, 将 (0,1) 之间的数负相关映射到 (0.0067, 0.5) 区间, 符合映射要求。如图 2 所示, 给出了式 (7) 与 Sigmoid 函数在 $[-3, 3]$ 区间内的函数曲线。

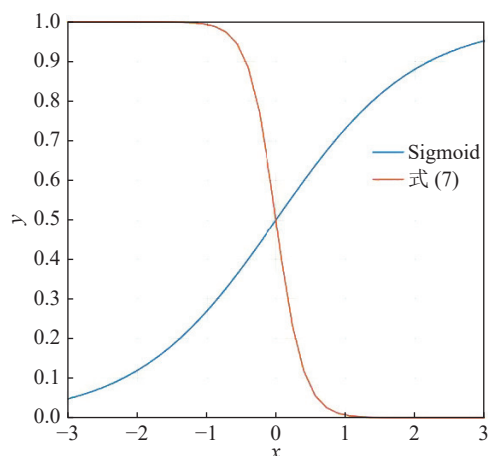


图2 Sigmoid 函数与本文使用映射函数的曲线图

4 应用案例及实验结果

4.1 集成软特征的人脸识别实验

文献 [9] 中分析, 在个体相貌变化不大的情况下, 异源相似关系 (来源于不同个体人脸的相似度) 是稳定的。我们在公开库 FacePix 中选择 3 个符合情况的用户人脸图像进行说明, 如图 3 所示, 直观上, 由于脸型不同, 个体 A 和个体 B 的人脸相似程度较低; 而由于都是瓜子型脸, 个体 A 与个体 C 人脸的相似程度会高一些。在这 3 个

个体相貌变化不大的情况下, 这种相似性关系是稳定存在的。因此, 异源相似度信息是值得挖掘和利用的具有区分能力的信息。文献 [9] 利用稀疏相似度序列对异源相似度信息进行挖掘并与主特征集成, 用于识别, 同时, 通过 Lasso 稀疏方法确定异源相似度的权重, 获得了较好的识别性能。异源相似度信息本质上属于软特征的范畴。

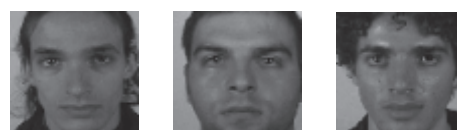


图3 异源个体人脸图像

这里, 我们将每一个异源相似度作为一个软特征, 利用本文提出的集成模型对其进行集成。本文实验在 FacePix 和 FaceMLA 两个公开人脸库上进行。在 FacePix 人脸库中包含从不同角度采集的 30 个不同个体的人脸图像。实验中用每个个体的最接近正面人脸的 5 幅图像作为模板样本, 其他 15 幅图像作为测试样本。FaceMLA 人脸数据库中包含 32 个不同个体的人脸图像。实验中取其中 30 个个体的 20 幅正面人脸图像进行实验, 其中 5 幅正面图像作为模板样本, 其他 15 幅正面图像作为测试样本。

对异源相似度这一软特征的提取需要一个基础的人脸识别方法实现, 文献 [9] 称之为“基方法”。这里将本文方法与文献 [9] 中方法在 4 个基方法的基础上进行对比。这 4 个基方法是较为常用的 4 个经典人脸识别方法: PCA (Eigenface) 方法^[12]、LDA (Fisherface) 方法^[13]、ICA 方法^[14] 和 LLE 方法^[15]。

首先验证本文对软、主特征之间“有效的互补性”分析的合理性和所提方法的有效性。分别在 FacePix 和 FaceMLA 两个公开人脸库上随机选择两个用户进行研究, 记为“用户 1”和“用户 2”。不失一般性, 选择最常用的两个人脸识别方法 PCA 和 LDA 作为基方法进行验证。因为相似度序列中的软特征数量较多, 只随机选取其中部分进行验证。这里为每个用户随机选择 5 个软特征, 分别记为 1~5 号软特征。为其中每一个软特征计算其与主特征之间的有效的互补性 δ , 随后只集成该软特征和主特征 (利用本文集成方法) 进行识别, 验证识别性能。利用最常用的衡量指标

