计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2021,30(4):227−233 [doi: 10.15888/j.cnki.csa.007857]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

安全人脸识别解决方案研究①

沈　玺, 康家梁, 王伟鹏

(中国银联, 上海 201201)
通讯作者: 沈　玺, E-mail: shenxi@unionpay.com

摘　要: 人脸面部特征, 与生俱来, 具有唯一性、自然性、终身不变等特点, 因此人脸识别作为一种身份鉴别方式, 相比于传统的认证技术具有巨大的便利性优势. 然而人脸识别容易受到图片、视频、面具等伪造攻击, 由于其不可更改特性, 若生物特征发生泄漏或篡改, 会造成难以挽回的风险和损失. 本文提出的安全人脸识别解决方案, 通过特殊的安全人脸采集模组, 完成活体检测, 模组中通过人脸加密专用密钥, 对图片进行加密、签名处理, 同时在识别流程中的多重安全设计, 保证人脸生物特征的安全.

关键词: 人脸识别; 活体检测; 生物特征; 人脸检测; 安全模组

引用格式: 沈玺,康家梁,王伟鹏.安全人脸识别解决方案研究.计算机系统应用,2021,30(4):227–233. http://www.c-s-a.org.cn/1003-3254/7857.html

Research on Secure Face Recognition

SHEN Xi, KANG Jia-Liang, WANG Wei-Peng

(China Unionpay, Shanghai 201201, China)

Abstract: Face biometrics are unique, natural, and invariant throughout one's life cycle. Therefore, face recognition, as an identity authentication method, is much more convenient than traditional techniques. However, it is susceptible to forgery attacks such as pictures, videos, and masks. The leakage or tampering of invariant biometrics will induce irreversible losses. The solution to secure face recognition proposed in this study uses a special module of secure face acquisition for liveness detection. A special secret key in the module is applied to encrypting every face image and signing the sensitive information. Additionally, multiple steps in the recognition process guarantees the security of face biometrics.

Key words: face recognition; liveness detection; biometrics; face detection; secure module

1　引言

人脸作为最常见的生物特征, 每个个体独一无二, 可作为身份鉴别的依据. 通过人脸图像识别用户身份, 具有识别速度快、精度高、无接触等优点, 被广泛应用于各行业的身份认证环节. 与其他生物特征不同, 人脸的生物特征, 具有外露性、自然性特点, 这使得复制人脸变得容易, 二维图片攻击、视频攻击、3D 头模攻击、甚至计算机人脸图像合成攻击, 是人脸识别必须考虑的安全问题. 通常在识别开始之前, 采取活体检测技术来对抗欺骗攻击[1]. 在完成活体检测, 认为用于识别的图像来自于生物活体后, 基于人脸的身份识别认证才能生效.

通常在人脸识别终端上安装有特定图像采集摄像头, 收集面部 RGB 图像和深度特征, 由终端应用完成活体检测后, 向云端发起人脸识别, 鉴别用户身份. 值得注意的是, 在多数场景下, 识别环境是开放的, 用户在进行人脸采集时, 屏幕中往往会出现多个人脸, 如何在多人脸情况下排除干扰, 准确选取识别的用户, 即识别主体, 是人脸识别产品体验设计中需要重视的一个关键环节[2].

① 收稿时间: 2020-08-03; 修改时间: 2020-08-28; 采用时间: 2020-09-08; csa 在线出版时间: 2021-03-30

现有的人脸识别装置, 对生物特征数据的处理一般运行于终端 REE (Rich Execution Environment) 环境的应用中, 存在被替换和泄露的风险. 出于对人脸生物特征数据的隐私保护, 通常对终端有较高的要求, 需配合特定的摄像头, 进行一系列安全验证检测, 过程复杂且成本高昂.

