

基于信号质量动态加权的多模生物特征识别研究

张闻彬¹, 刘培顺¹, 薛峰会²

(1. 中国海洋大学信息科学与工程学院, 山东 青岛 266100;
2. 青岛黄海学院国际电子商务学院, 山东 青岛 266427)

摘要: 将人脸识别和说话人识别进行决策层级的融合, 为应对外界环境对识别结果的影响, 引入图像质量和声音质量评价方法, 通过对信息质量进行评估, 去除信息质量较差的特征, 根据信息质量动态调整模块的权重比例, 并对单模特征识别匹配度低的个体做拒绝处理, 然后根据 D-S 理论将各个证据合并成为一个新的证据体, 实现对用户身份识别。实验结果显示, 这种考虑特征信号质量的融合方法可以有效提升识别的准确率和安全性。

关键词: 人脸识别; 说话人识别; 特征信号质量评估; D-S 证据理论; 多生物特征识别融合

中图分类号: TP391.41

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2018022

Multi-biometrics fusion based on dynamic weighting of characteristic signal quality

ZHANG Wenbing¹, LIU Peishun¹, XUE Fenghui²

1. College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China
2. Institute of International Electronic Commerce, Qingdao Huanghai University, Qingdao 266427, China

Abstract: Face recognition and speaker recognition were integrated at the decision-making level. In order to cope with the influence of the external environment on the recognition result, image quality and sound quality assessment methods were introduced. By evaluating the quality of information, features with poor information quality were removed. Information quality dynamically adjusted the weight proportion of the module, and rejected individuals with low matching degree of single-mode feature recognition. Then, according to DS theory, each evidence was merged into a new body of evidence to realize user identification. The experimental results show that the fusion method which takes into account the characteristic signal quality can effectively improve the recognition accuracy and security.

Key words: face recognition, speaker recognition, characterization of signal quality, D-S evidence theory, multi-biometrics fusion

1 引言

在众多身份认证方法中, 基于生物特征的身份认证因其唯一性、稳定性和易用性受到越来越

多的关注, 并广泛应用于现实生活^[1]。现阶段, 较为成熟的生物特征识别主要包括人脸、掌纹、指纹、指静脉、虹膜、声纹等, 在特征信号质量较好的情况下具有很高的识别准确率。受外界环

收稿日期: 2017-12-17; 修回日期: 2018-01-25

通信作者: 刘培顺, liups@ouc.edu.cn

基金项目: 国家重点研发计划基金资助项目 (No.2017YFC0806200)

Foundation Item: The National Key R&D Plan Program of China (No.2017YFC0806200)

境影响,单一生物特征识别方式在性能上会出现不同程度的下降,一些生物特征可能会被复制或伪造,难以满足系统安全要求^[2,3]。文献[4]提出采用融合理论,通过整合多个生物特征属性提供的证据,形成多模生物特征识别方法,因其使用多个特征信息源而展现出诸多优点:增强了识别性能、降低了验证失败率和提高了可靠性^[5]。

为增强多模生物特征识别系统的顽健性,文献[6]给出一种基于图像质量的多生物特征融合系统,将生物特征图像质量和特征匹配值相关联,用贝叶斯网络实现融合识别;文献[7]将指纹和虹膜进行融合识别,提出了一种基于似然比的方法实现多生物特征系统中匹配分数的质量融合,实验得到了较好的识别结果;文献[8]提出基于虹膜和人脸图像质量的匹配分级融合策略,根据虹膜遮挡程度以及人脸表情、角度变化调整单模识别匹配器的权重,在 IV^2 多模数据库的实验中,等错误率相较于不考虑特征质量的融合方案降低 30%。

本文将人脸识别与说话人识别进行融合,主要基于以下原因:1) 人脸识别技术和说话人识别技术已相对成熟,在环境良好的情况下识别率理想,且识别过程中只需要有摄像头和麦克风就可以完成检测(目前智能手机都具备);2) 使用范围也非常广泛,无论在手机移动端或在电脑客户端,在复杂多变的自然条件下或较为稳定的室内环境中均能运行;3) 在安全方面,将人脸识别和说话人识别相结合,也可以提高抵御照片欺骗、录音播放等攻击的能力,增强系统的安全性。

目前针对说话人识别和人脸识别融合的研究相对较少,文献[9]从人脸图像亮度和声音信号信噪比两方面确定识别模式的加权比例,将2种识别方式的结果按照权值比例进行求和,融合方式较为直接,对于信号质量评价存在局限性,文献[10]的研究指出,可以在 D-S 证据理论的框架下进行说话人识别和人脸识别的融合,并用实验证明了该方法能有效提高识别的准确度。基于上述工作,为更好地应对生物特征的信号质量对识别效果的影响,本文引入人脸图像和声音信号质量评价方法,并将特征信号质量进行量化处理,结合 D-S

证据理论,提出一种基于特征信号质量动态加权的多模生物特征识别的认证技术,以提高身份认证系统的准确率和容错率。

2 单模生物特征识别技术

单模生物特征识别技术使用单一的生物特征对目标进行身份识别。用来进行身份识别的生物特征可分为两类:生理特征和行为特征。生理特征主要包括人脸、指纹、虹膜、声纹等生物特征,而行为特征则包括步态、字迹等。由于本文要实现的是人脸识别和说话人识别融合,下面简要介绍这2种识别技术。

