

基于多光谱图像融合的掌纹识别方法

许学斌^{1,2}, 邢潇敏^{1,2*}, 安美娟^{1,2}, 曹淑欣^{1,2}, 孟 堪^{1,2}, 路龙宾^{1,2}

1. 西安邮电大学计算机学院, 陕西 西安 710121

2. 西安邮电大学陕西省网络数据分析与智能处理重点实验室, 陕西 西安 710121

摘 要 生物特征识别在信息安全领域发挥着重要作用, 掌纹识别作为一种新型生物特征识别方式, 具有低失真、非侵入性和高唯一性等优势。传统掌纹研究大多使用自然光成像系统以灰度格式获取, 识别精度很难进一步提升。为了获得更多的身份鉴别信息, 提出利用多光谱掌纹图像代替自然光掌纹图像。针对现有掌纹识别算法由于没有考虑到不同光谱的特性而导致纹理细节丢失, 识别精准率低的问题, 提出了一种基于多光谱图像融合的掌纹识别算法。该方法通过对不同光谱下的掌纹图像进行快速自适应二维经验模式分解(FABEMD), 将多光谱掌纹图像分解成一系列频率由高到低的二维固有模式函数(BIMF)和一个残余分量, 残余分量可被视为该光谱图像低频信息的初步估计。图像采集过程中光照条件很难保持稳定, 而近红外光谱图像在进行 FABEMD 分解时对光照变换敏感, 容易导致分解后的 BIMF 背景信息过于冗余; 因此对分解后的近红外掌纹图像进行背景重建及特征细化, 在对背景冗余信息进行平滑处理的同时可以有效增强高频信息的特征表达。为避免直接融合处理后引发的图像过度曝光问题, 提出对近红外特征压缩后再融合。此外, 提出了一种结合了注意力机制的改进残差网络(IRCANet), 用于融合后的掌纹图像分类, 在网络中引入分阶段残差结构, 缓解了网络的退化问题, 在学习过程中有效地减少信息丢失, 对于融合后的多光谱掌纹图像, 分阶段残差结构能够稳定地将图像信息在网络间传输, 但对图像中的高低频信息区分效果不够显著, 为了使网络关注更多区分性特征, 利用特征通道间的相互依赖性, 在分阶段残差结构中结合了通道注意力(Channel Attention)机制。最终, 在香港理工大学(PolyU)多光谱掌纹数据集上进行的综合实验表明, 该方法可以取得良好的效果, 算法识别准确率能达到 99.67%且具有良好的实时性。

关键词 多尺度分解; 图像融合; 多光谱掌纹识别; 注意力机制

中图分类号: TP391.9

文献标识码: A

DOI: 10.3964/j.issn.1000-0593(2022)11-3615-11

引 言

生物特征识别在现代信息社会中扮演着越来越重要的角色^[1]。掌纹作为一种新兴的生物特征, 具有识别率高、用户友好性优秀、稳定性强等显著优点, 被认为是最有效的身份鉴别方案之一^[2]。掌纹识别^[3]是基于手掌信息(包括主线、皱纹和细纹)来验证一个人的身份, 传统掌纹识别方法主要从单一的光谱波段获取特征, 无法获得足够的身份鉴别信息。多光谱掌纹图像利用不同波长的光谱对皮肤的吸收和反射率不同, 捕获每个波段中特定和互补的掌纹特征, 从而有效地提升掌纹识别效果。在最近的研究中, 使用多光谱掌纹

图像而非利用单个光谱图像来提高掌纹识别系统的准确性是一种日益增长的趋势^[4-5]。

多光谱掌纹识别技术主要集中在两个方向, 即在像素级和匹配分数级融合多光谱信息。对于像素级融合, 基本思想是在一定的尺度空间上提取掌纹图像的显著特征, 再将其集成在一起, 最后通过执行逆变换重构要识别的融合图像。研究主要集中在基于金字塔和基于小波变换。Hao 等^[6]对小波变换、梯度金字塔、形态金字塔和 Curvelet 变换四种融合方案进行比较, 分别在可见光和近红外掌纹图像的两个不同光谱带上进行了评估。此外, 其他一些创新方法, 如非采样轮廓波变换(NSCT)、离散小波变换(DWT)、平移不变数字小波变换(SIDWT)和数字剪切波变换(DST)在多光谱掌纹图

收稿日期: 2022-05-08, 修订日期: 2022-07-22

基金项目: 国家自然科学基金面上项目(61673316), 陕西省重点研发计划项目(2017GY-071), 陕西省教育厅项目(16JK1697), 陕西省技术创新引导计划项目(2021YFBT-108-02), 西安邮电大学研究生创新基金项目(CXJJYL2021024)资助

作者简介: 许学斌, 1974 年生, 陕西省网络数据分析与智能处理重点实验室硕士生导师 e-mail: ccp9999@126.com

* 通讯作者 e-mail: xingxiaomin@stu.xupt.edu.cn

像融合中得到了广泛和成功的应用, Meng^[7]等提出利用梯度下降的同时增强融合图像分辨率, 效果显著。在匹配分数级融合的情况下, 分别从不同的光谱带提取掌纹特征, 然后使用比较器获得匹配分数。这些匹配分数依次使用求和规则进行融合, 然后对融合结果进行验证。例如, Xu 等^[8]提出了主成分分析(QPCA)方法来分析在四个不同光谱波段采集的掌纹图像, 使用加权和规则在四个不同等级之间进行比较分数级融合, 与单个等级相比, 加权和规则表现出更好的性能。Khan 等^[9]在执行匹配分数级融合之前, 将轮廓代码(CC)应用于多光谱图像的表达。

在上述工作中, 像素级多光谱掌纹图像融合方案更具吸引力, 因为它可以有效地去除掌纹图像采集过程中可能存在的噪声。在多光谱掌纹识别中, 近红外光谱下捕获的掌纹图像信息量少但较为重要, 可见光谱下更关注掌纹图像的背景信息。近红外和可见光谱图像往往具有较低的相关性, 因此将多光谱掌纹图像进行融合的可靠方法是将近红外光谱下处理后的显著特征与可见光谱下提取到的特征进行融合。

提出了一种基于多尺度分解的多光谱掌纹图像识别算法。首先, 将快速自适应二维经验模式分解(FABEMD)分别应用在不同光谱的掌纹图像上提取显著特征。其次, 考虑到近红外光谱图像的特征, 通过 Bezier 插值对图像进行背景重建, 再利用光谱间差异估计可能的近红外背景细化红外特征。再次, 考虑到直接对图像进行加权融合可能会产生过度曝光的问题, 提出对近红外特征压缩后再加权融合, 使融合图像包含足够的身份鉴别信息。最终, 为了提高识别精度降低计算成本, 提出了一种结合了注意力机制的改进残差网络对融合图像进行分类。

1 实验部分

1.1 多光谱掌纹成像与采集

数据集来源于香港理工大学生物特征识别研究中心^[10], 多光谱掌纹采集设备结构如图 1, 由 CCD 摄像机(灵敏度: 1

Lux)、镜头(焦距: 3.6 mm)、A/D 转换器、多光谱光源和光控制器组成。该设备可以捕获分辨率为 352×288 或 704×576 的掌纹图像。在采集过程中, 用户需要将手掌放在设备面板上, 通过面板上的控制点固定手掌。由于连续灯光间切换时间非常短, 设备可以在一秒内采集四种光谱下的掌纹图像。两幅图像之间的平移或旋转非常小, 因此图像融合时不再需要配准。