本文提出的安全人脸识别方案, 旨在提供安全、友好的综合人脸识别服务. 方案提出一种安全人脸识别模组, 在模组内嵌入安全模块, 集成活体检测算法, 完成对人脸特征的采集、活体计算、数据签名、加密, 有效防止特征数据泄露和篡改, 保证人脸生物特征数据安全. 方案无需终端应用参与, 降低了对终端的安全性依赖, 为人脸识别快速集成提供了新思路. 此外, 方案兼顾安全和用户体验, 在人脸采集过程中, 设计一种人脸识别主体选择判断方法, 在多用户的情况下智能判断业务识别主体, 给用户带来良好的使用体验.

2 关键技术研究

2.1 活体检测

活体检测技术伴随着人脸识别技术的发展, 近几年有了长足的进步, 总结有关研究[3–5], 活体检测技术可被分为两大类: (1) 基于特征描述子的判断, 非活体与活体样本之间, 存在诸多物理特征差异, 结合一种或多种特征, 如材质、运动、深度等, 用于分析判断, 进而达到区分效果; (2) 基于分类器的判断, 指收集可观数量的活体、非活体数据作为正负样本, 构建训练集, 使用机器学习或深度学习方法, 对算法模型进行训练, 得到活体判断模型; 随着技术的应用发展, 上述两种方式也常进行有机结合, 达到更好的活体检测效果.

本文实现活体判断的方式属于第 2 类, 通过收集人脸样本的深度数据, 基于深度数据开展特征提取和训练, 最终将活体检测看做对深度特征的二分类决策过程.

本文使用 TOF (Time Of Flight) 摄像头进行人脸深度信息的采集. 通过特殊装置, 向人脸表面发射特殊调制后的光束, 经人脸反射, 由接受装置捕获, 得到光束往返的时间, 即可计算出对应的距离, 进而换算人脸的深度数据. TOF 具有响应时间快、识别距离远等优点, 越来越多的设备和应用, 选择 TOF 进行人脸原始深度信息的采集.

机器学习常用的判别式模型中, 支持向量机 (Support Vector Machine, SVM) 以稳定的性能得到了广泛的应用[6,7]. 通过核函数的变换后可适应非线性划分的问题. 模型的训练过程实际上是一个凸的二次规划的求解过程, 在高维特征空间中寻找决策超平面, 并使划分尽可能的扩大类间差距. 在二分类任务中支持向量机的表现通常也优于其他机器学习模型, 同时训练结果只与支持向量有关, 模型会有更小的体积, 这在模型最终部署时也具有明显的优势.

2.1.1 数据集

本文数据集来自内部采样, 共 50 人, 使用 TOF 相机进行人脸深度数据采集, 尽可能提供丰富的深度特征. 为避免单一采集场景影响, 采集过程对光照、背景等场景因素做随机变换.

正样本: 人脸正对相机, 变换表情、姿势、角度, 进行图像采集;

负样本: 覆盖多种类型的攻击手段, 包括: 纸张人脸折叠、弯曲、裁剪、挖洞、简易人脸面具.

本文在整体数据样本上拆分数据集, 并作交叉验证, 交叉验证进行 10 轮随机采样, 训练过程覆盖了约 2/3 的整体样本, 验证了模型的泛化性能, 测试与训练样本规模如表 1.

表 1 数据集样本数统计

样本	总数	训练集	测试集
正样本	5200	400	4800
负样本	4700	400	4300

2.1.2 模型训练

鉴于图像深度像素值一定程度上反应距离的性质, 使用 HOG (Histogram of Oriented Gradient) 梯度特征可以很好地进行数据抽象, 本文在使用支持向量机进行分类之前, 对图像进行 HOG 特征提取[8].

核函数将输入特征映射到特征空间, 以内积的形式表示, 使得二分类问题在高维空间中线性可分. 该映射过程对支持向量机的最终决策起到关键作用. 不同的核函数对模型训练有着较大的影响. 本文经过多次试验, 最终选用高斯核函数进行高维映射.

