# Learning Multilayer Feature Projection for Homogeneous and Heterogeneous Palmprint Recognition

Lunke Fei, *Senior Member, IEEE*, Kaiting Huang, Shuping Zhao, *Member, IEEE*, Qi Zhu, Bob Zhang, *Senior Member, IEEE*, and Wei Jia, *Member, IEEE*

***Abstract*—Owing to its remarkable convenience, weak invasiveness, and strong private security, palmprint recognition has become one of the most promising biometric methods and has attracted increasing attention in both academia and industry. Although considerable recognition performance has been achieved by existing palmprint learning methods, they generally require the use of substantial labeled datasets and involve substantial computational overhead for feature learning. In this article, we propose a novel multilayer projection learning (MLPL) method to achieve efficient palmprint feature learning and recognition. First, we transform the palmprint images into their direction-specific representations by computing the difference in the multiple directional responses. Then, we learn three layers of feature projections for robust feature learning, including low-rank projection for image noise decoupling, feature projection for discriminative feature exploration, and quantization projection for information preservation during feature encoding. With multilayer feature projections, palmprint images can be transformed into discriminative feature representations through a single-step process for efficient palmprint recognition. Moreover, we extend the proposed MLPL, referred to as E-MLPL, by minimizing the representation discrepancy between heterogeneous palmprint images to make it applicable for heterogeneous palmprint recognition. The results obtained from five widely adopted databases confirm the superior performance of the proposed method in terms of both accuracy and efficiency.**

***Index Terms*—Discriminative directional feature, heterogeneous palmprint recognition, homogeneous palmprint recognition, multilayer projection learning (MLPL).**

Received 2 November 2025; accepted 20 December 2025. Date of publication 1 January 2026; date of current version 19 March 2026. This work was supported in part by the National Natural Science Foundation of China under Grant 62576112 and Grant 62476077; in part by the Natural Science Foundation of Beijing, China, under Grant 4252002; and in part by the Science and Technology Development Fund (FDCT), Macao, under Grant 0028/2023/RIA1. This article was recommended by Associate Editor X. Peng. *(Corresponding author: Bob Zhang.)*

Lunke Fei, Kaiting Huang, and Shuping Zhao are with the School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China (e-mail: flksxm@126.com; huangkt2000@163.com; yb77458@connect.um.edu.mo).

Qi Zhu is with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China (e-mail: zhuqi@nuaa.edu.cn).

Bob Zhang is with the PAMI Research Group, Department of Computer and Information Science, Faculty of Science and Technology, and the Centre for Artificial Intelligence and Robotics, Institute of Collaborative Innovation, University of Macau, Taipa, Macau, China (e-mail: bobzhang@um.edu.mo).

Wei Jia is with the School of Computer and Information, Hefei University of Technology, Hefei 230002, China (e-mail: china.jiawei@139.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2025.3647704.

Digital Object Identifier 10.1109/TSMC.2025.3647704

## Nomenclature

| | |
|---|---|
| $X^m \in R^{d\times N}$ | Training DRD data of the $m$th modality. |
| $X \in R^{d\times MN}$ | All training DRD data of $M$ modalities. |
| $\hat{X}^m \in R^{d\times N}$ | Discriminative features of the $m$th modality. |
| $Z^m \in R^{N\times N}$ | Low-rank representation (LRR) of the $m$th modality. |
| $Z \in R^{MN\times MN}$ | Block-diagonal LRR of all $Z^m$. |
| $E^m \in R^{d\times N}$ | Noise matrix of $X^m$. |
| $E \in R^{d\times MN}$ | Concatenated noise matrix from all $E^m$. |
| $W \in R^{d\times d}$ | Common discriminative feature projection matrix. |
| $P \in R^{d\times d}$ | Common feature rotation matrix for quantization. |
| $B \in R^{d\times N}$ | Common binary code representation. |
| $H^m \in R^{N\times N}$ | Auxiliary variable for $Z^m$. |
| $H \in R^{MN\times MN}$ | Block-diagonal auxiliary matrix of all $H^m$. |
| $C^m \in R^{d\times N}$ | Lagrange multiplier for constraint $X^m = X^m Z^m + E^m$. |
| $C \in R^{d\times MN}$ | Concatenated Lagrange multiplier from all $C^m$. |
| $D^m \in R^{N\times N}$ | Lagrange multiplier for constraint $Z^m = H^m$. |
| $D \in R^{MN\times MN}$ | Block-diagonal Lagrange multipliers of all $D^m$. |
| $\mathbf{1} \in R^{MN\times 1}$ | All-one column vector. |
| $\Xi \in R^{N\times N}$ | Constant matrix with entries $\frac{1}{MN}$. |
| $\beta,\ \mu$ | Penalty parameters in augmented Lagrangian. |
| $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ | Weighting parameters in objective function. |
| $d$ | Dimension of DRD vector. |
| $N$ | Number of DRD vectors per modality. |
| $M$ | Total number of heterogeneous modalities. |

## I. Introduction

OVER the years, biometrics have been widely applied for personal authentication due to their uniqueness and reliability [1]. As one of the relatively new biometric traits, palmprints can be easily imaged with the cooperation of humans due to the easy self-positioning of hands, while they are hard to capture in images without one's authorization due to their strong privacy, and moreover, palmprints show less invasiveness than conventional biometrics such as faces [2]. As a result, palmprint recognition has become one of the most promising biometric directions in both academia and industry. An increasing number of publications have been published in recent years [3], and several palmprint-based e-payment systems have been developed, such as Amazon One [4] and Tencent WePalm [5].

In the early years, palmprint recognition methods usually focused on hand-crafted feature engineering, primarily including lines, texture, and direction-based palmprint descriptors. For example, Huang et al. [6] extracted palmprint principal lines based on the modified finite radon transform (MFRAT) for palmprint verification, and Wang et al. [7] proposed a 2-D Gabor wavelet-based texture representation for palmprint verification. Inspired by the robustness of line directions, Zhang et al. [8] and Xu et al. [9] used Gabor filters to obtain salient orientation information from palmprints, and a number of multidirection palmprint feature representations were designed, such as local multiple directional pattern (LMDP) [10] and apparent and latent direction code (ALDC) [11]. In addition, Bounneche et al. [12] utilized a fusion strategy involving hand-crafted features across diverse spectral bands to achieve better palmprint recognition. Most hand-crafted palmprint features can achieve encouraging recognition performance. However, they usually require substantial prior knowledge for feature engineering, and the resulting performance is often restricted to specific scenarios, making it difficult to adaptively represent palmprint images captured under diverse conditions.