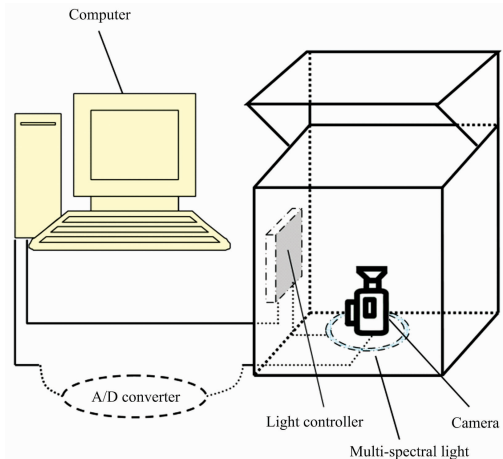


图 1 多光谱掌纹成像设备

Fig. 1 Schematic diagram of the multispectral palmprint imaging device

从采集的图像中提取感兴趣区域(ROI)是多光谱掌纹识别的关键步骤, ROI 是后续图像处理的根本运算区域。采用基于最大内切圆法的分割方法^[11], 如图 2 所示, 在对掌纹图像二值化后, 以手掌外侧轮廓作为定位参考线, 以手掌上最大内切圆的圆心为原点来构建掌纹图像坐标系, 裁剪内接正方形即掌纹 ROI。图 2(d)为分割后近红外光谱下的掌纹 ROI, 大小为 128×128 。这种分割方法能有效地增强掌纹图像中主线和褶皱线的对比度, 降低掌纹采集时旋转和平移的影响。

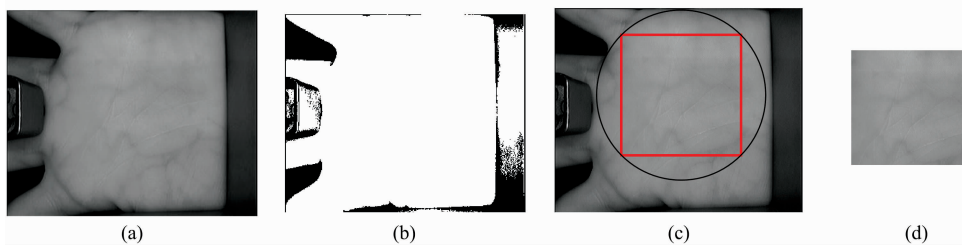


图 2 多光谱掌纹 ROI 提取过程

(a): Near-infrared palmprint; (b): Binarization;

Fig. 2 Multispectral palmprint ROI extraction process

(c): Maximum incision circle; (d): Near-infrared palmprint ROI

(c): Maximum incision circle; (d): Near-infrared palmprint ROI

1.2 基于多尺度分解的多光谱掌纹图像识别算法

图 3 为所提出的基于多尺度分解的多光谱掌纹图像识别算法流程, 该算法主要包括四个步骤, 首先多光谱掌纹图像执行快速自适应二维经验模式分解(FABEMD)。其次, 对近

红外光谱下分解后的二维固有模态函数(BIMF)通过 Bezier 插值重建背景, 利用光谱特性差异预估可能的背景从而对显著特征进行细化。再次, 为避免各光谱特征直接融合产生过度曝光的问题对近红外特征进行压缩后加权融合。最终, 利

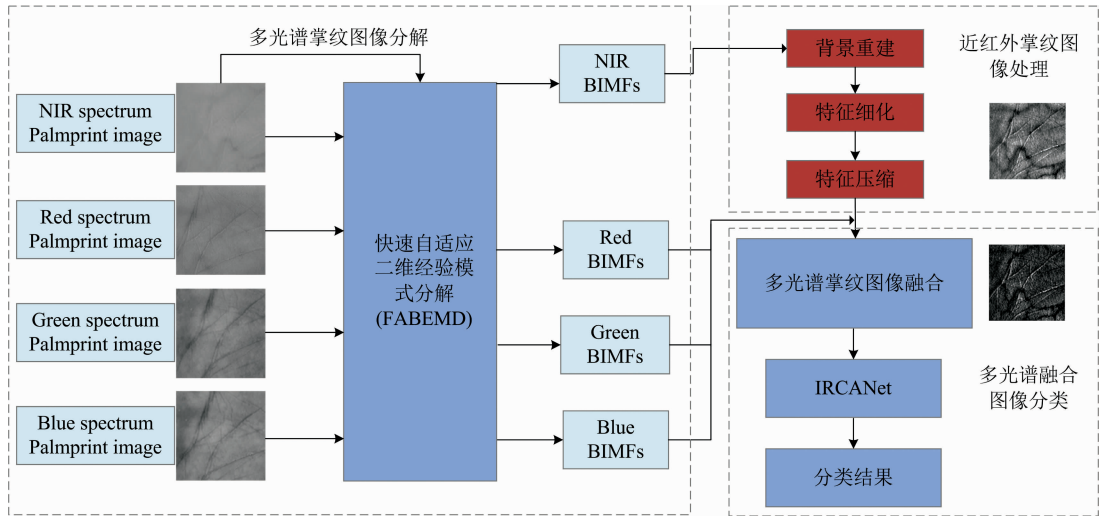


图 3 基于多尺度分解的多光谱掌纹图像识别算法流程
Fig. 3 Multispectral palmprint ROI extraction process

用结合了注意力机制的改进残差网络对多光谱掌纹融合图像进行分类。

1. 2. 1 快速自适应二维经验模式分解

快速自适应二维经验模式分解 (FABEMD) 是一种新型信号自适应处理工具^[12-13], 目的是将信号分解成一系列频率由高到低的二维固有模态函数 (BIMF) 和一个残余分量。FABEMD 算法采用顺序统计滤波器和平滑滤波器生成上下包络曲面, 其中滤波器大小由数据推导而来, BIMF 的生成基于迭代筛选。令源信号为 I , BIMF 为 S , 残余量为 R , 2D 信号 I 分解过程可由式 (1) 表示

$$I = \sum_{i=1}^K S_i + R \quad (1)$$

FABEMD 具体分解步骤如下:

Step1: 设置 $i=1$ 并初始化 $J_1 = I$ 。

Step2: 如图 4 所示, 利用相邻窗口搜索策略识别出局部极大值矩阵 M_i 和局部极小值矩阵 N_i 。如果一个数据点的值

高于或低于邻域窗口内的所有值, 则该数据点被视为局部极值。需要注意的是, 在边界或角点处寻找极值点时, 一般忽略图像外的邻域。通常, 3×3 是最佳窗口大小。

Step3: 根据局部极值 M_i 和 N_i 确定统计排序滤波器的最佳窗口大小, 对每个局部极大值 M_i 计算到其他局部极大值的欧式距离, 并将其存储在数组 $d_{adj-max}$ 中。同样地, 向量组 $d_{adj-min}$ 也是这样计算的。其中数组大小与极值数量相同。考虑到窗口形状, 统计排序滤波器总的窗口尺寸取决于式 (2)~式 (4) 四种方式