2.1 人脸识别

人脸识别的研究兴起于20世纪60年代,经过几十年的发展,已经成为一项广泛应用于社会生活各个领域的生物特征识别技术。传统的人脸识别的方法主要有几何结构法^[11]、代数特征法^[12]等。尽管上述方法在人脸识别领域取得了较好的成果,但均难以应对复杂多变的外部环境。近年来,深度学习在人脸特征学习中表现出来的突出性能使其成为人脸识别领域的主流。2014年,Facebook^[13]提出了 DeepFace 人脸识别方法,该方法利用含有400万数据量的人脸数据库训练一个9层CNN,并将全连接层 FC7 作为人脸表达,在户外脸部检测数据库(LFW, labeled faced in the wild)中识别准确率达到97.25%。此后,汤晓鸥团队的 DeepID 系列人脸识别方法^[14-17]和谷歌的 FaceNet 人脸识别系统^[18]相继提出,在 LFW 上的准确率分别达到99.53%和99.63%。和传统特征提取方式不同,深度学习主要从大量的训练数据中自动学习特征的表示,构建一个具有一定深度的网络,网络中包含大量经逐层训练以后确定的参数,更能表征数据特征。

2.2 说话人识别

说话人识别技术根据采集到的声音信号实现对说话人的身份鉴定。根据识别任务的不同,说话人识别可分为说话人确认(SV, speaker verification)和说话人辨认(SI, speaker identification)。一直以来,人们将高斯混合模型(GMM, gaussian mixture model)作为说话人识别应用的主要方法^[19],通过构建高斯混合模型-通用背景模板(GMM-UBM)实现

对识别系统中说话人模型的建立。Kenny^[20]在此基础上提出联合因子分析（JFA），把说话人与信道的差异限制在 GMM 超矢量高维空间的 2 个子空间中，进而抽取只与说话人有关的特征以解决信道问题，但此时信道空间中仍然包含少量的说话人语音信息。基于上述理论，Dehak 在文献[21]中提出了总变因子技术，也就是在现阶段说话人识别领域使用较为广泛的 i-vector 方法，该方法是将说话人和信道 2 部分同时进行建模，得到一个总变化子空间（TVS, total variability space），然后应用 PLDA 等方法对 i-vector 空间进行信道补偿，最后计算待测身份的 i-vector 和目标模型的 i-vector 的余弦距离，并将其与阈值进行比较，得到最终的判决结果。

3 基于特征信号质量动态加权的多模生物特征识别方法

D-S 证据理论的关键是确定每条证据（即各生物特征识别方法）的可信程度。在生物特征识别过程中，生物信号的质量下降可能导致识别性能的下降^[22]。文献[23~26]研究表明，信号质量对身份识别结果有显著的影响。基于以上研究结果，根据特征信号质量，动态调整该识别方法在基于 D-S 证据理论多生物特征识别过程中的权重分配，会使结果更加准确。

本文采用的是一种对生物特征信号质量顽健性的多生物特征识别系统。该系统流程如图 1 所示。

3.1 图像质量评价

在人脸识别中，图像质量对于识别结果会产生较大影响^[27]。在融合过程中，为减少环境影响，对图像进行合理评价显得尤为重要。

常用的图像质量评价可以根据图像对比度、图像亮度、结构信息相似度等指标进行分析，根据文献[28]中对于人脸图像质量影响因素的研究，本文选择图像光照和清晰度这两方面对人脸图像进行无参考评价。

3.1.1 光照质量评价指标^[29]

存在光照失真的人脸图像往往有不均匀的质量分布，根据它们在图像空间中重要性的不同，对人脸图像不同区域赋予不同的权重。把图像 I 分成左右 2 个子图像，分别为 IL 和 IR，然后将 IR 进行水平翻转得到 IFR，从左上角开始计算 IFR 中 8×8 窗口的 LQI（luminance quality index）^[30]，在每一块图像的 LQI 计算完毕后，通过式(1)便可以得到光照质量评价指标 Q_L 。

$$Q_L = \frac{\sum_{i=1}^N m_i w_i}{\sum_{i=1}^N w_i} \quad (1)$$

其中， m_i 表示第 i 个块的光照质量值， w_i 表示对应第 i 个块的权值，所提出的相关性方法权值求法如式(2)和式(3)所示。

$$w_i = g(x_i, y_i) \quad (2)$$

$$g(x, y) = \sigma_x^2 + \sigma_y^2 + c \quad (3)$$

x_i 和 y_i 分别表示第 i 个块中的像素， σ_x 和 σ_y 为 X 和 Y 像素上的标准差，而 c 代表基线最小权重的常数。

图 2 总结了如何将该方法应用于给定的人脸图像计算其照明质量。 Q_L 的取值范围为[0,1]，当且仅当左右区域的质量相等时其值为 1。其中，每个区域被分成一组大小为 8×8 像素的区块。