In recent years, a growing body of research has focused on developing learning-based methods that aim to adaptively learn palmprint descriptors. For example, Zhao and Zhang [13] learned a salient and discriminative palmprint descriptor based on LSR and LRR for general scenario palmprint recognition, Fei et al. [14] developed a single-layer convolution network (SLCN) aimed at adaptively capturing discriminative features common to various palmprint images, and Li et al. [15] designed an iterative optimization model via supervised learning to generate robust binary palmprint codes across different views. In addition, Zhao et al. [16] proposed the multiview hierarchical graph learning-based palmprint recognition method (MVHG_PR) by leveraging multiview hierarchical graph learning to handle hand rotations, noise, and low illumination. More recently, the superior feature extraction capabilities of neural networks have prompted a surge in deep learning-based approaches for palmprint recognition in the literature. For example, Fan et al. [17] introduced a Gabor convolution network to facilitate the extraction of structurally salient and distinctive features from palmprint images via multilevel Gabor feature fusion, and Fei et al. [18] developed a Fourier-based palmprint feature learning network for cross-spectra palmprint recognition by reducing feature differences in the frequency domain. Zhong et al. [19] proposed a palmprint registration framework that unifies palmprint orientations and learns pairwise spatial registration to mitigate pattern variance. Chai et al. [20] designed a contactless palmprint recognition system by enhancing feature discriminability through recurrent layer aggregation and angular center-neighbor loss. More details can be found in the latest survey paper on deep learning-based palmprint recognition [21]. Despite their impressive performance in palmprint recognition, learning-based methods are inherently constrained by their dependence on large quantities of labeled training data [22]. The inherently sensitive nature of palmprint biometrics makes the collection and annotation of such datasets expensive and impractical. How to efficiently capture the discriminative features from diverse palmprint classes with limited labeled training samples remains a challenging problem for palmprint recognition.

In this article, we propose a multilayer projection learning (MLPL) method to learn robust palmprint features for recognition across both homogeneous and heterogeneous scenarios. The basic idea of the proposed method is illustrated in Fig. 1. First, we extract the directional intensity information of the training palmprint images via Gabor convolution and further form directional response difference (DRD) maps to explore the inherent palmprint information representation. Then, we learn three layers of feature projections to convert the DRD maps into discriminative binary feature codes, comprising an LRR mapping for noise decoupling from the raw palmprint samples, a feature mapping to project noise-free palmprint representation into discriminative feature space, and a quantization mapping to minimize the information loss during feature encoding. Moreover, we extend the proposed MLPL method by jointly learning the modality-invariant palmprint discriminative features, making our method applicable to heterogeneous palmprint recognition. Extensive experiments on five widely used palmprint databases are conducted to demonstrate the effectiveness and efficiency of the proposed method.

The main contributions of this article can be summarized as follows.

1) We propose a three-layer projection learning (MLPL) method for palmprint recognition that first corrects palmprint image noise and then explores discriminative palmprint features. By integrating low-rank denoising, discriminative feature enhancement, and binary quantization in a unified framework, MLPL efficiently constructs a compact binary feature space with significantly fewer parameters than most existing learning networks, enabling robust recognition with limited training samples.
2) We develop a more general MLPL model to make the proposed method applicable to heterogeneous palmprint recognition. Using the extended MLPL (E-MLPL) method, the modality gap between heterogeneous palmprint images is effectively reduced at the feature level, providing a more robust and generalizable solution for cross-device, cross-lighting, and cross-spectrum scenarios.
3) To validate the effectiveness of the proposed MLPL and E-MLPL methods, we conducted comprehensive

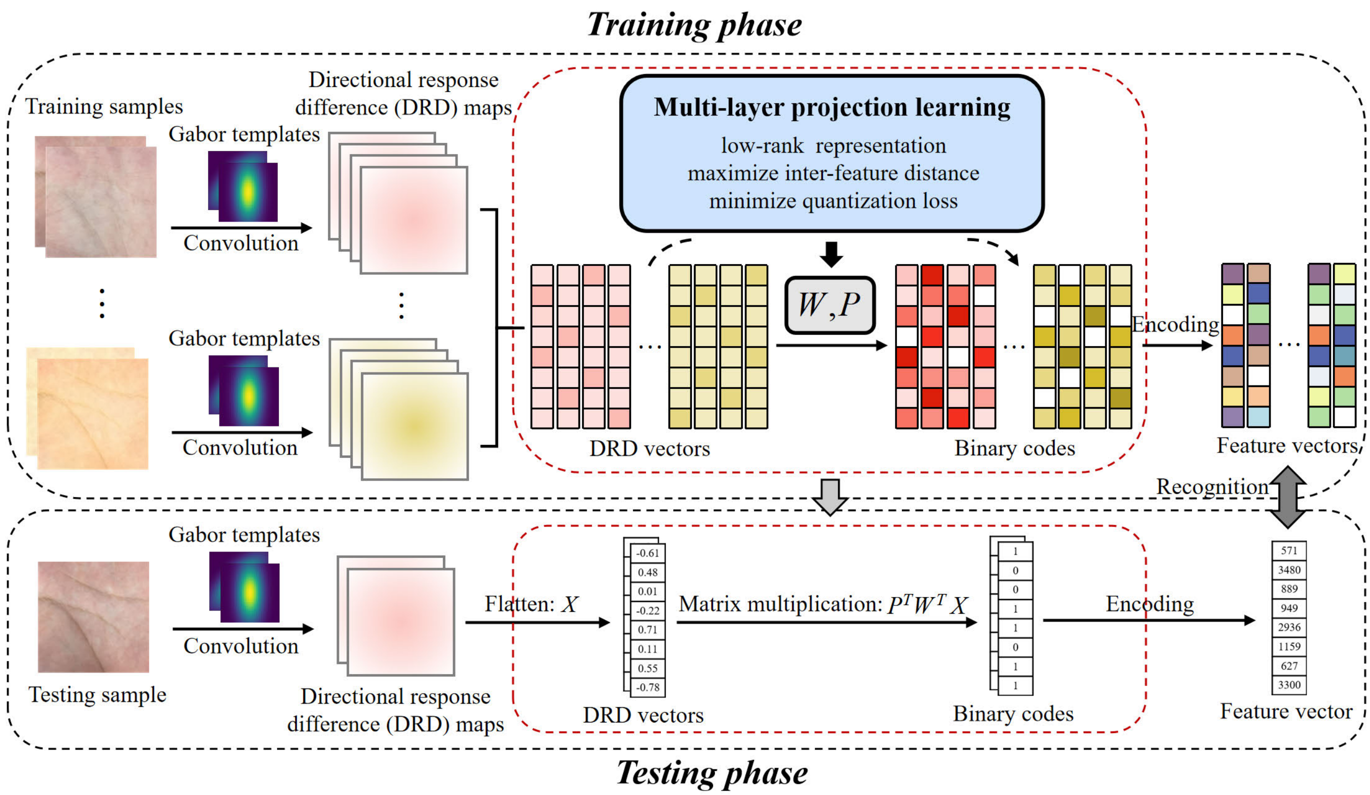



Fig. 1. Overall procedure of the proposed method. We first compute DRD maps of the training palmprint images by convolving them with a bank of Gabor templates with diverse directions. Then, we learn multiple layers of feature projections that jointly convert the DRD maps into discriminative binary feature codes for robust palmprint recognition.

experiments on five benchmark palmprint databases. The results clearly demonstrate that both methods achieve outstanding performance across homogeneous and heterogeneous palmprint recognition tasks, confirming their robustness and efficiency.

The rest of this article is structured as follows. Section II provides a brief review of three related topics. Section III details the proposed MLPL method for homogeneous palmprint recognition, while Section IV further extends the MLPL method for application in heterogeneous palmprint recognition. Section V presents the experimental results. Finally, Section VI concludes this article.

## II. Related Topics

In this section, we provide a brief overview of three relevant topics, including homogeneous and heterogeneous palmprint recognition, DRD maps of palmprint images, and LRR.