$$W_{erg} = d_1 = \min\{\min\{d_{adj-max}\}, \min\{d_{adj-min}\}\} \quad (2)$$

$$W_{erg} = d_2 = \max\{\max\{d_{adj-max}\}, \min\{d_{adj-min}\}\} \quad (3)$$

$$W_{erg} = d_3 = \max\{\min\{d_{adj-max}\}, \min\{d_{adj-min}\}\} \quad (4)$$

$$W_{erg} = d_4 = \min\{\max\{d_{adj-max}\}, \max\{d_{adj-min}\}\} \quad (5)$$

最终窗口大小 W_{en} 是对 W_{erg} 奇数取整, 多光谱掌纹图像的中窗口范围为 $3 \sim 69$ 。

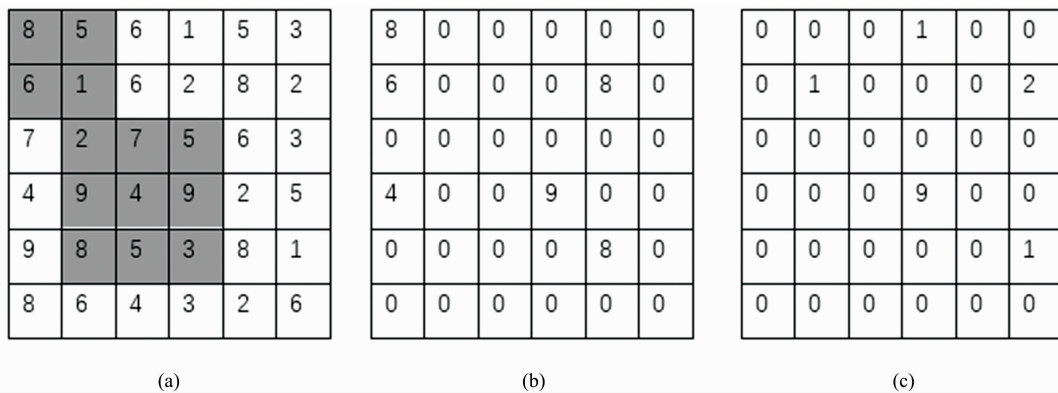


图 4 局部极大极小映射

(a): 源信号矩阵; (b): 局部最大值矩阵; (c): 局部最小值矩阵

Fig. 4 Mapping of local maxima and local minima

(a): Source signal; (b): Local maxima map; (c): Local minima map

Step4: 首先, 将窗口大小为 $\omega_{en} \times \omega_{en}$ 的 max 和 min 滤波器根据式(6)和式(7)生成上下包络。

$$U_{E_i}(m, n) = \max_{(s, t) \in z_{nm}} \{ J_i(s, t) \} \quad (6)$$

$$L_{E_i}(m, n) = \min_{(s, t) \in z_{nm}} \{ J_i(s, t) \} \quad (7)$$

其中, z_{nm} 是大小为 $\omega_{en} \times \omega_{en}$ 的方形区域, 中心为矩阵 J_i 中点 (m, n) , $U_{E_i}(m, n)$ 为区域 z_{nm} 中的最大值, 即点 (m, n) 的上包络值。同理, $L_{E_i}(m, n)$ 为点 (m, n) 在 z_{nm} 区域的最小值, 即点 (m, n) 的下包络值。为了获得上下包络线的光滑连续曲面, 对 $U_{E_i}(m, n)$ 和 $L_{E_i}(m, n)$ 进行平滑操作

$$U_{E_i}(m, n) = \frac{1}{\omega_{en} \times \omega_{en}} \sum_{(s, t) \in z_{nm}} U_{E_i}(s, t) \quad (8)$$

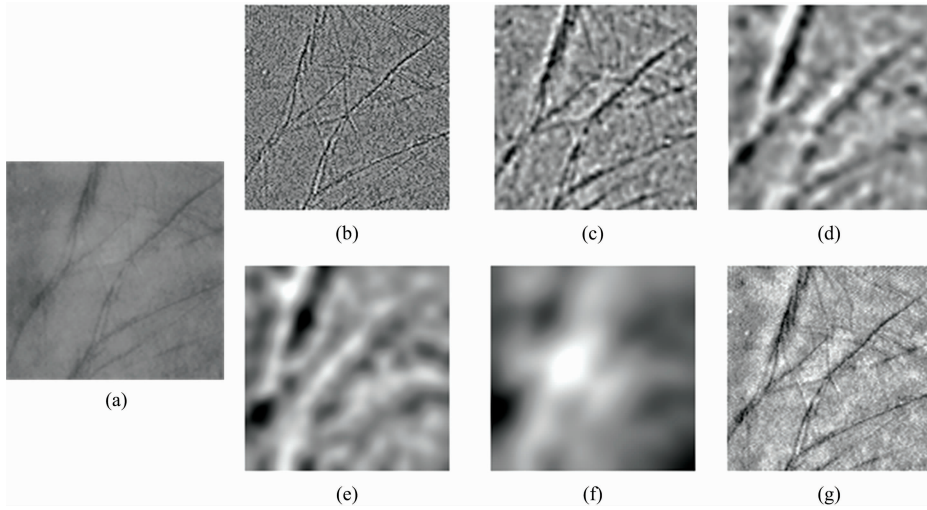


图 5 使用 FABEMD 对掌纹图像进行分解

(a): 绿色光谱下掌纹图像; (b): BIMF₁; (c): BIMF₂; (d): BIMF₃; (e): BIMF₄; (f): 残余余量; (g): S_G

Fig. 5 Decompositions of a palmprint image using FABEMD

(a): Green ROI; (b): BIMF₁; (c): BIMF₂; (d): BIMF₃; (e): BIMF₄; (f): The residue; (g): The sum of K BIMF

1.2.2 近红外背景重建及特征细化

在实际应用中, 多光谱图像采集过程并不像图 1 中描述的那样受到限制, 并且采集过程中光照条件不稳定, 而近红外光谱图像在进行 FABEMD 分解时对光照变换非常敏感, 容易导致分解后的 BIMF 背景信息过于冗余。由此, 对近红外光谱分解求和后的显著图像重建背景, 通过平滑背景来解决近红外光谱图像背景信息过于冗余的问题。

形态学算子和 Bezier 插值是两种应用广泛的背景重建技术。形态学开放算子可以有效地估计红外背景, 但估计的背景往往会产生遮挡伪影, 影响融合后图像识别准确性。而 Bezier 插值可以生成一个平滑的凸背景, 该背景与实际背景非常相似且能有效剔除冗余背景。因此, 采用 Bezier 插值算法来重建近红外背景。Bezier 插值是通过插值一些已知的控制点来恢复大规模矩阵的算法, 这些控制点可以自适应地估计图像背景。为了准确地重建红外背景, 在近红外掌纹图像中均匀采样九个控制点, 并插值重建 Bezier 曲面如式(10)

$$\tilde{S}_i(m, n) = MPS_i P^T N^T \quad (10)$$

式(10)中, (m, n) 为插值点, 其位置由 0 和 1 之间的插值比表示。 S_i 为插值点 (m, n) 的特征矩阵。 P 表示常量插值系数

$$L_{E_i}(m, n) = \frac{1}{\omega_{en} \times \omega_{en}} \sum_{(s, t) \in z_{nm}} L_{E_i}(s, t) \quad (9)$$

Step5: 通过 $S_i = (U_{E_i} + L_{E_i})/2$ 计算第 i 个 BIMF。令 $i = i+1$, $J_i = J_{i-1} - S_{i-1}$, 重复 Step2 至 Step5, 直至分解的 BIMF 个数为 k , 停止分解。

基于上述步骤, 2D 信号 I 被分解为 k 个 BIMF 再根据式(1)计算剩余残量 R 。图 5 为绿色光谱下的掌纹图像通过 FABEMD 算法分解的结果, 可以看出分解后的 BIMFs 包含稳定的掌纹特征, 其中图 5 (g) 是对 k 个 BIMFs 求和得到的特征显著的图像。

$$\text{矩阵} \begin{bmatrix} 1 & 0 & 0 \\ -3 & -6 & 0 \\ 1 & 3 & 1 \end{bmatrix}.$$