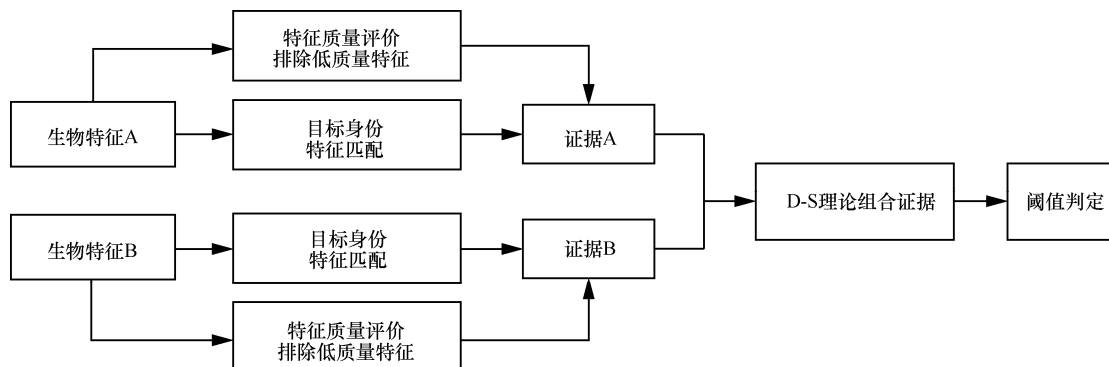


图 1 对生物特征信号质量顽健性的多生物特征识别系统流程

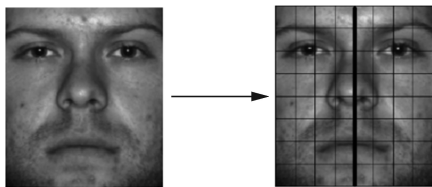


图 2 人脸图像被分成左右 2 个区域

3.1.2 图像清晰度评价指标

图像模糊的一般特点是边缘的模糊和细节的丢失,文献[31]将边缘宽度定义为在边缘点的梯度方向上离该点最近的 2 个局部极值之间的距离,并将边缘宽度定义为局部模糊。文献[32]基于此提出了感知模糊概率积累(CPBD, cumulative probability of blur detection)的清晰度估计算法,对于给定的局部对比度 C ,边缘点 e_i 的感知模糊概率 $P_{\text{blur}}(e_i)$ 建模如下。

$$P_{\text{blur}}(e_i) = 1 - \exp\left(-\left|\frac{w(e_i)}{w_{\text{JNB}}(e_i)}\right|\right) \quad (4)$$

其中, $w(e_i)$ 是边缘点 e_i 的宽度; $w_{\text{JNB}}(e_i)$ 是边缘点 e_i 的 JNB 阈值对应的边缘宽度, w_{JNB} 和 β 的取值均是通过将实验数据用最小二乘法拟合式(4)的指数模型得到的。当边缘点感知模糊概率 P_{blur} 小于等于阈值 P_{JNB} 时,这些边缘点不能被人眼检测到模糊,即为清晰的边缘点。当清晰边缘点占边缘点总数比例越高,图像就越清晰。CPBD 算法由各边缘感知模糊概率 P_{blur} 的概率密度分布函数得到最终的图像清晰度,无参考图像清晰度指标 Q_C 即为 CPBD 的值。

$$Q_C = \text{CPBD} = \sum_{P_{\text{blur}}=0}^{P_{\text{blur}}=P} P(P_{\text{blur}}) \quad (5)$$

3.1.3 人脸图像质量评价指标

根据人脸图像质量评价指标^[33],结合图片光照质量评价和清晰度量评价,定义人脸识别图像整体质量评价 Q_F 。

$$Q_F = m_1 Q_L + m_2 Q_C \quad (6)$$

图 3 是人脸图像质量由好到差的一组样本,除去个体差异,主要因为光照条件和人脸图像采集装置的不同等原因,导致存在质量较差的样本。其中,图 3(a)人脸图像光照条件良好、图像较为清晰,图 3(b)人脸图像虽然光照情况较好,但图像清晰度不佳,图 3(c)和图 3(d)人脸图像光照条件和图片清晰度均不理想。

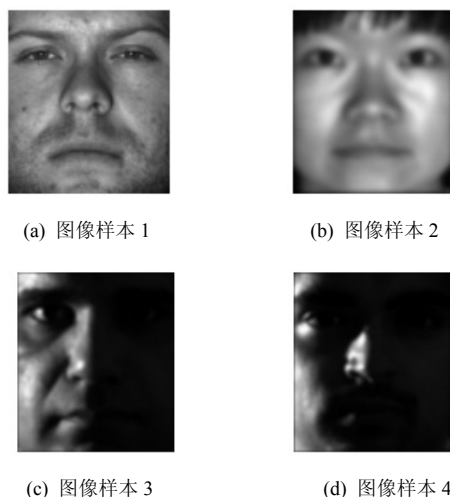


图 3 不同质量水平的人脸图像样本

根据主观评价和计算结果,赋予 m_1 、 m_2 分布为 0.5、0.5,上述 4 张图片评价结果如表 1 所示,所得结果反映了人脸图像的光照水平和清晰程度,符合人眼对图像质量水平的主观感受。

表 1 质量指标评估结果

图像序号	质量指标
图 3(a)	0.79
图 3(b)	0.58
图 3(c)	0.30
图 3(d)	0.18

3.2 声音质量评价

目前,说话人识别最主要的挑战来自处理测试语料的缺失以及获取语音时环境和信道的差异问题^[34]。需要在决策融合之前,将说话人的声音信息进行质量评价,提高系统对声音特征的顽健性。