### A. Homogeneous and Heterogeneous Palmprint Recognition

Homogeneous palmprint recognition refers to matching palmprint images captured with similar types of imaging devices in a stable and uniform environment, which generally exhibit consistent styles, similar visual characteristics, and relatively small intraclass variations. Early palmprint recognition studies primarily focused on such scenarios, leading to the establishment of several homogeneous databases, such as the PolyU [23], IITD [24], and TJU [25] databases. Due to similar imaging conditions, homogeneous palmprint images usually show consistent image styles with relatively small intraclass feature variances. As a result, it is relatively easy to achieve high recognition accuracy for homogeneous palmprint recognition [16]. In contrast, heterogeneous palmprint recognition involves the comparison of diverse styles of palmprint images with relatively large feature gaps that are collected under different imaging conditions, such as using diverse imaging devices and under the environments of varying illuminations. Heterogeneous images exhibit pronounced modality discrepancies, large intraclass variations, and diverse visual quality. In recent years, research on heterogeneous palmprint recognition, such as cross-domain [26], cross-illumination [27], and cross-dataset [28] palmprint recognition, has attracted significant attention, leading to the establishment of several heterogeneous palmprint databases, such as the XJTU-UP [29] and CASIA multispectral palmprint databases [30]. While most existing methods perform well on homogeneous datasets, their effectiveness deteriorates under heterogeneous conditions due to substantial modality and domain gaps, and most approaches lack mechanisms to explicitly compensate for such differences [16]. How to extract the common personal-invariant features of heterogeneous palmprint images to minimize intrapalm variation while maximizing interpalm margin remains a challenging problem. In this study, we propose a new and efficient MLPL framework for both homogeneous and heterogeneous palmprint recognition.

### B. DRD Maps of Palmprint Images

Inspired by the fact that a palmprint contains rich informative direction information [8], many methods [31] exploit direction-specific representation discriminative palmprint feature learning by calculating the difference in the directional response of palmprint images. Specifically, motivated by the strong palmprint direction modeling capability of the Gabor filters [12], a bank of Gabor-based templates, referred to as $G(\theta_i)(\theta_i = (i-1)\pi/N_\theta)$, is first employed to convolve with the palmprint image, obtaining the palmprint multiple directional response representation, as follows:

$$r_i = I * G(\theta_i) \tag{1}$$

where "$*$" represents the convolution operation and $r_i$ denotes the directional response of the palmprint image (i.e., $I$) in the direction $\theta_i$. $N_\theta$ denotes the total number of Gabor templates, with a typical value of 12 [32], such that the directional interval of two neighboring Gabor-based templates is $\pi/12$. After that, the direction-response difference between each pair of adjacent directions is calculated as follows:

$$\mathrm{DRD}_i = r_i - r_{i \bmod N_\theta + 1} = I * \left(G(\theta_i) - G\left(\theta_{i \bmod N_\theta + 1}\right)\right) \tag{2}$$

where ($i$ mod $N_\theta + 1$) actually denotes the next neighboring direction index of $\theta_i$, and $\mathrm{DRD}_i$ denotes the DRD map between the directions of $\theta_i$ and $\theta_{i \bmod N_\theta+1}$. In this way, multiple DRD maps with intensity knowledge can be obtained for a palmprint image, which can not only measure the directional intensity of each pixel in multiple orientations but also implicitly describe how intensity changes along neighboring directions.

### C. Low-Rank Representation

LRR aims to uncover the latent structures in high-dimensional data by recovering the intrinsic low-dimensional subspace embedded within noisy sample sets, making it effective for tasks such as subspace clustering, dimensionality reduction, and robust feature extraction. For example, Liu and Yan [33] proposed latent LRR (LatLRR) for subspace clustering by recovering the unobserved data hidden using two low-rank matrices. Fu et al. [34] leveraged spatial structure analysis alongside tensor low-rank modeling to advance subspace clustering solutions, and Yang et al. [35] designed a sparse LRR for incomplete multiview clustering. In addition, He et al. [36] proposed an LRR decomposition method for image recovery by decomposing the original images into sparse low-rank matrices, and Li et al. [37] introduced an LRR-based approach enhanced with supervised LDA for discriminative feature extraction. In this study, we propose recovering noise-free information from raw palmprint images via LRR constraint mapping to better exploit discriminative features for diverse homogeneous and heterogeneous palmprint images.

## III. MLPL for Homogeneous Palmprint Recognition

In this section, we begin by formulating the objective function of our proposed multistage feature projection. Then, we detail the optimization of the MLPL. Finally, we show how to use the MLPL for homogeneous palmprint recognition. For a better readability, we provide the notations used in this study in the Nomenclature.

### A. MLPL Objective Function

Our proposed MLPL aims to learn multiple layers of hash functions that orderly map the DRD maps of a palmprint image into a strong discriminative feature representation. Specifically, let $X = [x_1, x_2, \ldots, x_n] \in R^{d\times N}$ be the DRD set of training palmprint images containing $N$ vectorized DRD, where $x_i \in R^{d\times 1}$ is the $k$th DRD vector and $d$ denotes the size of the DRD vector. It is noted that palmprint images in real-world applications are often affected by noise, uneven lighting, and background interference. We design the first layer of the MLPL framework to correct such degradations via LRR [38] as follows:

$$\begin{aligned} &\min \|Z\|_* + \lambda \|E\|_F^2 \\ &\text{s.t.} \quad X = XZ + E,\ \operatorname{diag}(Z) = 0,\ E \geq 0 \end{aligned} \tag{3}$$

where $\|.\|_*$ denotes the nuclear-norm constraint to compute the rankness of a matrix, and $Z$ represents the LRR of the training set $X$. $E$ denotes the error information matrix decomposed from the training samples. $\lambda$ is a balance parameter. By minimizing the rankness of $Z$, we aim to capture the clean palmprint image information with as little noise impact as possible. $\operatorname{diag}(Z) = 0$ is imposed such that the diagonal elements of $Z$ are zeros, overcoming the problem of self-representation [39]. $E \geq 0$ makes the error elements nonnegative.

Having obtained the corrected training palmprint image information, i.e., $XZ$, we further learn a feature projection function in the second layer to convert them (i.e., $XZ$) into a discriminative feature space, as follows:

$$\hat{X} = W^T XZ \tag{4}$$

where $W^T \in R^{d\times d}$ denotes the discriminative feature projection function and $\hat{X} \in R^{d\times N}$ represents the discriminative feature representation of the palmprint images. To make the palmprint images discriminative in the feature space, we enlarge the feature variances as follows:

$$\begin{aligned} &\max \left\| \hat{X} - \bar{\hat{X}} \right\|_F^2 \\ &\qquad \Rightarrow \max \left\| W^T XZ\text{-}W^T XZ\Xi \right\|_F^2 \\ &\text{s.t.} \quad W^T W = I \end{aligned} \tag{5}$$

where $\bar{\hat{X}} \in R^{d\times N}$ denotes the mean feature matrix containing $N$ column vectors and where each vector is the mean of the $N$ feature vectors in $\hat{X}$. $\Xi = 1/N 1 1^T \in R^{N\times N}$ and 1 is an all-one column vector such that $\Xi$ is a matrix with all the elements being $1/N$. Hence, the mean feature matrix $\bar{\hat{X}}$ can be calculated as: $W^T XZ\Xi$. The constraint $W^T W = I$ is to make the basic discriminative feature projection vectors independent.