经过多次调参迭代, 最终模型的整体准确率 (*ACC*) 达到了 99.7%, 拒检率 (*FRR*) 0.5%, 误检率 (*FAR*) 维持在 0.2% 以下, 模型最终在整体样本上的混淆矩阵如图 1 所示.

混肴矩阵中, 将样例的真实类别与预测类别组合划分为真正例、假正例、真反例、假反例 4 种情形. 如表 2 所示, 其中正例为活体样本, 反例为攻击样本:

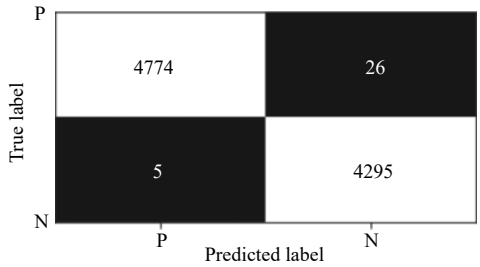

图 1 结果混淆矩阵

表 2 二分类混淆矩阵

真实标记	预测结果	
	正例	反例
正例	*TP* (真正例)	*FN* (假反例)
反例	*FP* (假正例)	*TN* (真反例)

评价标准的计算公式如下:

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$FRR = \frac{FN}{TP+FN} \tag{2}$$

$$FAR = \frac{TN}{FP+TN} \tag{3}$$

拒检率体现对正样本 (人脸) 的拒绝程度, 误检率则体现对负样本 (攻击) 的误检程度, 可以看出该模型在内部数据集上的误检率和拒检率都维持在较低的水平, 分类效果比较理想.

2.2 识别主体判断

在采集用户人脸过程中, 若屏幕中出现多个人脸, 需进行识别主体的判断, 我们认为, 最有可能的识别主体人脸, 应在用户使用过程中满足以下条件:

(1) 用户主动靠近, 人脸足够正对屏幕, 人像图片质量较高, 人脸清晰可见、无遮挡、未闭眼;

(2) 识别主体用户相比被拍摄到的人脸用户, 距离屏幕最近;

(3) 用户识别使用过程中未离开, 人脸始终出现在屏幕中.

将上述条件转换为特征指标, 识别主体人脸应具备以下特征:

(1) 人脸角度不应过大、人脸完整、无遮挡、睁眼、屏幕中人脸大小足够大;

(2) 在满足上述条件的多个人脸中, 第二大人脸需小于第一大人脸的 n (n 可根据业务情况和实际模型调整, 达到足够的区分度);

(3) 在上述两个条件判断后的最大人脸, 对其进行人脸跟踪, 结合业务流程, 在屏幕中持续足够的时间 T, 完成识别主体的最终确认.

识别主体的判断, 可分为 3 个过程, 如图 2 所示.

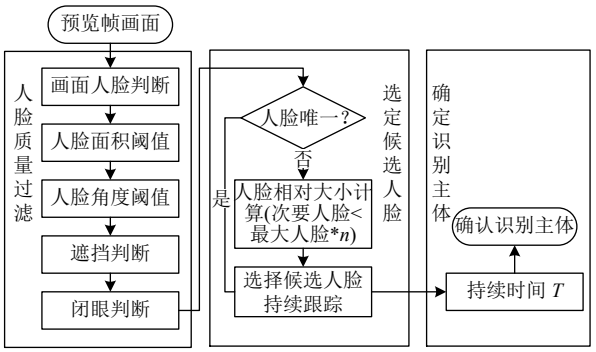

图 2 识别主体选择流程

(1) 人脸质量过滤: 通过预设参数, 筛选过滤角度过大、不完整、闭眼、模糊以及过小的人脸图片.

(2) 选定候选人脸: 在符合条件的多个人脸中, 主体人脸相比其他次要人脸, 要占有绝对大小优势, 选择多个人脸中的最大人脸, 以其 60% 作为阈值, 判断其他次要人脸的大小是否超过阈值, 若次要人脸均不超过阈值, 则将最大人脸作为候选人脸, 若存在超过阈值的最大人脸, 则提示无法确认识别主体, 请其他用户配合后退.