本文采用主观语音质量评估(PESQ, perceptual evaluation of speech quality)作为评价方法,主要基于 2 个原因:1) PESQ 是国际电联(ITU)提出的语音质量评价算法之一,采用改进的听觉模型和认知模型技术,能够更好地反映通信延迟和环境噪音对信道的影响,与其他算法相比,PESQ 与主观评价结果相关度最高;2) BLATNIK 等^[35]将 PESQ 应用于自动说话人识别系统(ASRS, automated speaker recognition system)中,结果显示了 PESQ 得分和 ASRS 错误率的相关性,指出 PESQ 可以有效地用于 ASRS 识别性能的预测。

PESQ 算法客观评价得分是平均对称干扰度

d_{SYM} 和平均非对称干扰度 d_{ASYM} 的线性组合^[36]。

$$PESQ = 4.5 - 0.1d_{\text{SYM}} - 0.0309d_{\text{ASYM}} \quad (7)$$

对于正规的主观测试, PESQ 得分在 1.0~4.5 之间。在失真情况严重时, 得分可能会低于 1.0。图 4 显示的是 3 段语音的波形图, 其中, 图 4(a) 是纯净语音, 图 4(b) 在图 4(a) 上加入了信噪比为 20 dB 的高斯白噪音, 图 4(c) 在图 4(a) 上加入了信噪比为 5 dB 的粉红噪音, 噪音取自噪声库 NoiseX-92。

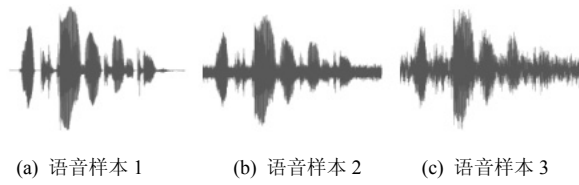


图 4 3 段语音的时域波形

经过测试, 图 4(a)、图 4(b)、图 4(c) 得分为 4.50、2.83、1.67, 结果反映出随着声音质量的下降, 得分逐渐降低。最后将得到的 PESQ 得分和 4.50 相比, 得到的便是说话人识别声音质量评价 Q_s 。

$$Q_s = \frac{PESQ}{4.50} \quad (8)$$

3.3 基于特征信息质量动态加权的融合方法

有关生物特征识别融合方式主要包含以下

3 类^[4]: 1) 特征提取级的融合; 2) 匹配分数级的融合; 3) 决策级的融合。由于人脸识别与说话人识别是 2 个独立的、互不影响的识别系统, 在特征层级并无关联。加上光照强度、噪音等因素的影响, 单模生物特征识别在识别系统中的结果可能会存在冲突, 影响识别结果, 因此, 如何能够表达和处理不确定性成为融合的关键。而 D-S 证据理论通过定义信任函数和似然函数把命题的不确定性问题转化为集合的不确定问题, 解决因信息不完备导致的不确定, 这为人脸识别和说话人识别的融合提供了理论支持。

在 D-S 证据理论应用于多模生物特征识别过程中, 人脸识别模块与说话人识别模块会提供给融合系统各自的证据, 通过对信息质量进行评估, 动态调整各个识别模块的权重比例, 在这一过程中要检测到信息质量较差的特征, 提示用户重新输入该特征信号; 对单模特征识别匹配度均较低的个体, 说明用户身份识别结果为否, 要做拒绝处理, 然后生成每条证据的基本概率分配函数, 对所有的证据赋予一个可信度, 并利用 D-S 理论的合并规则将各个证据合并成为一个新的证据体。图 5 为该融合方法的完整流程。