Motivated by the compactness and inherent robustness of binary representations to minor illumination variations [40], we project the discriminative palmprint features into binary codes through a third-layer quantization mapping, minimizing binarization error and preserving feature information to generate compact and efficient representations for storage and matching. To do this, we formulate the following binary feature encoding objective function:

$$\begin{aligned} &\min \left\| B - P^T W^T X Z \right\|_F^2 \\ &\text{s.t.} \quad B \in \{0,1\}^{d\times N} \end{aligned} \tag{6}$$

where $B \in \{0,1\}^{d\times N}$ represents the binary code representation of the learned features. $P^T \in R^{d\times d}$ denotes a feature rotation function, which makes the quantization loss between the original features and the encoded binary codes as small as possible.

To achieve the above goals, the overall objective function for our feature learning framework is defined as follows:

$$\begin{aligned} \min \, & \left\| B - P^T W^T X Z \right\|_F^2 - \lambda_1 \left\| W^T X Z\text{-}W^T X Z \Xi \right\|_F^2 \\ & + \lambda_2 \|Z\|_* + \lambda_3 \|E\|_F^2 + \lambda_4 \|P\|_F^2 \\ \text{s.t.} \quad & X = XZ + E, \quad W^T W = I, \quad B \in \{0,1\}^{d\times N} \\ & \operatorname{diag}(Z) = 0, \quad E \geq 0 \end{aligned} \tag{7}$$

where $\lambda_1$, $\lambda_2$, $\lambda_3$, and $\lambda_4$ are the balance parameters.

### B. MLPL Optimization

To obtain the optimal three-layer feature projection functions of our proposed MLPL, we employ the widely used alternating direction strategy [13] to optimize the multiple variables of the objective function (7). Specifically, we first introduce an auxiliary variable $H \in R^{N\times N}$ to separate the multiple $Z$ in the objective function (7) as follows:

$$\begin{aligned} \min \, & \left\| B - P^T W^T X Z \right\|_F^2 - \lambda_1 \left\| W^T X Z\text{-}W^T X Z \Xi \right\|_F^2 \\ & + \lambda_2 \|H\|_* + \lambda_3 \|E\|_F^2 + \lambda_4 \|P\|_F^2 \\ \text{s.t.} \quad & X = XZ + E, \;\; Z = H, \;\; \operatorname{diag}(Z) = 0, \;\; W^T W = I \\ & B \in \{0,1\}^{d\times N}, \;\; E \geq 0. \end{aligned} \tag{8}$$

**Algorithm 1** MLPL

**Input:** The training samples: $X \in R^{d\times N}$; Parameters: $\lambda_1$, $\lambda_2$, $\lambda_3$, and $\lambda_4$; Maximum number of iterations: $R$.

**Initialization:** $Z$, $H$, $E$, $W$, $P$, $B$, $C$, and $D$ with random values; $\beta = 0.1$; $\beta_{\max} = 1e+5$; and $\eta = 1.01$.

**for** *iteration* $< R$ **do**

- Update $Z$ by using (12);
- Update $H$ by using (14);
- Update $E$ by using (16);
- Update $W$ by using the GPI algorithm to optimize (17);
- Update $P$ by using (19);
- Update $B$ by using (21);
- Update $C$, $D$, and $\beta$ by using (22);

**Output:** $W$ and $P$.

Then, we convert (8) into its augmented Lagrange function form, which can be expressed as follows:

$$\begin{aligned} \min \, & \left\| B - P^T W^T X Z \right\|_F^2 - \lambda_1 \left\| W^T X Z\text{-}W^T X Z \Xi \right\|_F^2 \\ & + \lambda_2 \|H\|_* + \lambda_3 \|E\|_F^2 + \lambda_4 \|P\|_F^2 \\ & + \langle C, X - XZ - E \rangle + \langle D, Z - H \rangle \\ & + \frac{\beta}{2}\left( \|X - XZ - E\|_F^2 + \|Z - H\|_F^2 \right) \\ \text{s.t.} \quad & \operatorname{diag}(Z) = 0, \, W^T W = I, \, B \in \{0,1\}^{d\times N}, \;\; E \geq 0 \end{aligned} \tag{9}$$

where $C$ and $D$ are the Lagrange multipliers. $\langle ., . \rangle$ denotes the inner product calculation, and $\beta > 0$ is a penalty parameter. By relaxing the inner product operation, we rewrite (9) into the following form:

$$\begin{aligned} \min \, & \left\| B - P^T W^T X Z \right\|_F^2 - \lambda_1 \left\| W^T X Z - W^T X Z \Xi \right\|_F^2 \\ & + \lambda_2 \|H\|_* + \lambda_3 \|E\|_F^2 + \lambda_4 \|P\|_F^2 \\ & + \frac{\beta}{2}\left( \left\| X - XZ - E + \frac{C}{\beta} \right\|_F^2 + \left\| Z - H + \frac{D}{\beta} \right\|_F^2 \right) \\ & - \frac{1}{2\beta}\left( \|C\|_F^2 + \|D\|_F^2 \right) \\ \text{s.t.} \quad & \operatorname{diag}(Z) = 0, \, W^T W = I, \, B \in \{0,1\}^{d\times N}, \;\; E \geq 0. \end{aligned} \tag{10}$$

After that, we alternately update each variable by temporarily fixing the other ones, and the details are as follows.

*Step 1 (Update Z):* By fixing $H$, $E$, $W$, $P$, $B$, $C$, and $D$, but not $Z$, we rewrite the objective function (10) as follows:

$$\begin{aligned} \min \, & \left\| B - P^T W^T X Z \right\|_F^2 - \lambda_1 \left\| W^T X Z\text{-}W^T X Z \Xi \right\|_F^2 \\ & + \frac{\beta}{2}\left( \left\| X - XZ - E + \frac{C}{\beta} \right\|_F^2 + \left\| Z - H + \frac{D}{\beta} \right\|_F^2 \right) \\ \min \operatorname{tr} & \left( B B^T - 2 Z^T X^T W P B + Z^T X^T W P P^T W^T X Z \right) \\ & - \lambda_1 \operatorname{tr}\left( W^T X Z Z^T X^T W - 2 W^T X Z \Xi H^T X^T W \right. \\ & \left. + W^T X Z \Xi \Xi Z^T X^T W \right) \\ & + \frac{\beta}{2}\operatorname{tr}\left( U U^T - 2 U X Z^T + Z^T X^T X Z \right) \\ & + \frac{\beta}{2}\operatorname{tr}\left( Z Z^T - 2 Z V^T + V V^T \right) \\ \text{s.t.} \quad & \operatorname{diag}(Z) = 0 \end{aligned} \tag{11}$$

where $U = X - E + C/\beta$, $V = H - D/\beta$. Let the derivative of the objective function (11) with respect to (w.r.t.) $Z$ be zero, we can obtain the solution of $Z$ as follows:

$$\begin{aligned} Z = & \left( X^T W P P^T W^T X - \lambda_1 X^T X + \lambda_1 \Xi X^T X \right. \\ & \left. - \lambda_1 \Xi \Xi X^T X + \frac{\beta}{2} X^T X + \frac{\beta}{2} I \right)^{-1} \\ & \left( X^T W P B + \beta X^T U + \beta V \right) \\ \text{s.t.} \quad & \operatorname{diag}(Z) = 0 \\ & \Rightarrow Z = Z - \operatorname{diag}(Z). \end{aligned} \tag{12}$$

*Step 2 (Update H):* Fixing $Z$, $E$, $W$, $P$, $B$, $C$, and $D$, but not $H$, (10) is transformed into the optimization problem described below

$$\min \lambda_2 \|H\|_* + \frac{\beta}{2}\left\|Z - H + \frac{D}{\beta}\right\|_F^2. \tag{13}$$

Then, we can directly optimize $H$ via the singular-value thresholding (SVT) operator [38] as follows:

$$H = \Theta_{\frac{\lambda_2}{\beta}}\left(Z + \frac{D}{\beta}\right) \tag{14}$$

where $\Theta$ represents the SVT operator and $(\lambda_2)/(\beta)$ is the threshold value of the SVT.