相较于其他光谱, 近红外光谱下的掌纹图像包含的特征对后续身份识别影响较大。通过对数据集分析处理, 发现近红外光谱下掌纹图像平均灰度值均大于其他光谱, 且与绿色光谱下的掌纹图像平均灰度值差异最大。因此, 利用插值后近红外图像与通过 FABEMD 算法分解后的绿色光谱图像灰度差异的正部分来估计可能的近红外背景区域, 并对估计的背景乘以适当的抑制比来进一步细化红外特征。

$$S'_i = \max(\tilde{S}_i - S_G, 0) - \alpha(\max(\tilde{S}_i - S_G, 0)) \quad (11)$$

如图 6 所示, S'_i 是最终处理后的近红外光谱下的特征图像, \tilde{S}_i 分别表示近红外光谱下经过 Bezier 插值后的图像, S_G 表示绿色光谱下经过 FABEMD 算法分解求和的图像。 α 表示背景抑制比, 通常取值在 0-1 之间, 在此设置为 0.6。

1.2.3 多光谱掌纹图像融合

直接融合所有多光谱掌纹图像有时会产生过度曝光, 如图 7(a)所示, 融合图像上的曝光点会影响后续身份识别, 为避免产生这一问题, 对处理后的近红外光谱掌纹特征进一步

压缩。

$$S''_I = \beta \times S'_I \quad (12)$$

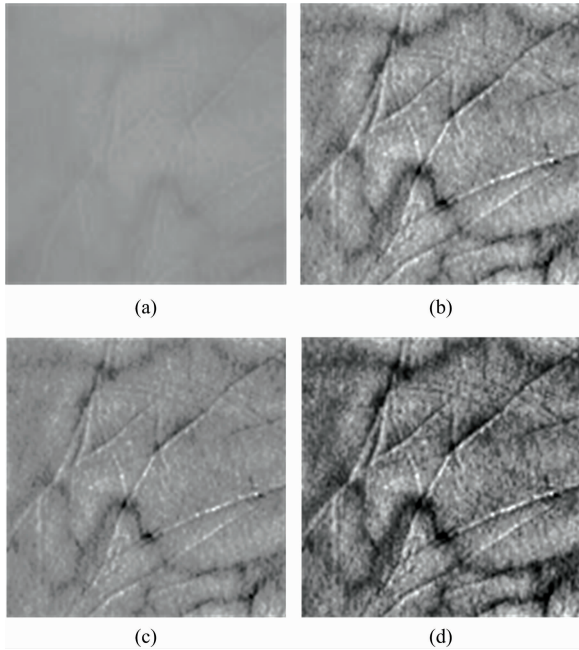


图 6 近红外掌纹图像处理过程

- (a): 近红外掌纹 ROI; (b): S_I (BIMF 求和);
- (c): \tilde{S}_I (Bezier 插值); (d): S''_I (特征细化)

Fig. 6 Near-infrared palmprint image processing

- (a): NIR ROI; (b): S_I ; (c): \tilde{S}_I (Bezier interpolation);
- (d): S''_I (feature refinement)

式(12)中, S''_I 为近红外光谱下压缩后的掌纹特征, β 为特征压缩比, 取融合后图像曝光度为前 1% 的平均值 Avg_1 来泛化整体融合后图像的过度曝光程度。

$$\beta = \min\left(\frac{255}{Avg_1}, 1\right) \quad (13)$$

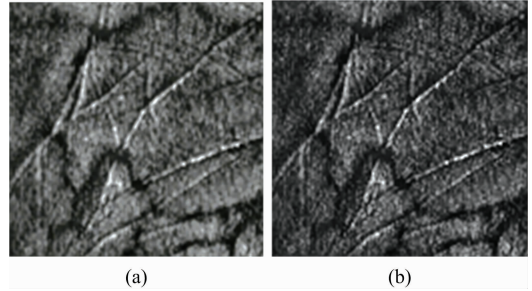


图 7 多光谱掌纹融合图像

- (a): 未经过特征压缩的融合图像; (b): 经过特征压缩的融合图像

Fig. 7 Multispectral palmprint fusion image

- (a): Fused images without feature compression;
- (b): Fused images with feature compression

至此, 将所有光谱下处理后的掌纹图像通过式(14)进行融合。

$$T^i = S''_I + \mu(S_R + S_G + S_B) \quad (14)$$

式(14)中, S_R , S_G 和 S_B 分别为在红、绿、蓝三色光谱下经过 FABEMD 算法分解求和的掌纹图像, S''_I 为经过特征压缩后的近红外光谱的掌纹图像, μ 为红、绿、蓝三色光谱在融合时的权重, 将 μ 设置为 0.2, T^i 为最终融合了各个光谱的掌纹图像, 如图 8 所示。

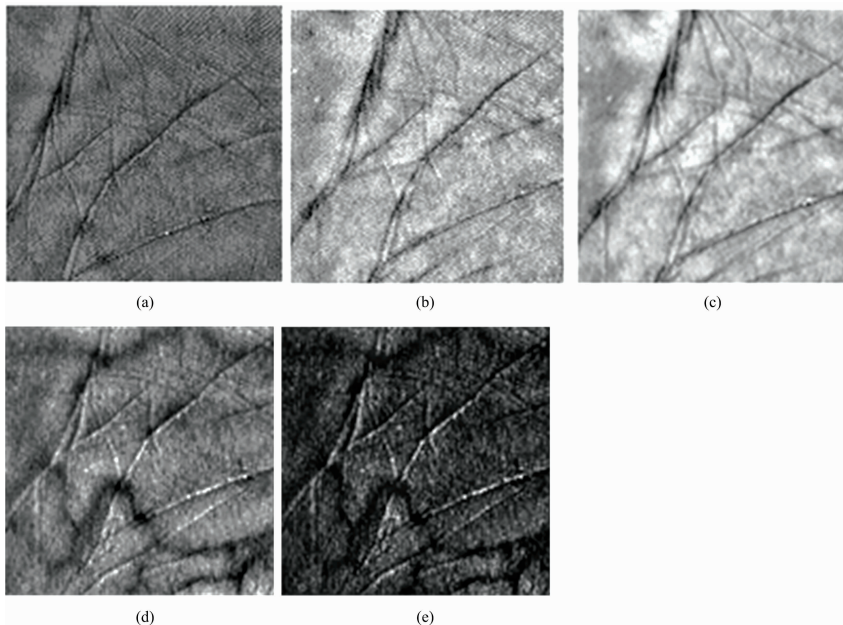


图 8 多光谱掌纹融合过程

- (a): S_R (红色光谱下分解求和的图像); (b): S_G ; (c): S_B ; (d): S''_I (经过压缩的红外特征); (e): 融合图像

Fig. 8 Multispectral palmprint fusion process

- (a): S_R (red in FABEMD); (b): S_G ; (c): S_B ; (d): S''_I (NIR with feature compression); (e): Fusion process

1.2.4 结合了注意力机制的改进残差网络(IRCANet)

随着卷积神经网络不断深入,由梯度爆炸和梯度弥散引起的网络退化问题使得学习难度不断增加。ResNet^[14]利用残差块作为网络基本单元,引入恒等映射来缓解这一问题。然而,简单地叠加残差块以构建更深的网络很难获得更好地改进,并且残差网络平等地对待每个通道特征,缺乏跨特征通道的学习能力,对融合后的多光谱掌纹图像中的高低频信

息不能有效区分,从而影响最终身份识别效果。为了解决这些问题,提出了一种基于通道注意力机制的改进残差网络(IRCANet)。

图 9 显示了本文所提出的 IRCANet 整体网络架构,它主要包括三个部分:在卷积层对浅层特征进行提取,再继续输入结合了注意力机制的改进残差模块(IRG)模块中提取有区分度的高频显著特征,最终通过全连接层进行分类识别。