具体实现如算法 1 所示。

算法 1 基于特征信息质量动态加权的 D-S 证据融合识别算法

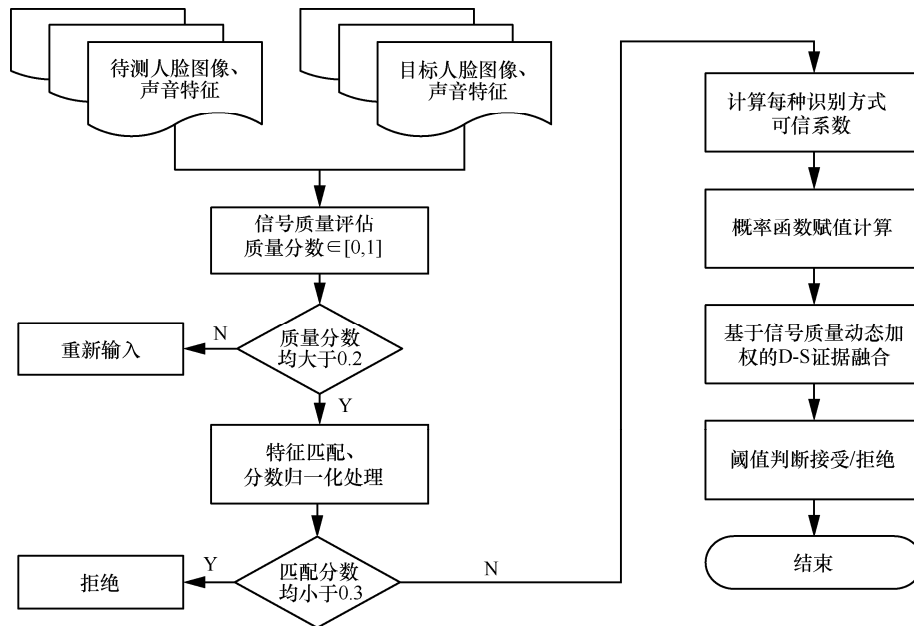


图 5 基于特征信息质量动态加权的 D-S 证据融合识别流程

输入 待测身份和目标身份的人脸图像、声音特征

输出 识别认证结果

Step1 对输入的特征进行合理评估, 由式(6)、式(8)得到人脸识别图像整体质量评价 Q_F 、说话人识别声音质量评价 Q_S , 此时如果质量评价取值低于 0.2, 该特征即为无效特征, 需要重新输入。

Step2 人脸识别使用基于 FaceNet 的深度学习方法进行识别, 说话人识别使用基于 i-vector 的方法进行识别, 得到相应匹配分数, 利用 Tanh 方法^[37]对所得结果进行归一化处理, 使匹配分数在区间[0,1]中。当 2 种识别方式的匹配分数均低于 0.3 时, 识别系统直接判定身份不匹配。

α_i 是使用第 i 种识别方法对第 j 个待测身份进行识别的匹配分数。

$$\alpha_i = S_i(j), i=1,2 \quad (9)$$

Step3 根据特征匹配分数与特征质量进行概率分配函数构造。 R_i 为使用第 i 种识别方法进行身份识别的可信系数^[38]。

$$R_i = \frac{\alpha_i w_i}{\sum_i \alpha_i w_i} \quad (10)$$

其中, w_i 为第 i 中识别方式占整个识别系统的权重, 根据待测身份特征信号质量 Q_i 进行动态调整, 相应计算方式为

$$w_i = \frac{Q_i}{\sum_i Q_i} \quad (11)$$

K 为识别方式的个数, 使用第 i 种识别方式对 j 个待测身份进行识别的基本概率函数为

$$m_i(j) = \frac{\alpha_i}{\alpha_i + K(1-R_i)(1-\alpha_i w_i)} \quad (12)$$

Step4 证据合并, 通过式(12)可以得到人脸识别和说话人识别所对应的概率赋值函数 $m_1(A)$ 和 $m_2(B)$, 进而计算焦元 C 的信任函数。

$$m(C) = \frac{\sum_{A \cap B = C} m_1(A) m_2(B)}{1 - \sum_{A \cap B = \emptyset} m_1(A) m_2(B)} \quad (13)$$

最后用阈值法进行判断, 这里取等错误率 EER 的阈值, 如果 $m(C) > \varepsilon$, 则身份认证通过, 否则认证失败。

Step5 重复 Step1~Step4, 直到完成全部输入的识别认证。

算法应用实例, 如图 6 所示, 从左至右依次为目标身份 1、待测身份 1、待测身份 2、待测身份 3、待测身份 4 的双模特征对, 其中, 待测身份 1 的人脸图像和声音特征、待测身份 2 的声音特征、待测身份 3 的人脸图像来自于目标身份。

根据算法 1 给出的方法, 将图 6 的 4 个待测身份代入, 得到表 2 的结果。

待测身份 1 的相关生物特征均来自目标身份, 虽然人脸图像特征信息质量较差, 但经过融合后依然能获得较高的目标身份概率; 待测身份 2、待测身份 3 分别采用照片欺骗和录音欺骗仿造目标身份, 最后融合结果的概率均低于 0.5, 攻击失败; 而待测身份 4 由于一开始 2 种识别方式匹配分数较低, 直接遭到系统拒绝。由此可以看出该融合识别算法具有较强的防御单一生物特征欺骗的能力。

4 实验结果与分析

由于现在尚无公开的多生物特征库, 在融合实验阶段选择将 Extended Yale B 人脸库和 TIMIT 语料库进行组合配对。Extend Yale B 人脸库共包含 38 个人的 2 432 幅人脸图像, 每个人有 10 种姿态、60 多种光照情况, 适合进行融合系统的人脸识别测试。从 TIMIT 语料库中选取 38 个人的语音资料, 与上述人脸库组成人脸-语音双模特征对。由于 TIMIT 语料库都是在纯净环境下录制的, 为了模拟真实情况, 在每人的 5 段语音中随机加入信噪比为 5 dB、10 dB、20 dB、40 dB 的白噪音、粉红噪音、工厂车间噪音, 其余 592 人的语音进行说话人通用背景模板 (UBM) 和全变换子空间矩阵 T 的训练。

在单一生物特征识别方法选择方面, 人脸识别是基于 FaceNet 的深度学习方法实现, 通过训练 MS-Celeb-1M 人脸库得到人脸模型, 在 LWF 人脸库中识别率为 99.1%; 说话人识别基于 i-vector 实现, 实验预处理系数: 预加重 0.96、汉明窗、双阈值端点检测、39 维 MFCC 特征、UBM 高斯混合度