*Step 3 (Update E):* We fix all the variables expect $E$ and rewrite (10) as follows:

$$\begin{aligned}&\min \lambda_3 \|E\|_F^2 + \frac{\beta}{2}\left\|X - XZ - E + \frac{C}{\beta}\right\|_F^2\\ &\text{s.t.}\quad E \geq 0.\end{aligned} \tag{15}$$

By computing the derivative of (15) w.r.t. $E$ and setting it to zero, we can optimize $E$ as follows:

$$E = \max\left(0, \frac{\beta}{2\lambda_3 + \beta}\left(X - XZ + \frac{C}{\beta}\right)\right). \tag{16}$$

*Step 4 (Update W):* By fixing the other variables, including $Z$, $H$, $E$, $P$, $B$, $C$, and $D$, but not $W$, we reformulate (10) into the following form:

$$\begin{aligned}&\min \left\|\left(P^T\right)^{-1} B - W^T XZ\right\|_F^2 - \lambda_1 \left\|W^T XZ - W^T XZ\Xi\right\|_F^2\\ &= \min \Big[\mathrm{tr}\left(\left(P^T\right)^{-1} BB^T P^{-1}\right) - 2\mathrm{tr}\left(W^T XZB^T P^{-1}\right)\\ &\quad + \mathrm{tr}\left(W^T XZZ^T X^T W\right) - \lambda_1 \mathrm{tr}\left(W^T XZZ^T X^T W\right.\\ &\quad \left. - 2W^T XZ\Xi Z^T X^T W + W^T XZ\Xi\Xi Z^T X^T W\right)\Big]\\ &= \min \mathrm{tr}\left(W^T RW - 2W^T Q\right)\\ &\text{s.t.}\quad W^T W = I\end{aligned} \tag{17}$$

where $R = XZZ^T X^T - \lambda_1 XZZ^T X^T + 2\lambda_1 XZ\Xi Z^T X^T - \lambda_1 XZ\Xi\Xi Z^T X^T$ and $Q = XZB^T P^{-1}$. After that, $W$ can be optimized by the GPI algorithm [41].

*Step 5 (Update P):* Having fixed the values of variables $Z$, $H$, $E$, $W$, $B$, $C$, and $D$, the objective function (10) with only the variable of $P$ becomes the following optimization problem:

$$\begin{aligned}&\min \left\|B - P^T W^T XZ\right\|_F^2 + \lambda_4 \|P\|_F^2\\ &\Rightarrow \min \mathrm{tr}\left(BB^T - 2P^T W^T XZB^T\right.\\ &\qquad \left. + P^T W^T XZZ^T X^T WP + \lambda_4 P^T P\right).\end{aligned} \tag{18}$$

Making the derivative of (18) w.r.t $P$ zero, we can obtain the solution of $P$ as follows:

$$P = \left(W^T XZZ^T X^T W + \lambda_4 I\right)^{-1} W^T XZB^T. \tag{19}$$

*Step 6 (Update B):* When other variables (such as $Z$, $H$, $E$, $W$, $P$, $C$, and $D$), but not $B$ are fixed, a modified version of objective function (10) is presented as follows:

$$\min \left\|B - P^T W^T XZ\right\|_F^2, s.t.\quad B \in \{0,1\}^{d \times N}. \tag{20}$$

Then, $B$ can be directly encoded on the basis of the sign of $P^T W^T XZ$ as follows:

$$B = \frac{1}{2}\left(\mathrm{sgn}\left(P^T W^T XZ\right) + 1\right). \tag{21}$$

*Step 7 (Update C, D, and β):* Finally, we fix $Z$, $H$, $E$, $W$, $P$, and $B$ and update $C$, $D$, and $\beta$ as follows [13]:

$$\begin{cases} C = C + \beta (X - XZ - E) \\ D = D + \beta (Z - H) \\ \beta = \min(\eta\beta, \beta_{\max}). \end{cases} \tag{22}$$

Algorithm 1 summarizes the basic optimization procedure of the proposed MLPL method.

### *C. Feature Representation and Matching*

After learning both the discriminative feature projection function $W$ and the feature rotation function $P$, we can easily map a new palmprint image into binary feature representations via (21). In this study, we refer to the effective block-wise histogram vector method [32] to represent the binary feature codes for palmprint recognition. Specifically, each of the $k$th binary feature maps is weighted by $2^{k-1}$, and the sum of these weighted maps is computed to obtain a decimal representation of the palmprint features. After that, we flatten the decimal feature map into a feature vector as the final palmprint feature descriptor. During the recognition stage, the commonly used distance measurements, such as Euclidean and cosine distances, can be calculated to measure the similarity between the vector-based palmprint feature descriptors.

## IV. E-MLPL for Heterogeneous Palmprint Recognition

Recently, increasing attention has been given to heterogeneous palmprint recognition because the matching palmprint images of a palm are possibly captured by different devices under different imaging conditions. In this section, we extend our proposed MLPL, referred to as E-MLPL, to a more general version, making it applicable for heterogeneous palmprint recognition.

### *A. E-MLPL Objective Function*

Based on the notation summarized in Nomenclature, let $X^m = [x_1^m, x_2^m, \ldots, x_n^m] \in R^{d\times N} (m = 1, 2, \ldots, M)$ denote the training vectored DRD sample set formed from the $m$th modality of palmprint images with $M$ different heterogeneous modalities, and let $X = [X^1, X^2, \ldots, X^M] \in R^{d\times MN}$ represent all the training DRD data of $M$ modalities. Given the significant visual differences among various palmprint image modalities, we employ LRR to decompose the noise component from the training data for each modality separately, thus obtaining the latent clean palmprint DRD information as follows:

$$\begin{aligned}&\min \left\|Z^m\right\|_* + \lambda \sum_{m=1}^{M} \left\|E^m\right\|_F^2 \\ &\text{s.t.} \quad X^m = X^m Z^m + E^m,\ \operatorname{diag}\left(Z^m\right) = 0,\ \ E^m \geq 0\end{aligned} \tag{23}$$

where $Z^m$ denotes the LRR of the $m$th modality samples and $E^m$ denotes the noise component matrix decomposed from $X^m$. Then, we learn a common discriminative feature projection function, referred to as $W$, aiming to convert each modality of palmprint images into discriminative feature representations, as follows:

$$\hat{X}^m = W^T X^m Z^m. \tag{24}$$

Following (12), to strengthen the discriminative capacity of the palmprint features for each modality, we maximize their feature variance using the following formulation:

$$\begin{aligned}&\max \left\|W^T X^m Z^m \text{-} W^T X^m Z^m \Xi\right\|_F^2 \\ &\text{s.t.} \quad W^T W = I.\end{aligned} \tag{25}$$

Finally, we convert discriminative palmprint features into binary codes for compact and robust feature representation. To make the intrapalm binary codes of palmprint images with different modalities consistent, we formulate the common binary code learning objective function as follows:

$$\begin{aligned}&\min \sum_{m=1}^{M} \left\|B - P^T W^T X^m Z^m\right\|_F^2 \\ &\text{s.t.} \quad B \in \{0, 1\}^{d\times N}\end{aligned} \tag{26}$$

where $B$ denotes a common binary code representation of palmprint images with different modalities. $P$ represents the feature rotation matrix, aiming to extensively minimize the information loss between the common binary codes and the multimodality palmprint features.