(3) 在选定候选主体后, 对其进行连续帧的人脸跟踪, 持续期望时间 T 后, 将其作为识别主体. 具体的, 在选定候选主体后, 对其分配唯一 *faceId*, 在连续帧的人脸跟踪过程中, 同一自然人脸在未离开的情况下, *faceId* 不变, 在进行连续的跟踪判断后, 达到期望时间 T, 则将其作为识别主体.

在上述过程中, 选定候选人脸阶段, 如果存在次要人脸大于阈值时, 此时往往屏幕前出现聚集, 各人脸相对靠近, 本文中引入人脸绝对位置和相对位置来优化判断结果, 在用户相对靠近的情况下仍然具有主体选择判断能力.

在实际业务场景中, 我们发现在某一固定场景的识别样本中, 识别主体在屏幕中出现的区域以及对于其他人脸的相对位置, 往往出现一定的规律性, 引入主体人脸在屏幕中的绝对位置和主体人脸在多人脸中的相对位置作为辅助参考, 可以进一步区分业务识别主体. 具体方法如下:

以画面中心像素点为坐标原点, 建立平面直角坐标系, 并划分①②③④共计 4 个象限. 同时, 以 1/2 * *width* 为宽, 1/2 * *height* 为高, 做内接矩形, 划分中心区域“C”, *width* 和 *height* 分别为图像的像素宽和像素高, 如图 3 所示. 中心区域与象限区域有重合, 对每个人脸, 进行区域标记.

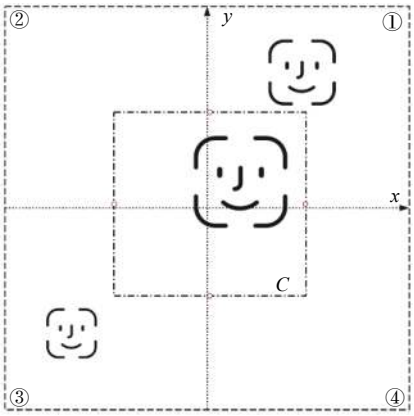

图 3 屏幕区域划分

绝对位置判断: 以画面中心点为原点 (0, 0), 构建平面坐标系. 对于每个人脸图片, 取其人脸外接矩形的中心点, 计算其坐标 (x, y), 根据人脸中心点坐标的位置, 判断其归属位置, 每个图片分配 2 位标记 *LC*, *L* 取值“1、2、3、4”表示坐落在不同象限, *C* 取值“0-在中心区域外”; “1-在中心区域内”, 可得分类如表 3.

表 3 人脸位置标记 *LC* 取值

位置	象限①	象限②	象限③	象限④
中心区域内	11	21	31	41
中心区域外	10	20	30	40

相对位置判断: 使用向量 $A(X, Y) = O_{\text{次人脸}}(x_1, y_1) - O_{\text{主体}}(x_2, y_2)$, 其中 $O_{\text{主体}}(x_2, y_2)$ 表示识别主体人脸的中心点坐标, $O_{\text{次人脸}}(x_1, y_1)$ 为次要人脸的人脸中心点坐标. 向量 $A(X, Y)$ 中, $X > 0$ 则表示主体人脸在横轴方向上, 在次要人脸的左侧, 反之则在右侧; $Y > 0$ 则表示主体人脸在纵轴方向上, 在次要人脸的下方, 反之则在上方.

在场景和机位固定后, 通常最佳人脸位置和相对人脸位置固定, 统计足够数量的 (5 万张) 正向交易且存在多人脸的交易场景图片, 统计绝对位置 *LC* 值, 以及相对位置向量 $A(X, Y)$, 根据统计分类, 寻找最佳绝对位置和相对位置, 在次要人脸的像素面积超过最大人脸的 n_1 时, 若最大人脸位置与次要人脸相对位置, 且最大人脸的绝对位置符合场景最佳位置规律, 此时可将阈值调整到 n_2 ($n_2 > n_1$), 提升识别主体选择区分强度.