等错误率(equal error rate, EER)衡量系统的识别性能。表 1 和表 2 分别给出两个人脸库上的实验结果。

表 1 FacePix 数据库上, 不同软特征的 δ 和集成主特征后的性能

用户	软特征	PCA		LDA	
		δ	集成主特征的识别性能 EER/%	δ	集成主特征的识别性能 EER/%
用户1	1号	0.78	0.114	0.75	0.078
	2号	0.85	0.105	0.81	0.072
	3号	0.67	0.135	0.64	0.086
	4号	0.86	0.107	0.90	0.064
	5号	0.73	0.125	0.80	0.074
用户2	1号	0.66	0.135	0.71	0.083
	2号	0.75	0.117	0.81	0.075
	3号	0.78	0.112	0.85	0.072
	4号	0.86	0.101	0.90	0.066
	5号	0.69	0.128	0.70	0.083

表 2 FaceMLA 数据库上, 不同软特征的 δ 和集成主特征后的性能

用户	软特征	PCA		LDA	
		δ	集成主特征的识别性能 EER/%	δ	集成主特征的识别性能 EER/%
用户1	1号	0.71	0.028	0.68	0.041
	2号	0.64	0.036	0.65	0.041
	3号	0.85	0.022	0.88	0.033
	4号	0.69	0.031	0.87	0.035
	5号	0.75	0.025	0.70	0.038
用户2	1号	0.75	0.023	0.77	0.034
	2号	0.85	0.019	0.86	0.035
	3号	0.64	0.038	0.68	0.042
	4号	0.91	0.016	0.85	0.035
	5号	0.79	0.022	0.70	0.037

表 1 和表 2 中的结果显示软特征和主特征之间有效的互补性 δ 越大, 集成该软特征和主特征能的识别性能越高的整体趋势。这有力地证明了本文对软特征与主特征之间“有效的互补性”分析的合理性, 以及本文所提的量化方法和集成方法的有效性。

为进一步验证本文方法的合理性以及集成模型的有效性, 实验中将本文提出的集成方法与文献 [9] 中的集成方法进行了对比, 同时给出不集成任何软特征的方法(主特征)的识别性能以及只使用软特征进行识别的识别性能。为方便进行实验, 在只利用软特征进行识别时, 利用平均加权

的集成方式将所有软特征的匹配得分进行集成, 用于识别。实验中就等错误率 EER、零误识率 (false accept rate, FAR) 前提下的拒识率 (false reject rate, FRR), 以及零拒识率 FRR 前提下的误识率 FAR 几个关键指标进行对比, 对比结果如表 3~表 5 所示。

表 3 基于不同的基方法, 各对比方法的 EER

数据库	基方法	EER/%			
		本文集成方法	文献[9]集成方法	不集成软特征的方法	只使用软特征的方法
FacePix	PCA	0.08	0.10	0.14	11.65
	LDA	0.04	0.05	0.09	13.33
	ICA	0.09	0.12	0.17	14.58
	LLE	0.08	0.11	0.19	14.32
	PCA	0.01	0.01	0.04	7.86
Face-MLA	LDA	0.02	0.04	0.05	7.43
	ICA	0.02	0.03	0.04	8.72
	LLE	0.01	0.03	0.08	9.55

表 4 基于不同的基方法, 各对比方法的 FAR(FRR=0%)

数据库	基方法	FAR/%			
		本文集成方法	文献[9]集成方法	不集成软特征的方法	只使用软特征的方法
FacePix	PCA	0.31	0.37	0.39	83.45
	LDA	0.68	0.83	0.98	79.92
	ICA	0.40	0.49	0.53	84.95
	LLE	0.25	0.35	0.47	88.96
	PCA	0.02	0.02	0.05	71.75
Face-MLA	LDA	0.05	0.08	0.09	67.83
	ICA	0.04	0.05	0.08	73.81
	LLE	0.05	0.09	0.16	78.85

表 5 基于不同的基方法, 各对比方法的 FRR(FAR=0%)