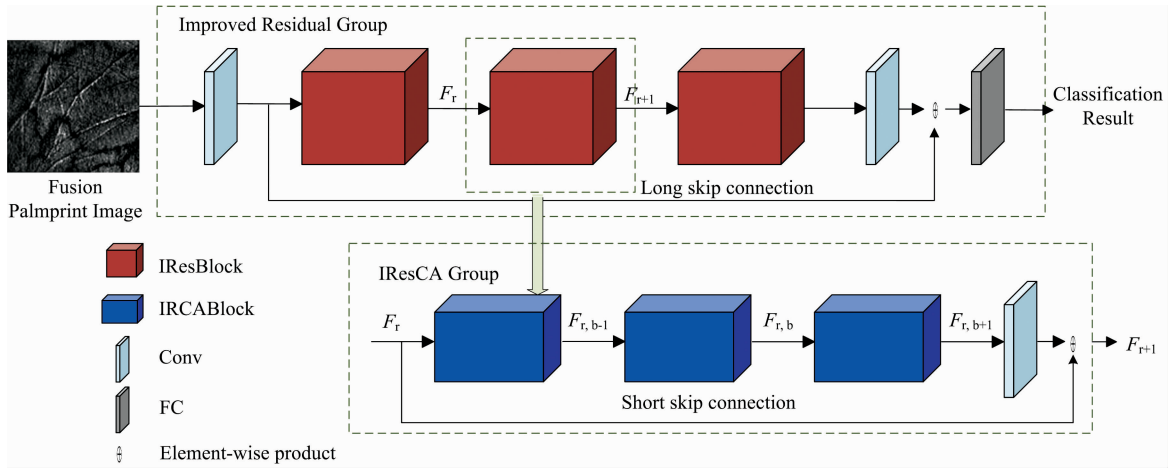


图 9 结合了注意力机制的改进残差网络的架构

Fig. 9 Network architecture of improved residual channel attention network (IRCANet)

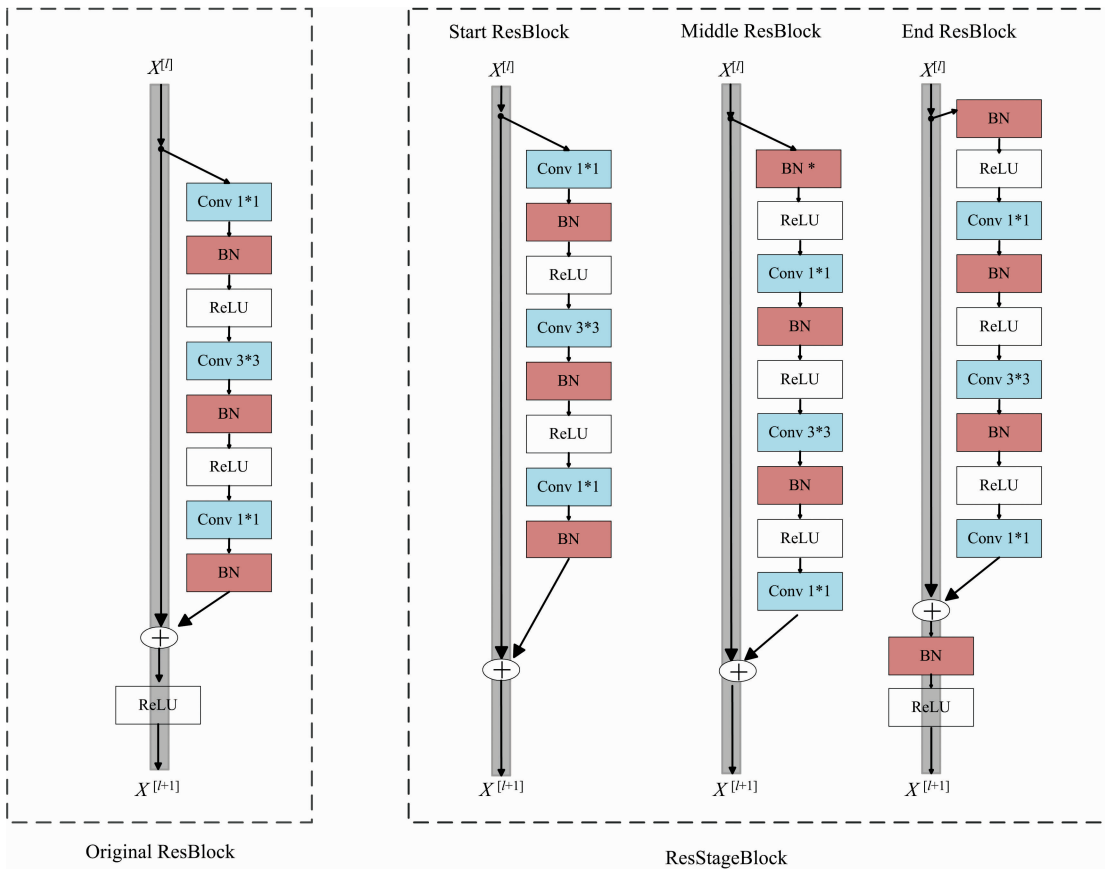


图 10 传统残差结构与分阶段残差结构

Fig. 10 Original ResBlock and ResStageBlock

首先, 利用一个卷积层(Conv)从融合后的掌纹图像 T 中提取出浅层特征 T_0 , 如式(15)

$$T_0 = U_{\text{conv}}(T) \quad (15)$$

在网络训练初期, 考虑到传统残差模块主传播路径的 ReLU 函数可能通过将信号归零的方式对信息在网络间的传播产生负面影响, 如图 10(a)。直接允许信号以不受控制的方式通过网络会使完整信号变得非标准化, 导致网络在后期训练困难。

由此, 引入了一种分阶段残差结构(ResStage Block), 为信息传播提供了更好的路径, 如图 10(b); 首先在每个阶段连接处引入批量归一化(BN)模块, 使信号稳定地进入下一阶段传播。其次, 通过固定主路径上的 ReLU 函数数量, 避免在网络加深时信号传输受阻。这种新的分阶段学习残差结

构相较于传统残差块堆叠的结构, 能够有效缓解网络的退化问题, 在学习过程中减少信息丢失, 促进了信息在网络中的传播。

对于融合后的多光谱掌纹图像, 分阶段残差结构能够稳定地将图像信息在网络间传输, 但对图像中的高低频信息区分效果不够显著。由此, 为了使网络关注更多区分性特征, 利用特征通道间的相互依赖性, 在改进的残差块中引入通道注意力(channel attention)机制。如图 11。在新的残差结构中, 融合图像里的高频信息在训练过程中更容易被关注。这种结合了注意力机制的改进残差结构能够在减少信息丢失的同时更关注通道间的特征信息, 不仅能促进信息在网络中的传播还能更进一步地增强网络的辨别能力。