3 安全人脸识别设计

3.1 安全人脸识别模组设计

为加强人脸生物特征数据的安全保护, 提高刷脸安全壁垒, 本文研究并设计了安全人脸识别模组, 使用 SE 芯片, 在模组中嵌入安全模块[9,10], 在安全模块中存储安全人脸密钥, 对采集图像进行活体判断, 并在判断为活体的情况下, 对结果进行签名, 图像进行加密, 保证数据安全.

装置的整体设计示意图如图 4 所示.

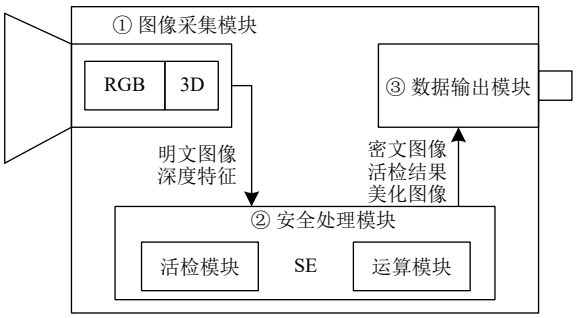

图 4 模组架构图

安全人脸识别模组主要包含 3 个功能模块: ① 图像采集模块; ② 安全处理模块; ③ 数据输出模块.

图像采集模块: 是摄像头基本功能模块, 用于图像信息采集, 生成图像信息流用于后续处理. 图像采集模块除 RGB 摄像头模组外, 配置用于活体检测的 3D 摄像头活检模组, 通过活检模组, 采集生物 3D 活体检测特征数据.

安全处理模块: 安全人脸识别模组的核心模块, 集成嵌入活检模块和运算模块两个子模块, 存储人脸专用安全密钥, 对数据进行安全处理, 如图 5 所示, 活检和运算模块的具体功能如下.

活检模块, 对输入深度特征数据进行运算, 判断活体结果, 将处理结果传递给运算模块.

运算模块, 对数据的处理可分为签名、加密、图像处理 3 部分, 如图 5 所示.

(1) 数据签名, 使用专用人脸安全密钥, 对原始人脸图像数据、活检结果、交易要素等数据, 进行签名, 保证数据在后续交易中不被替换;

(2) 数据加密, 使用专用人脸安全密钥, 对原始人脸图像数据进行加密, 加密后密文只有同样存储密钥的人脸识别服务后台可解密, 保证生物特征数据不被泄露;

(3) 图像处理, 对人脸图像进行裁剪、压缩、美颜、引入噪声等处理, 用于交互展示, 处理后图像不参与签名和识别, 与原始数据隔离, 解决交互展示需求.

数据输出模块: 数据输出模块作为对外统一接口模块, 接收安全处理模块传输的相关数据, 通过标准接口对外提供如下数据: 活体检测结果、结果签名、加密的图像数据, 用于展示的美化图片明文数据.

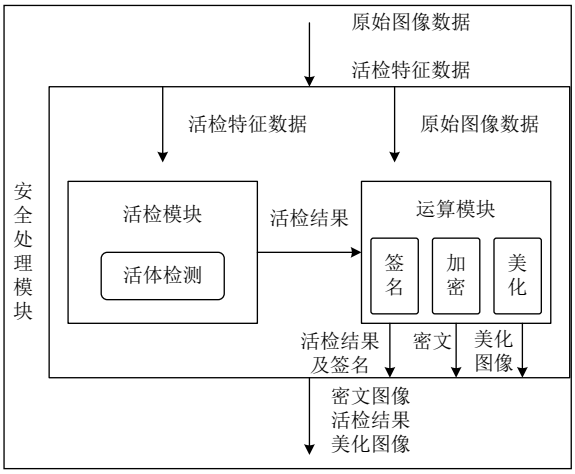

图 5 安全处理流程

3.2 系统架构设计

综合考虑安全、成本和用户体验, 系统整体架构设计如图 6 所示.