Consequently, we achieve the overall objective function of the heterogeneous palmprint feature learning model as follows:

$$\begin{aligned}\min \sum_{m=1}^{M} \Big[ &\left\|B - P^T W^T X^m Z^m\right\|_F^2 \\ &- \lambda_1 \left\|W^T X^m Z^m - W^T X^m Z^m \Xi\right\|_F^2 \\ &+ \lambda_2 \left\|Z^m\right\|_* + \lambda_3 \left\|E^m\right\|_F^2 \Big] + \lambda_4 \left\|P\right\|_F^2 \\ \text{s.t.} \quad &X^m = X^m Z^m + E^m,\ \ W^T W = I,\ \ B \in \{0, 1\}^{d\times N} \\ &\operatorname{diag}\left(Z^m\right) = 0,\ \ E^m \geq 0.\end{aligned} \tag{27}$$

### B. E-MLPL Optimization

Similar to the optimization of the MLPL, we use the alternating direction method to obtain the optimal solutions of (27). Specifically, to separate the multiple $Z^m$ in the objective function (27), we define $M$ auxiliary variable $H^m$ for $Z^m$ such that (27) can be rewritten as follows:

$$\begin{aligned}\min \sum_{m=1}^{M} \Big[ &\left\|B - P^T W^T X^m Z^m\right\|_F^2 \\ &- \lambda_1 \left\|W^T X^m Z^m - W^T X^m Z^m \Xi\right\|_F^2 \\ &+ \lambda_2 \left\|H^m\right\|_* + \lambda_3 \left\|E^m\right\|_F^2 \Big] + \lambda_4 \left\|P\right\|_F^2 \\ \text{s.t.} \quad &X^m = X^m Z^m + E^m,\ \ W^T W = I,\ \ B \in \{0, 1\}^{d\times N} \\ &\operatorname{diag}\left(Z^m\right) = 0,\ \ E^m \geq 0,\ \ Z^m = H^m.\end{aligned} \tag{28}$$

Then, we combine $M$ modalities of LRRs (i.e., $Z^m$) to form a composite block-wise LRR, referred to as $Z$, as follows:

$$\begin{aligned}Z &= \operatorname{blkdiag}\left(Z^1, Z^2, \ldots, Z^M\right) \\ &= \begin{bmatrix} Z^1 & O & \cdots & O \\ O & Z^2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & Z^M \end{bmatrix} \in R^{MN\times MN}\end{aligned} \tag{29}$$

where blkdiag denotes the matrix composite operation. Similarly, we combine the multiple auxiliary variables $H^m$ into a composite block-wise auxiliary matrix $H$ as: $H = \operatorname{blkdiag}(H^1, H^2, \ldots, H^M)$. By doing so, we can reformulate (28) into the following optimization problem:

$$\begin{aligned}\min &\left\|\hat{B} - P^T W^T X Z\right\|_F^2 - \lambda_1 \left\|W^T X Z - W^T X Z \Xi\right\|_F^2 \\ &+ \lambda_2 \|H\|_* + \lambda_3 \|E\|_F^2 + \lambda_4 \|P\|_F^2 \\ \text{s.t.} \quad &X = XZ + E,\ \ W^T W = I,\ \ \hat{B} \in \{0, 1\}^{d\times MN} \\ &\operatorname{diag}(Z) = 0,\ \ E \geq 0,\ \ Z = H\end{aligned} \tag{30}$$

where $E = [E^1, E^2, \ldots, E^M] \in R^{d\times MN}$. $\Xi = (1)/(MN) 11^T \in R^{MN\times MN}$ and all the elements in $\Xi$ are $(1)/(MN)$. $\hat{B} = [B, B, \ldots, B] \in R^{d\times MN}$ contains $M$ common binary feature codes. After that, we introduce the Lagrange multipliers $C = [C^1, C^2, \ldots, C^M] \in R^{d\times MN}$ and $D = \operatorname{blkdiag}(D^1, D^2, \ldots, D^M) \in R^{MN\times MN}$ to convert (30) into the following form:

$$\begin{aligned}\min &\left\|\hat{B} - P^T W^T X Z\right\|_F^2 - \lambda_1 \left\|W^T X Z - W^T X Z \Xi\right\|_F^2 \\ &+ \lambda_2 \|H\|_* + \lambda_3 \|E\|_F^2 + \lambda_4 \|P\|_F^2 \\ &+ \frac{\mu}{2}\left(\left\|X - XZ - E + \frac{C}{\mu}\right\|_F^2 + \left\|Z - H + \frac{D}{\mu}\right\|_F^2\right) \\ &- \frac{1}{2\mu}\left(\|C\|_F^2 + \|D\|_F^2\right) \\ \text{s.t.} \quad &\operatorname{diag}(Z) = 0, W^T W = I,\ \ \hat{B} \in \{0, 1\}^{d\times MN},\ \ E \geq 0.\end{aligned} \tag{31}$$

where $\mu > 0$ is a penalty parameter. In the following, we detail the optimization procedure for each variable of (31).

*Step 1 (Update $Z^m$, $H^m$, $E^m$, $C^m$, $D^m$, and $\mu$):* Following the optimization of MLPL, we can individually update $Z^m$,

$H^m$, $E^m$, $C^m$, $D^m$, and $\mu$. Specifically, referring to (12), we can optimize $Z^m$ as follows:

$$\begin{aligned} Z^m = &\Big( \left(X^m\right)^T WPP^TW^TX^m - \lambda_1 \left(X^m\right)^T X^m \\ &+ \lambda_1 \Xi \left(X^m\right)^T X^m - \lambda_1 \Xi\Xi \left(X^m\right)^T X^m \\ &+ \frac{\mu}{2} \left(X^m\right)^T X^m + \frac{\mu}{2} I \Big)^{-1} \\ &\Big( \left(X^m\right)^T WPB + \mu \left(X^m\right)^T U^m + \mu V^m \Big) \\ \text{s.t.}\quad & \operatorname{diag}\left(Z^m\right) = 0 \\ &\Rightarrow Z^m = Z^m - \operatorname{diag}\left(Z^m\right) \end{aligned} \tag{32}$$

where $U^m = X^m - E^m + (C^m)/(\mu)$, $V^m = H^m - (D^m)/(\mu)$. Similarly, by referring (14), (16), and (22), we can compute the solutions of $H^m$, $E^m$, $C^m$, $D^m$, and $\mu$, respectively, as follows:

$$H^m = \Theta_{\frac{\lambda_2}{\mu}} \left( Z^m + \frac{D^m}{\mu} \right) \tag{33}$$

$$E^m = \max \left( 0, \frac{\mu}{2\lambda_3 + \mu} \left( X^m - X^m Z^m + \frac{C^m}{\mu} \right) \right) \tag{34}$$

and

$$\begin{cases} C^m = C^m + \mu \left( X^m - X^m Z^m - E^m \right) \\ D^m = D^m + \mu \left( Z^m - H^m \right) \\ \mu = \min \left( \eta\mu, \mu_{\max} \right). \end{cases} \tag{35}$$

Then, we can obtain combination matrices of $Z$, $H$, $E$, $C$, and $D$ of diverse modalities via *blkdiag* or concatenation.