数据库	基方法	FRR/%			
		本文集成方法	文献[9]集成方法	不集成软特征的方法	只使用软特征的方法
FacePix	PCA	0.18	0.25	0.43	59.50
	LDA	0.24	0.35	0.65	65.97
	ICA	0.21	0.29	0.42	68.89
	LLE	0.31	0.39	0.46	66.92
	PCA	0.02	0.03	0.11	57.85
Face-MLA	LDA	0.05	0.09	0.29	60.89
	ICA	0.04	0.05	0.17	66.91
	LLE	0.07	0.10	0.21	51.88

结果显示, 软特征有一定的区分能力, 但只使用软特征进行识别, 识别性能较差, 不能满足一般的应用需求。主特征的识别性能优于软特

征, 与软特征之间具有互补性。两个集成了软特征的方法在几个关键指标上均优于未集成软特征的方法, 可见对软特征进行集成的有效性。本文集成方法在两个数据库上, 均获得了优于文献 [9] 中集成方法的识别性能。由于算法的特殊要求, 两个数据库上人脸均趋近于正面人脸, 因此, 在特定情况下, 两个方法的 EER 均达到了 0.01%。本文方法体现出更高的识别性能是因为本文在分析软特征特点的基础上, 根据不同个体的不同软特征设计集成权重, 而文献 [9] 中方法统一将软特征权重进行稀疏化处理, 使一些本具有较好辅助作用的软特征的权重也在稀疏化过程中减弱, 甚至被置为零。

4.2 集成软特征的指纹识别实验

文献 [16] 中定义了指纹上的“未匹配细节点特征”, 并将其与主特征进行集成, 验证了其在提高指纹识别准确率上的有效性。本质上, 未匹配细节点特征是一种软特征。图 4 给出两对指纹的匹配过程中出现的匹配细节点和未匹配细节点, 其中图 4(a) 为异源匹配(属于不同个体的指纹之间的匹配), 图 4(b) 为同源匹配(属于同一个体的指纹之间的匹配)。红色点为匹配细节点, 蓝色点为未匹配细节点。文献 [16] 从“相对数量”“位置对应关系”和“全局分布一致性”3 个角度提取不同的软特征参与识别, 采用线性集成模型进行实验, 并利用交叉验证的方法确定集成参数。这种集成实验方法是目前软特征集成实验中较为常用的方法, 具有一定的代表性。这里我们用本文提出的集成模型将文献 [16] 提取的软特征与主特征进行集成。为方便表述, 下面实验中记“相对数量”软特征为软特征 I、“位置对应关系”软特征为软特征 II、“全局分布一致性”软特征为软特征 III。实验在 3 个国际指纹认证竞赛库上进行验证, 分别是 FVC2000、FVC2002 和 FVC2004。每一个竞赛库包含 DB1、DB2、DB3 和 DB4 共 4 个子库, 实验在所有 12 个子库上进行。每个子库包含采集自 100 个手指的指纹, 其中每个手指有 8 幅指纹。实验中, 将数据库中的每个手指视为一个用户, 每个子库中合计有来自 100 个用户的 800 幅指纹, 针对每个用户随机选择 4 幅指纹作为模板指纹, 其他 4 幅指纹用作测试指纹。

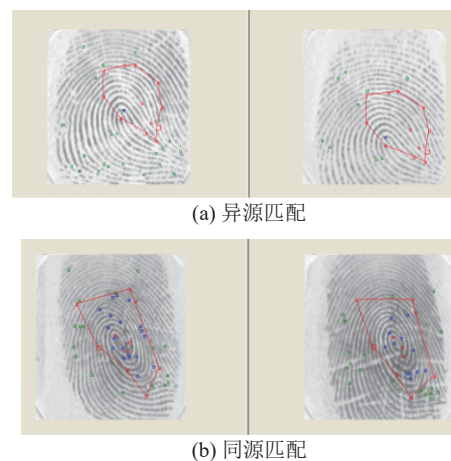


图 4 指纹匹配中的匹配细节点和未匹配细节点

同样, 首先验证本文对软、主特征之间“有效的互补性”分析的合理性和所提方法的有效性。不失一般性, 在 3 个竞赛库的 DB1 和 DB2 子库上分别随机选择两个用户, 记为用户 1 和用户 2。针对每个用户, 为每一个软特征计算其与主特征之间的有效的互补性 δ , 随后只集成该软特征和主特征(利用本文集成方法), 验证识别性能。这里, 同样用等错误率 EER 衡量系统的识别性能。3 个竞赛库上的实验结果如表 6~表 8 所示。