安全人脸识别模组: 独立人脸采集摄像头装置, 支持即插即用, 支持人脸深度信息采集. 集成 SE 的安全芯片, 存储人脸专用安全密钥, 在安全芯片中完成活体检测、数据加密和结果签名等操作, 从源头保证密钥安全、数据安全; 安全人脸识别模组加密后的数据直接用于后台识别, 对应的明文图片, 仅用于 APP 展示.

智能刷脸终端: 指安装人脸识别应用, 提供 REE 运行环境, 多为安卓系统, 形态不限, 可直接安装安全人脸识别模组进行人脸识别业务.

人脸识别应用: 面向人脸识别用户提供的人脸识别应用, 开启摄像头采集用户人脸, 向用户展示识别结果, 并根据识别结果进行业务处理.

人脸识别后台: 指人脸识别应用后台, 负责处理人脸识别应用请求, 提供人脸照片密文解密、验签、人脸识别等服务.

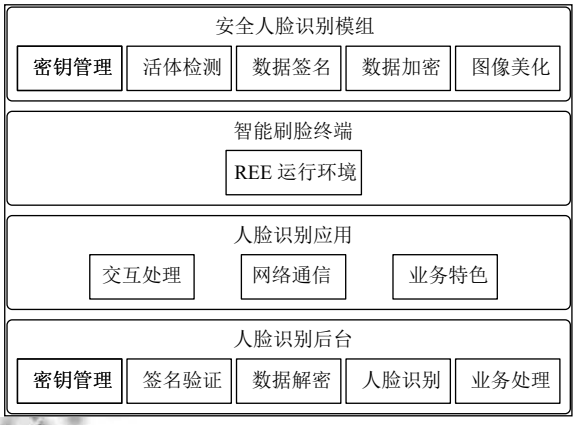

图 6 系统架构图

3.3 流程设计

系统整体数据流程如图 7.

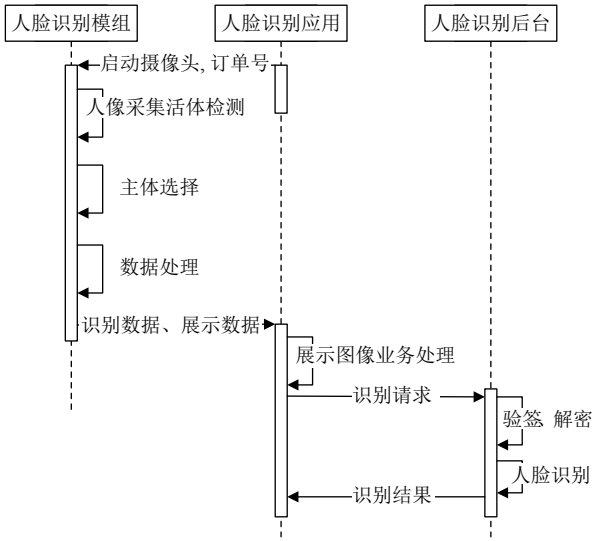

图 7 数据流程图

总体步骤如下:

(1) 人脸识别应用发起识别流程, 调起人脸识别模组采集人脸, 传入交易订单号及其他交易要素;

(2) 由安全人脸识别模组进行人脸数据采集, 通过安全模块, 进行活体判断, 检测来自生物活体的人脸图像;

(3) 若发现多个活体用户人脸, 则按照前文介绍方法进行识别主体选择;

(4) 在确认单一识别主体后, 在安全模块中对数据进行处理: 首先, 对通过活检的原始图、活检结果、订单信息, 使用专用人脸签名密钥进行签名, 并使用专用

人脸加密密钥, 对原始图数据加密; 其次, 输出作为应用展示的识别主体人脸图, 该图像在原始图像上经过处理不参与识别;