*Step 2 (Update $W$ and $P$):* Referring to (17), we can optimize the following objective function to update $W$ via the GPI algorithm:

$$\begin{aligned} &\min \operatorname{tr} \left( W^T K W - 2 W^T J \right) \\ &\text{s.t.}\quad W^T W = I \end{aligned} \tag{36}$$

where $K = XZZ^TX^T - \lambda_1 XZZ^TX^T + 2\lambda_1 XZ\Xi Z^TX^T - \lambda_1 XZ\Xi\Xi Z^TX^T$ and $J = XZ\hat{B}^TP^{-1}$. Then, referring to (19), we can obtain the solution of $P$ as follows:

$$P = \left( W^T XZZ^TX^TW + \lambda_4 I \right)^{-1} W^T XZ\hat{B}^T. \tag{37}$$

*Step 3 (Update $\hat{B}$):* $\hat{B}$ is composed of $B$, while $B$ can be directly obtained by encoding the learned features of different modalities of the same palm as follows.

$$B = \frac{1}{2} \left( \operatorname{sgn} \left( \frac{1}{M} \sum_{m=1}^{M} P^T W^T X^m Z^m \right) + 1 \right). \tag{38}$$

Algorithm 2 summarizes the primary optimization steps of the E-MLPL. Having obtained the common feature projection functions of $W$ and $P$, we can easily obtain the compact binary feature codes of heterogeneous palmprint images and further convert them into histogram-based feature vectors for heterogeneous palmprint recognition.

**Algorithm 2** E-MLPL

**Input:** The training samples: $X \in R^{d \times MN}$;
Parameters: $\lambda_1$, $\lambda_2$, $\lambda_3$, and $\lambda_4$; Maximum number of iterations: $T$.
**Initialization:** $Z^m$, $H^m$, $E^m$, $W$, $P$, $B$, $C^m$, and $D^m$ with random values; $\mu = 0.1$;
$\mu_{\max} = 1e + 5$; and $\eta = 1.01$.
**for** *iteration* $< T$ **do**
Update $Z^m$ by using (32);
Update $H^m$ by using (33);
Update $E^m$ by using (34);
Update $W$ by using the GPI algorithm to optimize (36);
Update $P$ by using (37);
Update $\hat{B}$ by using (38);
Update $C^m$, $D^m$, and $\mu$ by using (35);
**Output:** $W$ and $P$.

## V. Experiments

In this section, we evaluate the performance of our method through comparisons conducted on five widely used palmprint databases, namely, CASIA [42], GPDS [43], PolyU_MS [33], XJTU-UP [29], and MPD [44], to evaluate our proposed MLPL and E-MLPL. All the experiments were implemented using MATLAB R2021b on a Windows 11.0 platform, with hardware comprising an Intel i7-13700KF CPU running at 3.40 GHz and 16 GB of RAM.

### A. Databases

The CASIA database comprises 5502 palmprint images obtained from 624 hands of 312 individuals, with each hand providing 8–17 palmprint images without constraints on hand posture and position during image capture.

The GPDS database contains 1000 palmprint images captured from the right palms of 100 individuals, with each palm contributing ten images acquired using a contactless camera setup without any hand constraints.

The PolyU_MS database comprises 24 000 palmprint images obtained from 500 palms of 250 volunteers. Each palm was imaged 12 times under four spectral illuminations, including green, red, blue, and near-infrared (NIR) bands, in a contact-based imaging manner.

The XJTU-UP database comprises a total of 8000 palmprint images acquired from 100 subjects, captured using different types of mobile phones, such as the Huawei Mate8 and iPhone 6S, under indoor natural light or flashlight conditions, respectively. Each palm was imaged 20 times under each collection condition (using one of two mobile phones and under one of the natural light and flashlight).

The MPD database contains 16 000 palmprint images provided by 400 palms of 200 volunteers. For each palm, 20 palmprint images were captured by Huawei mobile phones, and another 20 images by Xiaomi phones. Thus, 40 samples were collected for each palm.

All the palmprint images have been cropped into ROIs with a resolution of 64 × 64 pixels, and some palmprint

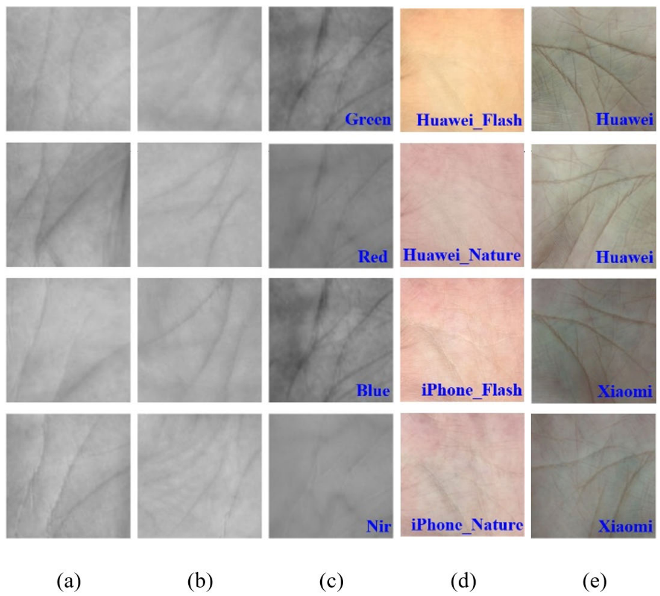



Fig. 2. Several exemplary palmprint ROI images randomly chosen from (a) CASIA, (b) GPDS, (c) PolyU_MS, (d) XJTU-UP, and (e) MPD databases.

images from the five databases are shown in Fig. 2. In this experiment, we use the CASIA, GPDS, and the NIR band of the PolyU_MS (PolyU_NIR) databases for homogeneous palmprint recognition evaluation and the XJTU-UP, MPD, and PolyU_MS databases for heterogeneous palmprint recognition testing.

### *B. Homogeneous Palmprint Recognition Results*

To assess the proposed MLPL, we perform palmprint identification and verification experiments under homogeneous conditions using the CASIA, GPDS, and PolyU_NIR databases. The evaluation involved comparisons with several state-of-the-art methods, including hand-crafted palmprint descriptors such as CompCode [8], LMDP [10], and LDDBP [45]; learning-based palmprint representation methods such as PCANet [46], DDBPD [32], LCMFC [47], RsBFL [48], SLCN [14], and LSIR [40]; and deep learning models such as VGG-F [49] and MobileNetV4 [50] pretrained on the ImageNet ILSVRC dataset. In palmprint identification, a single sample from each palm in every database is randomly chosen to create the training set, while the remaining samples serve as query data. All comparisons are conducted ten times, and the average rank-one identification accuracy is presented in Table I. For palmprint verification, we compare every two palmprint images in each database and calculate the genuine acceptance rate (GAR) and false acceptance rate (FAR). The ROC curves for the various methods evaluated are shown in Fig. 3.