表 6 FVC2000 数据库上, 不同软特征的 δ 和集成主特征后的性能

用户	软特征	DB1		DB2	
		δ	集成主特征的识别性能 EER/%	δ	集成主特征的识别性能 EER/%
用户 1	软特征 I	0.74	2.68	0.71	0.75
	软特征 II	0.82	1.83	0.68	0.79
	软特征 III	0.79	2.57	0.83	0.68
用户 2	软特征 I	0.68	1.89	0.72	0.77
	软特征 II	0.83	1.77	0.67	0.83
	软特征 III	0.65	2.34	0.80	0.66

表 7 FVC2002 数据库上, 不同软特征的 δ 和集成主特征后的性能

用户	软特征	DB1		DB2	
		δ	集成主特征的识别性能 EER/%	δ	集成主特征的识别性能 EER/%
用户 1	软特征 I	0.67	1.21	0.73	1.21
	软特征 II	0.81	0.98	0.84	1.03
	软特征 III	0.79	1.19	0.76	1.16
用户 2	软特征 I	0.64	1.27	0.72	1.07
	软特征 II	0.83	0.95	0.87	0.95
	软特征 III	0.81	1.02	0.68	1.28

表 8 FVC2004 数据库上, 不同软特征的 δ 和集成主特征后的性能

用户	软特征	DB1		DB2	
		δ	集成主特征的识别性能 EER/%	δ	集成主特征的识别性能 EER/%
用户1	软特征I	0.68	2.87	0.78	3.25
	软特征II	0.89	1.94	0.85	2.88
	软特征III	0.75	2.64	0.71	3.40
用户2	软特征I	0.68	2.58	0.74	3.15
	软特征II	0.84	2.21	0.80	3.05
	软特征III	0.76	2.46	0.69	3.65

从实验结果可以看出, 当软特征和主特征之间有效的互补性 δ 更大时, 集成该软特征和主特征能获得更好的识别性能。这与人脸识别中的实验结果相一致, 既验证了我们对“有效的互补性”分析的合理性, 也同时验证了本文量化互补性方法的有效性。对照表 9~表 11 中的结果, 可以看到用本文方法集成软特征后系统性能较单一使用主特征得到了提高, 这也验证了本文集成方法的有效性。

表 9 FVC2000 数据库上, 各对比方法的识别性能
单位: %

方法	评价指标	子库			
		DB1	DB2	DB3	DB4
本文集成方法	EER	1.33	0.52	2.01	0.41
	FAR(FRR=0%)	16.14	26.87	25.14	0.73
	FRR(FAR=0%)	3.94	0.40	7.05	1.97
文献[16]集成方法	EER	1.50	0.67	2.62	0.58
	FAR(FRR=0%)	19.07	30.18	27.84	0.83
	FRR(FAR=0%)	6.00	1.00	10.25	2.24
不集成软特征的方法	EER	3.38	1.01	6.51	1.32
	FAR(FRR=0%)	64.10	66.05	50.98	11.90
	FRR(FAR=0%)	18.00	5.25	21.75	14.42
软特征I	EER	6.28	3.04	8.81	4.12
	FAR(FRR=0%)	78.68	85.44	74.35	66.36
	FRR(FAR=0%)	43.35	34.56	45.17	36.87
软特征II	EER	5.92	4.45	10.93	7.01
	FAR(FRR=0%)	73.38	79.96	83.95	80.04
	FRR(FAR=0%)	44.68	44.65	56.70	73.44
软特征III	EER	5.09	7.07	16.44	7.00
	FAR(FRR=0%)	70.26	80.78	87.65	82.36
	FRR(FAR=0%)	52.79	67.73	57.94	79.60