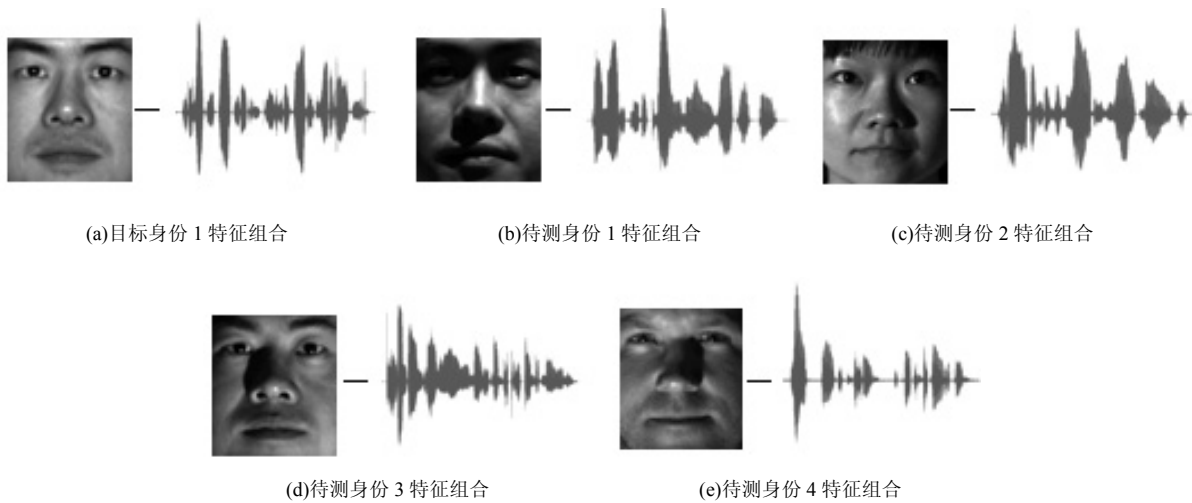


图6 目标身份1、待测身份1、待测身份2、待测身份3、待测身份4的双模特征对

表2

身份测试结果

人脸\声音	归一化后匹配分数 a_i	特征信息质量 Q_i	结果可信系数 R_i	赋值概率函数	目标身份概率
待测身份1	0.21\0.80	0.35\0.96	0.08\0.92	0.11\0.92	0.57
待测身份2	0.06\0.69	0.53\0.78	0.05\0.94	0.03\0.91	0.25
待测身份3	0.85\0.10	0.67\0.48	0.96\0.04	0.95\0.03	0.38
待测身份4	0.08\0.17	—	—	—	—

为512, T 矩阵维度400维,在TIMIT语料库中进行与性别无关的验证,等错误率为1.85%。

实验根据3.3节中提供的算法进行融合,设置3组测试,将每组的击中率(TP, true positive rate)和误报率(FP, false positive rate)在同一坐标轴表现出来,绘制出在不同特征信息质量组合下该融合系统分类器的ROC曲线,同时与单模生物特征识别方式以及不考虑特征信息质量的融合方式进行对比,结果如图7~图9及表3所示。

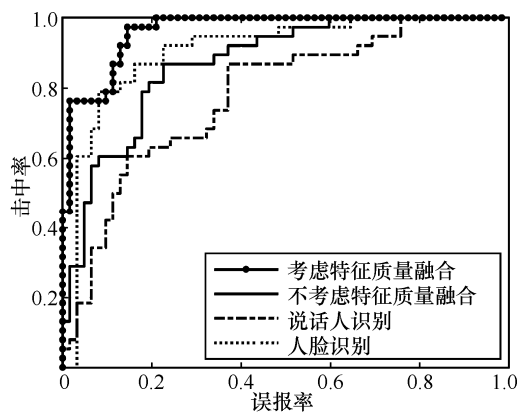
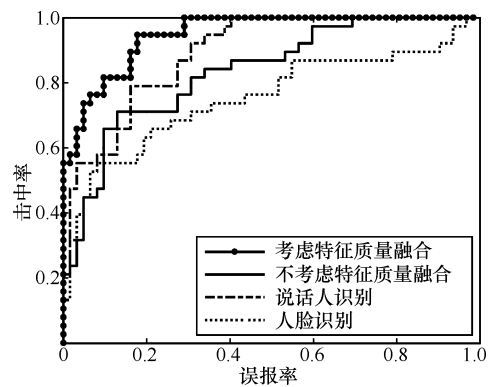
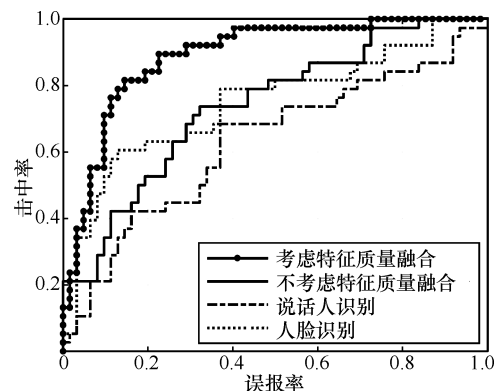
图7 待测身份特征信息质量 $Q_F \in [0.5, 0.7]$ 、 $Q_S < 0.4$ 图8 待测身份特征信息质量 $Q_S \in [0.5, 0.7]$ 、 $Q_F < 0.4$ 图9 待测身份特征信息质量 $Q_S < 0.3$ 、 $Q_F < 0.3$