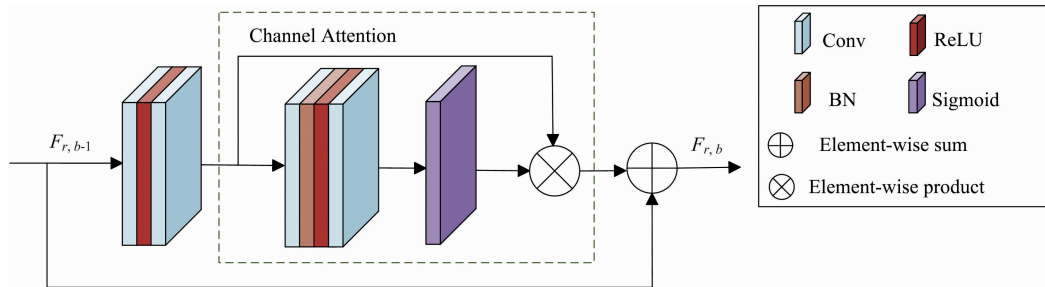


图 11 结合了注意力机制的残差模块

Fig. 11 Improved residual block with channel attention

由此, 将提取出的浅层特征 T_0 输入结合了注意力机制的改进残差模块即 IRG 模块。

$$T_{\text{DF}} = U_{\text{IRG}}(T_0) \quad (16)$$

最终, 通过全连接层对输出的深层特征进行分类处理

$$F = U_{\text{FC}}(T_{\text{DF}}) = U_{\text{IRCAN}}(T_0) \quad (17)$$

式(17)中, U_{FC} 和 U_{IRCAN} 分别表示全连接层和本文所提出的 IRCANet 结构, F 表示网络分类结果。

2 结果与讨论

对实验结果进行分析, 并评估了所提出方法的性能。利用精准度(Accuracy)作为评估标准, 定义如式(18)

$$\text{Accuracy} = \frac{\text{Num}_{\text{TP}+\text{TN}}}{\text{Num}_{\text{TP}+\text{FP}+\text{TN}+\text{FN}}} \quad (18)$$

式(18)中, $\text{Num}_{\text{TP}+\text{TN}}$ 为正确识别的样本数, $\text{Num}_{\text{TP}+\text{FP}+\text{TN}+\text{FN}}$ 为表示样本总数。

实验所使用的计算机配置是: 64 位的 Windows 10 操作系统, 仿真软件采用 MATLAB 2020a, CPU 为 Intel Core i7-10750H CPU, 内存 64 GB, GPU 是 RTX 2080 Ti。

2.1 多光谱掌纹数据集

采用香港理工大学提供的 PolyU 多光谱掌纹数据库^[8]。图像来自 20~60 岁的 195 名男性志愿者和 55 名女性志愿者, 总共分两次采集, 时间间隔约 9 d, 每次采集左右手掌共 12 张图像。掌纹图像采集于四个光谱波段, 即红色、绿色、蓝色和近红外波段。每个波段有 6 000 张来自 500 个不同手掌的图像。图 12 展示了一些多光谱掌纹 ROI 样本。

2.2 实验参数讨论

通过实验对 FABEMD 算法中顺序统计滤波器的窗长大小(d_1, d_2, d_3, d_4)及分解的 BIMF 个数进行讨论, 结果如图 13 所示。为了测试 BIMF 数量的影响, 将分解个数 K 值从 1 逐渐增加到 6, 可以观察到 K 值为 4 时并且当滤波器窗长为 d_1 时, 效果最好。

为了获得 IRCANet 最优参数, 按照训练集、测试集 7:3 的比例随机对融合图像进行划分, 再根据所设置的参数进行不同的实验对比, 得到的结果如表 1 所示, 当 IRCANet 使用最优的参数设置时, 融合图像平均精准度最高可以达到 99.67%。

2.3 实验结果分析

每个光谱带可以捕获一些特定的和互补的手掌特征, 为掌纹识别提供不同的信息。近红外光谱的掌纹图像包含的特征更为显著, 对后续身份识别干扰性强。为了验证本文所提出的对多光谱掌纹图像融合策略有效, 将近红外光谱图像处理后再融合与直接加权融合进行比较。表 2 展示了三个光谱带与近红外光谱不同组合测试本文所提出的融合策略结果。图 14 表示在融合所有光谱的条件下, 三种融合策略最优的识别精准度。可以观察到, 采用直接加权的融合策略近红外与红色光谱融合效果最好, 但也仅仅达到 92.04%, 对一些高精度识别很难符合要求。在对近红外图像进行背景重建后, 融合图像识别效果显著增强, 融合所有光谱后精准度能够达到 95.12%。由此, 进一步对红外特征细化, 很明显, 对于任何融合组合, 本文所提出的对近红外特征进行背景重建

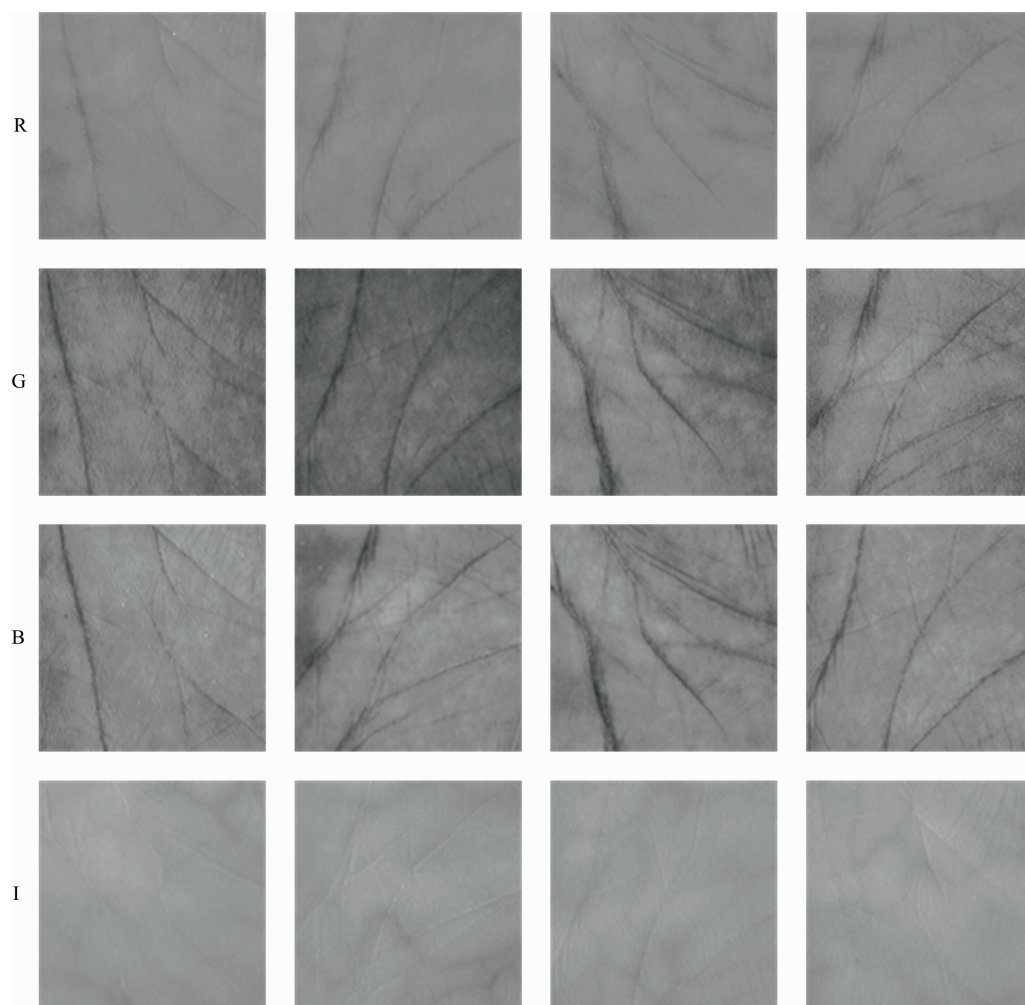


图 12 PolyU 掌纹数据集

Fig. 12 Multispectral palmprint samples in the PolyU database

及特征细化后再融合的策略效果显著, 在融合所有光谱的情况下, 识别精准度为 99.67%。

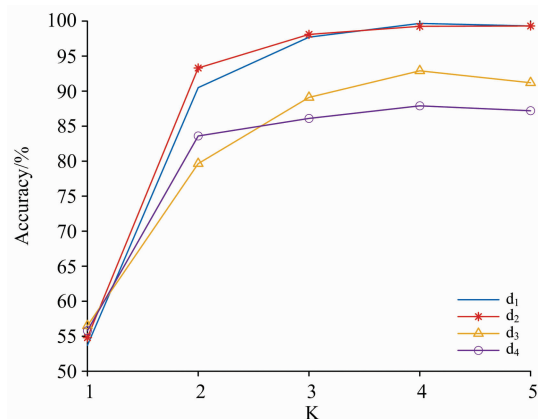


图 13 FABEMD 算法中设置不同参数的效果

Fig. 13 Accuracies of different K values in FABEMD

表 1 调整 IRCANet 参数平均准确率

Table 1 Average accuracies of different IRCANet parameters

学习率	批大小	迭代次数	平均精准度/%
0.001	4	100	97.88
0.001	4	200	98.93
0.001	6	100	99.67
0.001	6	200	99.59
0.000 1	8	100	99.62