(5) 人脸识别模组将原始图像密文数据、签名结果和展示图像明文数据, 返回给人脸识别应用;

(6) 人脸识别应用展示美化后的图像, 并根据业务需要, 决定是否需要采集额外辅助信息, 发起识别交易;

(7) 人脸识别后台在收到识别请求后, 对上送签名字段进行验签, 验签通过, 表明图片未被篡改, 属于该笔订单下的原始图像, 然后使用专用安全密钥, 对密文图片解密, 安全密钥保护图像原始数据不会泄露;

(8) 在解密后, 使用原始图像数据进行特征提取, 进行人脸识别;

(9) 人脸识别后台将识别结果返回人脸识别应用, 人脸识别应用展示识别结果, 发起后续业务处理.

3.4 安全设计

在人脸识别应用的使用过程中, 安全始终是重要的考量因素, 尤其在人脸识别与金融领域相结合时,

需要考虑生物特征安全与资金安全. 系统安全设计涉及秘钥安全、生物特征安全、业务安全、身份安全、存储安全 5 个维度, 如图 8 所示.

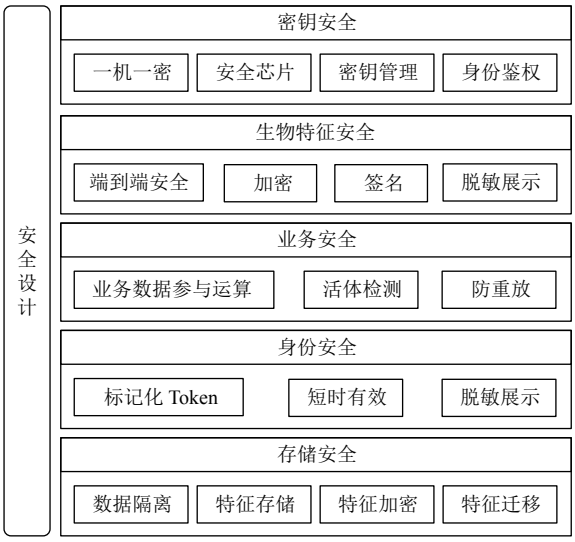

图 8 安全设计

3.4.1 密钥安全

通过人脸识别模组的安全芯片, 保证人脸专用签名密钥、人脸专用加密密钥安全, 一机一密, 即每个模组单独设置一套安全密钥, 存储于安全芯片中, 通过密钥管理体系完成鉴权管理, 维护模组密钥.

3.4.2 生物特征安全

人脸生物特征在采集后, 通过专用密钥在安全芯片中完成图像处理, 用于识别的数据在第一时间完成签名、加密; 密钥只在模组安全芯片和人脸识别后台中存在, 人脸明文数据不参与报文传输, 保证数据无法被篡改和盗取, 实现端到端安全.

出于人脸识别业务的用户体验特殊要求, 需要用户图像进行交互展示, 在原始识别图像中, 对识别主体进行人像裁剪, 并配合人像美化、防 Hack 技术, 对图片进行处理, 处理后的图片, 不能进行识别, 仅作为交互展示使用, 如此解决了生物特征隐私安全需要和前端交互展示的矛盾, 保证用户生物特征安全.

3.4.3 业务安全

在进行图像采集处理之前, 会将订单号等业务有关的交易信息传入模组中, 在对图像进行签名时, 业务信息一同参与签名计算, 如此, 保证交易上送的图像一定来自于该交易采集, 防止数据替换和重放.

在人像采集过程中, 通过活体检测, 能够有效抵御多种形式的非活体攻击, 保证交易安全.

3.4.4 身份安全

在用户身份识别成功后, 与之绑定的身份信息及其他业务敏感信息, 通过标记化技术以 Token 替代, 作为识别成功的短时有效标识, 使用 Token 发起后续业务处理, 保证用户身份和信息安全, 用户姓名等进行脱敏展示.