为进一步验证本文所提方法的有效性, 将本文方法与文献 [16] 中的集成方法就 EER、零 FAR 前提下的 FRR, 以及零 FRR 前提下的 FAR 几个

关键指标进行对比。同时给出未集成软特征的方法(主特征)以及只使用某一种软特征进行识别的识别性能。实验结果如表 9~表 11 所示。

表 10 FVC2002 数据库上, 各对比方法的识别性能
单位: %

方法	评价指标	子库			
		DB1	DB2	DB3	DB4
本文方法	EER	0.38	0.54	0.56	0.30
	FAR(FRR=0%)	10.95	37.8	11.02	0.31
	FRR(FAR=0%)	1.26	1.30	2.20	0.84
文献[16]方法	EER	0.50	0.73	0.76	0.33
	FAR(FRR=0%)	14.21	43.67	14.11	0.57
	FRR(FAR=0%)	1.50	1.50	2.75	1.25
不集成软特征的方法	EER	1.47	1.38	3.31	1.37
	FAR(FRR=0%)	41.77	74.62	79.64	9.33
	FRR(FAR=0%)	4.25	4.00	18.00	6.00
软特征I	EER	4.49	2.46	4.53	4.12
	FAR(FRR=0%)	89.57	75.65	84.38	56.17
	FRR(FAR=0%)	44.79	52.67	68.54	69.31
软特征II	EER	5.92	4.45	10.93	7.01
	FAR(FRR=0%)	84.90	83.33	82.67	71.79
	FRR(FAR=0%)	57.88	68.47	88.98	74.14
软特征III	EER	4.57	2.25	16.05	7.00
	FAR(FRR=0%)	86.59	71.33	88.90	81.27
	FRR(FAR=0%)	53.13	56.35	67.85	63.54

表 11 FVC2004 数据库上, 各对比方法的识别性能
单位: %

方法	评价指标	子库			
		DB1	DB2	DB3	DB4
本文方法	EER	0.98	1.85	3.05	0.42
	FAR(FRR=0%)	14.95	34.31	38.70	1.15
	FRR(FAR=0%)	3.97	4.26	14.30	1.95
文献[16]方法	EER	1.32	2.20	3.64	0.62
	FAR(FRR=0%)	19.07	45.42	50.70	1.92
	FRR(FAR=0%)	6.00	7.00	16.75	2.25
不集成软特征的方法	EER	3.39	4.10	7.94	1.61
	FAR(FRR=0%)	64.10	68.99	68.48	23.37
	FRR(FAR=0%)	18.00	12.75	30.50	11.25
软特征I	EER	7.23	6.51	9.74	9.50
	FAR(FRR=0%)	84.33	89.35	84.35	76.33
	FRR(FAR=0%)	49.34	44.53	59.18	67.87
软特征II	EER	11.69	8.34	9.61	10.71
	FAR(FRR=0%)	89.96	83.25	88.90	90.33
	FRR(FAR=0%)	76.54	67.53	71.33	69.68
软特征III	EER	7.63	15.99	10.44	14.68
	FAR(FRR=0%)	83.25	92.33	87.77	91.35
	FRR(FAR=0%)	52.33	71.07	70.85	78.67

从表9~表11中可以看出, 软特征识别性能较弱, 与主特征具有一定的互补性。两个集成了软特征的方法在几个关键指标上均优于未集成软特征的方法, 可见软特征集成的有效性。另外, 本文的集成模型在实验中得到了明显优于文献[16]方法的性能, 可见, 本文在充分分析软特征特点的前提下, 利用合理方法设计集成模型, 优于传统的通过交叉验证等技术确定集成模型的方法。