利用识别精准度作为度量对结合了注意力机制的改进残差模块、分阶段学习残差模块和传统残差模块进行比较。表 3 描述了不同光谱组合在三种残差模块下的识别精准度, 图 15 表示在融合所有光谱的条件下, 各残差模块最好的效果。可以发现, 本文所提出的结合了注意力机制的改进残差模块识别精度最高。对于任何谱带组合, 结合了注意力机制的改进残差模块对高低频信息地有效区分使得识别效果优于其他残差模块。

表 2 多光谱组合在不同融合策略下的精准度

Table 2 Accuracies of multispectral combinations under different fusion strategies

融合光谱带	精准度/%		
	直接加权融合	红外背景重建	红外背景重建及特征细化
Red, NIR	92.04	95.08	99.49
Green, NIR	90.63	93.35	99.22
Blue, NIR	91.72	94.41	99.37
Red, Green, NIR	91.12	94.92	99.05
Blue, Red, NIR	90.76	95.01	99.45
Blue, Green, NIR	90.68	93.77	99.19
Blue, Green, Blue, NIR	90.92	95.12	99.67

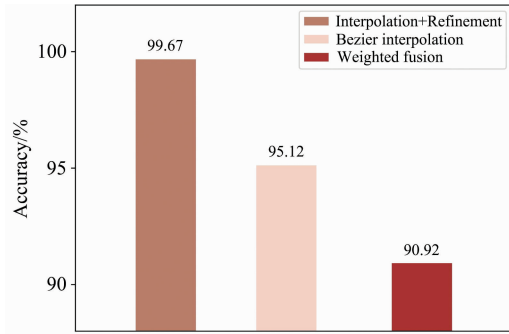


图 14 多光谱 (R, G, B, N) 掌纹图像在不同融合策略下最优效果

Fig. 14 Accuracies of different fusion strategies for multispectral (R, G, B, N) palmprint images

表 3 多光谱组合在不同残差结构下的精准度

Table 3 Accuracies of multispectral combinations under different residual structures

融合光谱带	精准度率/%		
	传统残差结构	分阶段残差结构	结合注意力机制分阶段残差结构
Red, Blue	96.37	91.23	99.25
Green, Red	95.57	90.55	99.41
Blue, Green	93.10	89.58	99.51
Red, NIR	96.93	92.66	99.09
Green, NIR	95.03	91.11	99.22
Blue, NIR	95.70	90.99	99.37
Red, Green, Blue	96.79	92.13	99.10
Red, Green, NIR	94.71	90.60	99.05
Blue, Red, NIR	95.36	91.51	99.55
Blue, Green, NIR	96.85	91.74	99.49
Blue, Green, Blue, NIR	94.03	90.80	99.67

为进一步评估所提出的结合了注意力机制的改进残差网络 (IRCANet), 与近年来主流网络 GoogleNet^[15], DenseNet^[16] 和 VggNet^[17] 对融合后的多光谱掌纹图像的分类识别作比较, 从图 16 可以看出, 相较于其他网络, 所提出的 IRCANet 对融合后的多光谱掌纹图像更具有针对性, 性

能更好。

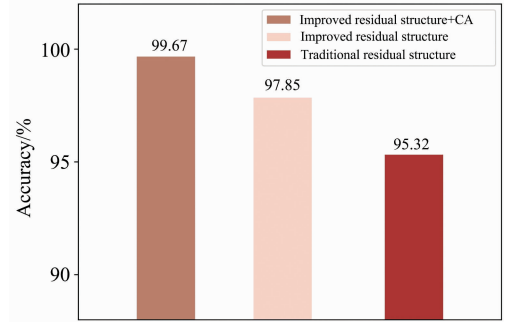


图 15 多光谱 (R, G, B, N) 掌纹图像在不同残差结构下最优效果

Fig. 15 Optimal results of different residual structures for multispectral (R, G, B, N) palmprint images

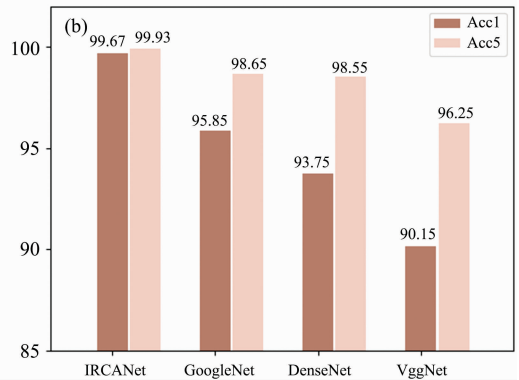
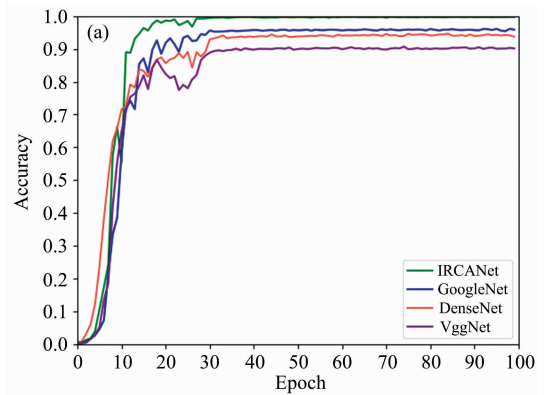


图 16 多光谱掌纹融合图像在不同主流网络下的识别效果 (a): 识别趋势; (b): 精确度对比

Fig. 16 Recognition effect of different mainstream networks for multispectral palmprint fusion images

(a): Recognition trends; (b): Accuracy contrast

将本算法与其他多光谱掌纹图像识别算法进行对比。其中, 文献[18]基于小波的图像融合方法作为数据级融合策略, 采用竞争编码方案作为匹配算法, 该算法对模糊图像是有效的, 可以很好地模拟无监控的图像捕获。文献[10]对多光谱掌纹图像的每个光谱带进行提取, 分析光谱带之间的相关性, 提出了一种分数级融合决策方案。文献[19]提出了一种基于四元数模型的多光谱图像处理方法, 利用应用主成分

分析(PCA)和离散小波变换(DWT)提取掌纹特征,利用欧几里得距离和最近邻分类器进行识别决策。文献[20]提出了一种散射网络,利用多类支持向量机和最小距离分类器进行分类。文献[21]提出一种新的联合像素级和特征级对齐(JPFA)框架进行识别。文献[22]提出了一种新的深度学习框架 Inception-ResNet。以上文献算法对所有光谱图像均使用单一算法,没有考虑到近红外光谱下掌纹图像的特殊性,在这些方法中,所提出的算法识别精度最高,具体对比如表 4 所示。