表3 不同识别方式 AUC (area under curve) 值

AUC 值	本文方法	不考虑特征质量融合	说话人识别	人脸识别
$Q_F \in [0.5, 0.7]$ 、 $Q_S < 0.4$	0.947 4	0.855 7	0.762 3	0.914 3
$Q_S \in [0.5, 0.7]$ 、 $Q_F < 0.4$	0.932 7	0.819 2	0.876 1	0.737 3
$Q_S < 0.3$ 、 $Q_F < 0.3$	0.874 4	0.723 3	0.627 8	0.731 3

从上述实验结果可以清楚地看到,在不同的特征信号质量组合情况下,本文提出的方法效果最佳,特别是在待测身份的人脸图像质量指标和声音质量指标达到较低水平时($Q_S < 0.3$ 、 $Q_F < 0.3$),这种考虑特征信号质量的融合方式依然保持较高的识别准确率,优于不考虑特征质量、只进行单纯 D-S 证据融合的方法以及单模生物特征识别。同时如表 4 所示,在每 100 个测试对中,本文方法在时间效率上与单独使用说话人识别、人脸识别几乎相同,完全满足实际识别检测需要。基于如上实验数据,本文提出的方法具有检测速度快、识别准确率高、对特征信号质量顽健性好等特点。

表4 不同识别方式身份识别所需时间(每 100 个测试对)

识别方式	时间/s
本文方法	16.81
说话人识别	16.58
人脸识别	15.37

5 结束语

由于环境等因素干扰,人脸识别和说话人识别的结果可能存在不确定性,而 D-S 理论则是表达和处理不确定性知识的数学工具,适合上述 2 种识别方式的融合。鉴于生物特征信息质量对识别准确度的影响,一般来说,越好的特征信息往往会有较高的识别率,根据特征信息质量对识别方式权重进行动态调整,减小质量较差的特征信息对识别的影响,采用基于特征信息质量动态加权的 D-S 证据融合的方法,可以有效地利用多模生物识别的优点,提高系统认证的安全性和可靠性。最后实验结果表明,该融合系统可以有效减少特征信息质量差异的影响,提升识别的准确率和安全性。

参考文献:

- [1] JAIN A, BOLLE R, PANKANTI S. Biometrics Personal Identification in Networked Society[M]. Boston: Kluwer Academic Publishers, 1999.
- [2] 绕过三星 Galaxy S8 的虹膜识别? 黑客表示非常简单[EB/OL]. <http://www.freebuf.com/news/135600.html>.
To bypass the iris recognition of samsung Galaxy S8? Hackers say it's very simple[EB/OL]. <http://www.freebuf.com/news/135600.html>.
- [3] 疯狂的身份证: 一张身份证如何攻破人脸识别技术[EB/OL]. <http://www.freebuf.com/articles/others-articles/129889.html>.
Crazy ID card: how does a card breach face recognition technology[EB/OL]. <http://www.freebuf.com/articles/others-articles/129889.html>.
- [4] ROSS A, JAIN A. Information fusion in biometrics[J]. Pattern Recognition Letters, 2003, 24(13): 2115-2125.
- [5] ROSS A, JAIN A. Multimodal biometrics: an overview[C]//European Signal Processing Conference. 2015: 1221-1224.
- [6] MAURER D E, BAKER J P. Fusing multimodal biometrics with quality estimates via a Bayesian belief network[J]. Pattern Recognition, 2008, 41(3): 821-832.
- [7] NANDAKUMAR K, CHEN Y, JAIN A, et al. Quality-based score level fusion in multibiometric systems[C]//The International Conference on Pattern Recognition. 2006: 473-476.
- [8] KHIARI-HILI N, MONTAGNE C, LELANDAI S, et al. Quality dependent multimodal fusion of face and iris biometrics[C]//The International Conference on Image Processing Theory TOOLS and Applications. 2017: 1-6.
- [9] 杨静波. 融合说话人识别和人脸识别的身份认证[D]. 兰州: 西北师范大学, 2016.
YANG J B. Research on authentication by fusing with speaker recognition and face recognition[D]. Lanzhou: Northwest Normal University, 2016.
- [10] 陈倩. 多生物特征融合身份识别研究[D]. 杭州: 浙江大学, 2007.
CHEN Q. Multimodal biometric personal identification based on data fusion[D]. Hangzhou: Zhejiang University, 2007.
- [11] LEE S Y, HAM Y K, PARK R H. Recognition of human front faces using knowledge-based feature extraction and neuro fuzzy algorithm[J]. Pattern Recognition, 1996, 29(11): 1,863-1,876
- [12] KIRBY M, SIROVICH L. Application of the karhunen-loeve procedure for the characterization of human faces[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 12(1): 103-108
- [13] TAIGMAN Y, YANG M, RANZATO M, et al. DeepFace: closing the gap to human-level performance in face verification[C]//IEEE Computer Vision and Pattern Recognition. 2014: 1701-1708.
- [14] SUN Y, WANG X, TANG X. Deep learning face representation from predicting 10,000 classes[J]. Computer Vision & Pattern Recognition. 2014: 1891-1898.
- [15] SUN Y, WANG X, TANG X. Deep learning face REPRESENTA-