5 结束语

在针对生物特征识别技术的研究和实验中, 人们发现对“软特征”的集成可以克服单一生物特征识别的固有缺陷和局限性, 获得更高的识别性能, 且比利用多生物特征进行识别具有更好的灵活性和方便性。目前, 在软特征与主特征的集成实验中, 主要借鉴主特征之间的集成机制, 还尚未有在深入分析软特征特性的基础上, 有针对性地提出软特征集成模型的工作。本文指出, 软特征具有其独有特性, 且与主特征在适用范围、识别性能等方面具有较大差异。直接使用主特征之间的集成机制无法有效地挖掘和利用软特征中的区分性信息, 造成软特征集成实验的结果不够理想和准确。基于此, 本文深入分析软特征的特性及软特征与主特征之间的关系, 提出并量化“有效的互补性”, 同时结合最小二乘法为每一个用户建立局部集成模型。本文提出的集成模型在人脸识别和指纹识别两个实验场景均获得了更好的识别性能。后续的研究中, 将进一步将“有效的互补性”与更多非线性集成模型进行结合应用, 并在更多的生物特征中验证实验方法的有效性。

参考文献

- [1] TRANQ N, TURNBULL B P, HU J. Biometrics privacy-preservation: How do they evolve?[J]. *IEEE Open Journal of the Computer Society*, 2021, 2(1): 179–191.
- [2] SHEKHARS, PATELV M, NASRABADI N M. Joint sparse representation for robust multimodal biometrics recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(1): 113–126.
- [3] PARK U, JAIN A K. Face matching and retrieval using soft biometrics[J]. *IEEE Transactions on Information Forensics and Security*, 2010, 5(3): 406–415.
- [4] YANG L, YANG G, YIN Y. Exploring soft biometric traits in finger vein recognition[J]. *Neurocomputing*, 2014, 135(8): 218–228.
- [5] OSORIO-ROIG D, RATHGEB C, DROZDOWSKI P. An attack on facial soft-biometric privacy enhancement[J]. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2022, 4(2): 263–275.
- [6] SAEED U. Micro-expressions as a soft biometric for person recognition[J]. *Pattern Recognition Letters*, 2021, 143: 95–103.
- [7] SADEDDINE K, CHELALI F Z, DJERADI R. Recognition of user-dependent and independent static hand gestures: Application to sign language[J]. *Journal of Visual Communication and Image Representation*, 2021, 79(C): 103193.
- [8] MHENNI A, CHERRIER E, ROSENBERGER C. User dependent template update for keystroke dynamics recognition[C]//2018 International Conference on Cyberworlds. Singapore: IEEE Press, 2018.
- [9] ZHANG Q, YIN Y, YANG G. Face recognition based on sparse similarity sequence[J]. *Journal of Computational Information Systems*, 2014, 10(22): 9629–9639.
- [10] ASSAAD F S, SERPEN G. Transformation based score fusion algorithm for multi-modal biometric user authentication through ensemble classification[J]. *Procedia Computer Science*, 2015, 61(1): 410–415.
- [11] BISOGNI C, NAPPI M. Multibiometric score-level fusion through optimization and training [C]//3rd International Conference on Bio-engineering for Smart Technologies. Paris: IEEE Press, 2019.
- [12] ZHOU C, WANG L, ZHANG Q. Face recognition based on PCA and logistic regression analysis[J]. *Optik*, 2014, 125(20): 5916–5919.
- [13] VINAY A, SHEKHAR V S, BALASUBRAMANYA K N. Performance study of LDA and KFA for Gabor based face recognition system[J]. *Procedia Computer Science*, 2015, 57(1): 960–969.
- [14] KUO S C, LIN C J, LIAO J R. 3D reconstruction and face recognition using kernel-based ICA and neural networks[J]. *Expert Systems with Applications*, 2011, 38(5): 5406–5415.
- [15] ZHAO S, ZHU S. Face recognition by LLE dimensionality reduction[C]//2011 International Conference on Intelligent Computation Technology and Automation. Shenzhen: IEEE Press, 2011.
- [16] ZHANG Q, YIN Y, YANG G. Unmatched minutiae: Useful information to boost fingerprint recognition[J]. *Neurocomputing*, 2016, 171(1): 1401–1413.