3.4.5 存储安全

为保证用户特征数据存储安全, 整个系统中不保存用户图像原始数据, 在人脸识别后台系统中, 对用户数据进行特征提取, 仅加密存储用户特征数据. 用户特征在识别算法和模型升级时, 进行特征迁移. 在人脸识别模组、人脸识别应用等前端不进行存储.

3.5 产品体验优化设计

(1) 人脸采集模组即插即用

人脸识别模组逻辑独立, 自维护安全密钥、模块应用体系, 安全模块具有较强的计算能力, 支持终端即插即用, 将活体检测、数据处理嵌入安全模块中, 降低对终端安全性的依赖. 理论上, 不考虑其他业务要求的情况下, 任意一款安卓终端均可使用人脸采集模组, 免开发, 即插即用, 降低成本的同时, 提供快速的安全人脸识别解决方案.

(2) 多人识别主体判断

在人脸采集过程中, 考虑多人脸情况下的识别主

体判断, 避免发生“误抓拍”“拍错人”等现象, 提升用户使用体验.

4 结论与展望

本文提出一种安全人脸识别解决方案, 通过人脸识别采集模组内嵌安全模块, 存储人脸专用安全密钥, 在模组中完成活体检测, 数据签名、加密、以及展示数据的后处理, 在满足业务展示需求的同时, 确保生物特征数据的安全性, 在传输和识别过程中, 数据无法被篡改或盗取.

通过一整套的安全解决方案, 能够在兼顾安全和体验的条件下, 降低设备硬件成本, 减少开发接入工作, 可快速产出安全的人脸识别解决方案. 方案设计具有通用性, 可在识别完成后, 自定义业务流程, 应用于不同行业的业务场景. 通过本研究, 有望大力拓展人脸识别有关业务发展, 是在生物特征识别应用领域, 对生物特征隐私保护的很好实践.

参考文献

1 杨巨成, 代翔子, 韩书杰, 等. 人脸识别活体检测综述. 天津科技大学学报, 2020, 35(1): 1–9, 17. [doi: 10.13364/j.issn.1672-6510.20190151]

2 张庆杰, 龚涵适. 人脸识别支付用户使用意愿研究. 财经理论与实践, 2018, 39(5): 109–115. [doi: 10.3969/j.issn.1003-7217.2018.05.016]

3 李冰. 人脸识别系统中的活体检测技术的研究 [硕士学位论文]. 天津: 天津大学, 2016.

4 刘茹莹. 浅谈人脸识别技术的发展. 信息记录材料, 2020, 21(4): 21–22.

5 Xie ZM, Li JJ, Shi H. A face recognition method based on CNN. Journal of Physics: Conference Series, 2019, 1395: 012006. [doi: 10.1088/1742-6596/1395/1/012006]

6 吴庆洪, 高晓东. 稀疏表示和支持向量机相融合的非理想环境人脸识别. 计算机科学, 2020, 47(6): 121–125. [doi: 10.11896/jsjkx.190500058]

7 Kumar S, Singh S, Kumar J. Live detection of face using machine learning with multi-feature method. Wireless Personal Communications, 2018, 103(3): 2353–2375. [doi: 10.1007/s11277-018-5913-0]

8 付智军. 基于 hog 图像特征的人脸识别系统设计与实现. 自动化技术与应用, 2020, 39(1): 117–120.

9 龚哲兮, 施彦媛. 基于 RK3288 国产化平台下嵌入式人脸识别系统的开发. 通信技术, 2020, 53(3): 781–785. [doi: 10.3969/j.issn.1002-0802.2020.03.041]

10 涂文军, 杨先明. 低功耗嵌入式人脸识别模组关键技术分析及研究. 五金科技, 2018, 46(4): 84–85. [doi: 10.3969/j.issn.1001-1587.2018.04.032]