表 4 多光谱掌纹识别算法对比

Table 4 Comparison of multispectral palmprint recognition algorithms

算法	Accuracy/%
Image-level fusion by DWT ^[18]	99.03
Matching score-level fusion ^[10]	99.43
QPCA+QDWT ^[19]	98.83
Deep Scattering Network ^[20]	99.40
Joint Pixel and Feature Alignment ^[21]	99.16
Inception-ResNet ^[22]	98.74
本算法	99.67

为评估本算法性能,给出了使用该算法识别多光谱掌纹图像时每个步骤的平均执行时间。如表 5 所示,本方法对于实时应用效果较好。

表 5 本算法的时间性能

Table 5 Time cost of the proposed method

算法步骤	平均运行时间/s
FABEMD	0.175×4
近红外背景重建及特征细化	0.186
图像融合	0.002 1
分类识别	0.001 5
总的识别时间	0.889 6

3 结 论

基于多尺度变换的多光谱掌纹图像融合算法。通过在像素级进行图像融合,将多个光谱波段(蓝色、绿色、红色和近红外)的信息结合起来。利用 FABEMD 在单个光谱带捕获的每幅图像分解为几个 BIMF 和一个残差余量。为了保证融合空间中图像的最终识别精度尽可能高,考虑到近红外光谱的特性,提出对分解求和后的近红外掌纹图像利用 Bezier 插值背景重建及特征细化,再将所有处理好的图像进行加权融合。在此基础上,利用结合了注意力机制的改进残差网络进行特征提取和分类,引入的残差块能够在减少信息丢失的同时更关注通道间的特征信息,能有效区分出掌纹融合图像中的高低频信息。在 PolyU 多光谱掌纹数据库上进行的实验表明,本方法可以获得非常有竞争力的结果。此外,在不同光照下的多光谱掌纹图像融合效果还可以进一步探索,在光照条件比较极端时,本算法仍有提升空间。

References

- [1] Jain Anil K, Flynn Patrick Joseph, Ross Arun Abraham (Eds). Handbook of Biometrics. Springer, USA, 2008.
- [2] Kong A, Zhang D, Kamel M. Pattern Recognition, 2009, 42(7): 1408.
- [3] Wang J G, Yau W Y, Suwandy A, et al. Pattern Recognition, 2008, 41(5): 1514.
- [4] Guo Z, Zhang D, Zhang L. IEEE Transactions on Information Forensics and Security, 2012, 7(3): 1094.
- [5] Xu Y, Fan Z, Qiu M, et al. Neurocomputing, 2013, 103: 164.
- [6] Hao Y, Sun Z, Tan T. Proceedings of International Conference on Asian Conference on Computer Vision. 2007, Part 11: 12.
- [7] MENG Xiang-chao, SHEN Huan-feng, ZHANG Hong-yan, et al(孟祥超, 沈焕锋, 张洪艳, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2014, 34(5): 1332.
- [8] Xu X, Guo Z, Song C, et al. Sensors, 2012, 12(4): 4633.
- [9] Khan Z, Mian A, Hu Y. In Proceedings of 2011 International Conference on Computer Vision (ICCV 2011), 2011, 1935.
- [10] Zhang D, Guo Z, Gong Y. IEEE Transactions on Instrumentation and Measurement, 2010, 59(2): 480.
- [11] WANG Yan-xia, RUAN Qiu-qi(王艳霞, 阮秋琦). Chinese Journal of Image and Graphics(中国图象图形学报), 2008, 8(6): 1115.
- [12] Bhuiyan Sharif M A, Reza R Adhami, Jesmin F Khan. EURASIP Journal on Advances in Signal Processing, 2008, 32(7): 1.
- [13] Demir B, Ertürk S. IEEE Transactions on Geoscience and Remote Sensing, 2010, 48(11): 4071.
- [14] He K, Zhang X, Ren S, et al. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 770. <https://doi.org/10.1109/CVPR.2016.90>.
- [15] Szegedy C, Liu W, Jia Y, et al. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, 1. <https://doi.org/10.1109/CVPR.2015.7298594>.
- [16] Huang G, Liu Z, Van Der Maaten L J P, et al. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, 4700. <https://doi.org/10.1109/CVPR.2017.243>.
- [17] Simonyan K, Zisserman A. 2014 Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014, ArXiv preprint: 1409.
- [18] Han D, Guo Z, Zhang D. Multispectral Palmprint Recognition Using Wavelet-Based Image Fusion, In Proceedings of the IEEE International Conference on Signal Processing (ICSP), 2008, 2074.
- [19] Xu X, Guo Z, Song C, Li Y. Sensors, 2012, 12(4): 4633.

- [20] Minaee S, Wang Y. Subspace Learning in the Presence of Sparse Structured Outliers and Noise. In 2017 IEEE International Symposium on Circuits and Systems (ISCAS), 2017, 1.
- [21] Shao H, Zhong D. IEEE Transactions on Image Processing, 2021, 30(2021): 3764.
- [22] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-V4, Inception-ResNet and the Impact of Residual Connections on Learning, Thirty-First AAAI Conference on Artificial Intelligence, 2017, 1.

Palmprint Recognition Method Based on Multispectral Image Fusion

XU Xue-bin^{1,2}, XING Xiao-min^{1,2*}, AN Mei-juan^{1,2}, CAO Shu-xin^{1,2}, MENG Kan^{1,2}, LU Long-bin^{1,2}

1. The Department of Data Science and Big Data Technology, Xi'an University of Posts & Telecommunications, Xi'an 710121, China
2. Shaanxi Key Laboratory of Network Data Analysis and Intelligent Processing, Xi'an University of Posts & Telecommunications, Xi'an 710121, China

Abstract Biometric identification plays an important role in the field of information security. As a new biometric identification method, Palmprint identification has the advantages of low distortion, non-invasiveness and high uniqueness. Traditional palmprint research mostly uses natural light imaging systems to acquire in grayscale format, and it is not easy to improve the recognition accuracy further. In order to obtain more identification information, a multispectral palmprint image is proposed to replace the natural light palmprint image. Aiming at the problem that the existing palmprint recognition algorithms do not consider the characteristics of different spectra, resulting in loss of texture details and low recognition accuracy, a palmprint recognition algorithm based on multi-spectral image fusion is proposed. This method decomposes the multispectral palmprint image into a series of two-dimensional intrinsic mode functions (BIMF) with frequencies from high to low and a Residual component, which can be regarded as a preliminary estimate of the low-frequency information of the spectral image. Since the illumination conditions are unstable during the image acquisition process, and the near-infrared spectral image is sensitive to illumination transformation during FABEMD decomposition, it is easy to cause the decomposed BIMF background information to be too redundant. Therefore, the background reconstruction of the decomposed near-infrared palmprint image is performed. And feature refinement, which effectively enhances the feature expression of high-frequency information while smoothing the background redundant information. In order to avoid the problem of image overexposure caused by the spectral information after direct fusion processing, it is proposed to compress the near-infrared features before fusion. In addition, an improved residual network (IRCANet) combined with an attention mechanism is proposed for palmprint image classification after fusion, and a staged residual structure is introduced into the network to alleviate the degradation problem of the network. For the fused multispectral palmprint image, the staged residual structure can stably transmit image information between networks, but the effect of distinguishing high and low-frequency information in the image is not significant enough. In order to make the network pay attention to More discriminative features, use the interdependence between feature channels and incorporate a channel attention mechanism in the staged residual structure. Finally, comprehensive experiments on the multispectral palmprint dataset of the Hong Kong Polytechnic University (PolyU) show that the method can achieve good results, and the algorithm recognition accuracy can reach 99.67% and has good real-time performance.

Keywords Multiscale decomposition; Image fusion; Multispectral palmprint recognition; Channel attention mechanism

(Received May 8, 2022; accepted Jul. 22, 2022)

* Corresponding author