- TION by joint identification-verification[J]. Advances in Neural Information Processing Systems, 2014, 27: 1988-1996.
- [16] SUN Y, WANG X, TANG X. Deeply learned face representations are sparse, selective, and robust[J]. Computer Vision & Pattern Recognition, 2015: 2892-2900.
- [17] SUN Y, LIANG D, WANG X, et al. DeepID3: face recognition with very deep neural networks[J]. Computer Science, 2015.
- [18] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: a unified embedding for face recognition and clustering[J]. Computer Vision & Pattern Recognition, 2015: 815-823.
- [19] DOUGLAS A, REYNOLDS, QUATIERI T F, et al. Speaker verification using adapted Gaussian mixture models[J]. Digital Signal Processing, 2000, 10(1-3):19-41.
- [20] KENNY P, BOULIANNE G, OUELLET P, et al. Speaker and session variability in GMM-based speaker verification[J]. IEEE Transactions on Audio Speech & Language Processing, 2007, 15(4): 1448-1460.
- [21] DEHAK N, DEHAK R, KENNY P, et al. Support vector machines versus fast scoring in the low-dimensional total variability space for speaker verification[C]//Conference of the International Speech Communication Association. 2009:1559-1562.
- [22] KRYSZCZUK K, DRYGAJLO A. What do quality measures predict in biometrics?[C]//European Signal Processing Conference. 2015:1-5.
- [23] HSU R L V, SHAH J, MARTIN B. Quality assessment of facial Images[C]//Biometric Consortium Conference. 2006:1-6.
- [24] CAPPELLI R, FERRARA M, MALTONI D. The quality of fingerprint scanners and its impact on the accuracy of fingerprint recognition algorithms[C]//The International Conference on Multimedia Content Representation. 2006:10-16.
- [25] ABAZA A, HARRISON M A, BOURLAI T, et al. Design and evaluation of photometric image quality measures for effective face recognition[J]. Biometrics Iet, 2014, 3(4):314-324.
- [26] BARTOS A L, NELSON D J. Enabling improved speaker recognition by voice quality estimation[C]//Signals, Systems & Computers. 2012: 595-599.
- [27] SELLAHEWA H, JASSIM S A. Image-quality-based adaptive face recognition[J]. IEEE Transactions on Instrumentation & Measurement, 2010, 59(4):805-813.
- [28] 高修峰, 张培仁, 李子青. 人脸图像质量评估标准[J]. 小型微型计算机系统, 2009, 30(1):95-99
GAO X F, ZHANG P R, LI Z Q. Standardization of face image sample quality[J]. Journal of Chinese Computer Systems, 2009, 30(1): 95-99.
- [29] ABBOUD A J, SELLAHEWA H, JASSIM S A. Quality based approach for adaptive face recognition[C]//The International Society for Optics and Photonics. 2009.
- [30] WANG Z, BOVIK A C. A universal image quality index[J]. IEEE Signal Processing Letters, 2002, 9(3):81-84
- [31] FERZLI R, KARAM L J. A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)[J]. IEEE Transactions on Image Processing, 2009, 18(4):717-28.
- [32] NARVEKAR N D, KARAM L J. A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)[J]. IEEE Transactions on Image Processing, 2011, 20(9): 2678-2683.
- [33] 肖宾杰. 基于图像质量加权的 D-S 证据理论多生物特征融合识别[J]. 计算机应用, 2012, 32 (1): 264-268.
- XIAO B J. Multi-biometric feature fusion identification based on D-S evident theory of image quality with different weights[J]. Journal of Computer Applications, 2012, 32 (1):264-268
- [34] 蒋晔. 基于短语音和信道变化的说话人识别研究[D]. 南京: 南京理工大学, 2013.
- JIANG Y. Research on speaker recognition over short utterance and varying channels[D]. Nanjing: Nanjing University of Science and Technology. 2013.
- [35] BLATNIK R, KANDUS G. Influence of the speech quality in telephony on the automated speaker recognition[C]//WSEAS International Conference on Circuits. 2011:115-120
- [36] RIX A W, BEERENDS J G, HOLLIER M P, et al. Perceptual evaluation of speech quality (pesq)[J]. Journal of the Audio Engineering Society, 2001, 50(10):755-764.
- [37] JAIN A, NANDAKUMAR, ROSS K A. Score normalization in multimodal biometric systems[J]. Pattern Recognition, 2004, 38(12):2270-2285.
- [38] 王延平, 袁杰, 廖原, 等. 利用信息融合技术的缺损目标识别方法[J]. 中国图象图形学报, 2000, 5(3):237-240.
- WANG Y P, YUAN J, LIAO Y, et al. Pattern recognition of occluded objects using information fusion[J]. Journal of Image and Graphics, 2000, 5(3):237-240.

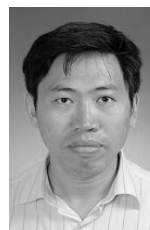
[作者简介]



张闻彬(1994-), 男, 山东济南人, 中国海洋大学硕士生, 主要研究方向为信息安全、模式识别。



刘培顺(1975-), 男, 山东菏泽人, 中国海洋大学讲师, 主要研究方向为网络与信息安全。



薛峰会(1982-), 男, 山东沂水人, 青岛黄海学院副教授, 主要研究方向为网络与信息安全。