The results indicate that our proposed method reliably attains greater identification accuracy rates and also higher GAR at the same FAR against the state-of-the-art methods. Unlike hand-crafted palmprint descriptors such as CompCode, LMDP, and LDDBP, which employ static encoding rules to convert palmprint images into features, our proposed MLPL first fully exploits the directional characteristics embedded in palmprint images and then adaptively learns discriminative features from the resulting directional representation, leading to better recognition performance. In addition, the performance of the proposed MLPL is competitive with that of existing learning-based palmprint descriptors such as PCANet, DDBPD, LCMFC, SLCN, and LSIR. This is possibly because the existing methods usually learn single-stage feature projection functions, while our proposed method learns not only feature projection but also feature rotation mappings, minimizing information loss during feature encoding. Finally, unlike popular deep learning networks such as MobileNetV4, which usually require large numbers of training samples to improve recognition performance, our proposed method incorporates an elaborate design aimed at learning compact binary codes for effective palmprint discrimination, achieving a higher recognition rate with limited training samples. The superiority of the proposed approach is particularly evident on the PolyU_NIR dataset, where stable NIR lighting and contact-based image acquisition yield consistent spectral characteristics and reduced intraclass variations, thereby promoting more discriminative feature learning.

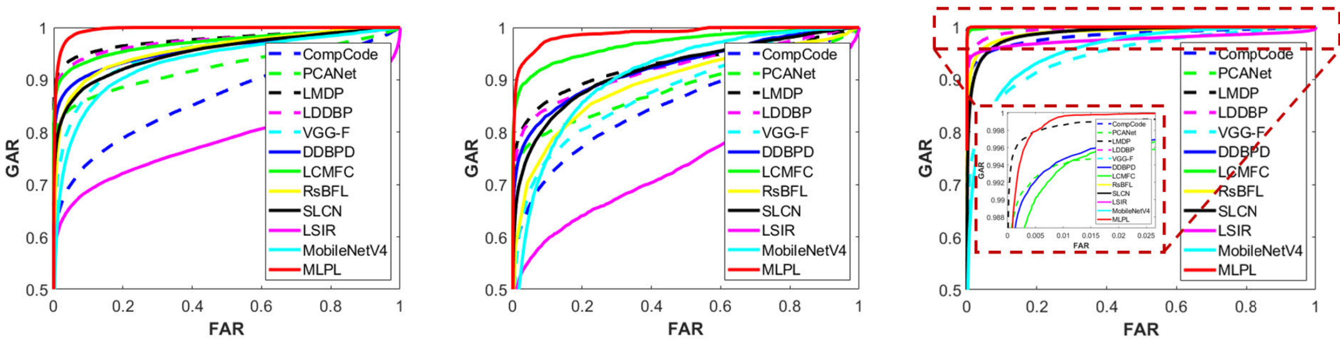
### *C. Heterogeneous Palmprint Recognition Results*

Moreover, we evaluate the E-MLPL by conducting heterogeneous palmprint identification and verification experiments on the PolyU_MS, XJTU-UP, and MPD databases, all of which include images collected under varying conditions. Specifically, we combine the NIR band with the green, blue, and red bands of PolyU_MS samples to form three PolyU heterogeneous palmprint image datasets, referred to as G_N (green+NIR), B_N (blue+NIR), and R_N (red+NIR), respectively. In addition, we combine the XJTU-UP palmprint images captured by using Huawei Mate8 under natural light and an iPhone 6S under flashlight to form the XJTU-UP heterogeneous image dataset, named HN_IF. Similarly, another heterogeneous image dataset, named HF_IN, is constructed by combining images captured with Huawei Mate8 under flashlight and iPhone 6S under natural light. In addition, we select the MPD palmprint images captured using the Huawei (HW) and Xiaomi (XM) phones to form the MPD heterogeneous palmprint image dataset, denoted as HW_XM.

In heterogeneous palmprint identification, for each palm in every "A_B" heterogeneous dataset, we randomly select half of the samples from "A" and half from "B" to form the training sample set and use the rest for identification testing. For a better evaluation, we compare the proposed method with state-of-the-art approaches, including CompCode [8], OLOF [42], ResNet [51], PCANet [46], DDBPD [32], LCMFC [47], RsBFL [48], SLCN [14], EfficientVit [52], and MobileNetV4 [50], and cross-domain recognition methods such as DHPFL [53]. ResNet, EfficientVit, and MobileNetV4 were pretrained on the ImageNet ILSVRC dataset to extract deep palmprint features. Table II tabulates the average rank-one identification accuracies of different methods on diverse heterogeneous datasets, where all methods are repeated ten times to report the average results. "A→B" indicates that all samples of the testing subset from A are used as probes, while all samples from B form the gallery. In heterogeneous palmprint verification, for each "A_B" dataset, every palmprint

TABLE I
AVERAGE RANK-ONE HOMOGENEOUS PALMPRINT IDENTIFICATION ACCURACY OBTAINED BY DIFFERENT METHODS ON THE CASIA, GPDS, AND POLYU_NIR DATABASES

| Databases | CompCode | PCANet | LMDP | LDDBP | VGG-F | DDBPD | LCMFC | RsBFL | SLCN | LSIR | MobileNetV4 | MLPL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CASIA | 63.44±5.48 | 81.08±2.41 | 90.46±1.76 | 76.25±3.75 | 67.38±3.96 | 84.65±1.57 | 88.30±2.57 | 84.53±2.79 | 75.03±3.29 | 56.53±2.61 | 59.96±3.28 | **91.04±0.95** |
| GPDS | 62.50±3.27 | 79.98±1.97 | 77.93±4.95 | 69.11±1.60 | 66.18±2.09 | 76.58±3.58 | 89.89±1.06 | 75.83±1.87 | 70.60±3.15 | 53.19±2.39 | 60.57±3.57 | **93.47±0.83** |
| PolyU_NIR | 90.08±0.80 | 99.40±0.08 | 99.67±0.09 | 94.26±0.55 | 71.24±0.85 | 99.09±0.18 | 99.30±0.12 | 92.84±0.51 | 85.68±1.22 | 60.06±2.82 | 64.09±2.50 | **99.78±0.12** |




(a) (b) (c)

Fig. 3. ROC curves of the GAR versus the FAR of homogeneous palmprint verification drawn by different methods on (a) CASIA, (b) GPDS, and (c) PolyU_NIR databases.

TABLE II
AVERAGE RANK-ONE HETEROGENEOUS PALMPRINT IDENTIFICATION RATES ACHIEVED BY DIFFERENT METHODS ON THE HN_IF, HF_IN, HW_XM, G_N, B_N, AND R_N HETEROGENEOUS PALMPRINT IMAGE DATASETS

| Databases | CompCode | OLOF | ResNet | PCANet | DDBPD | LCMFC | RsBFL | SLCN | DHPFL | EfficientVit | MobileNetV4 | E-MLPL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IF→HN | 83.82±0.30 | 83.70±0.28 | 80.28±0.12 | 84.80±0.22 | 83.95±0.12 | 84.47±0.11 | 80.17±0.10 | 81.46±0.29 | 84.18±0.14 | 80.38±0.21 | 80.52±0.14 | **94.96±0.28** |
| HN→IF | 83.38±0.27 | 82.85±0.20 | 80.35±0.16 | 84.73±0.10 | 83.91±0.19 | 84.58±0.11 | 80.51±0.07 | 81.58±0.13 | 83.76±0.20 | 83.76±0.20 | 80.69±0.15 | **94.58±0.33** |
| IN→HF | 89.32±0.30 | 89.37±0.54 | 80.96±0.25 | 92.12±0.25 | 91.34±0.24 | 91.95±0.40 | 84.02±0.32 | 81.43±0.21 | 88.95±0.33 | 80.22±0.17 | 80.54±0.14 | **94.68±0.44** |
| HF→IN | 89.51±0.22 | 89.02±0.20 | 80.71±0.38 | 92.29±0.09 | 90.86±0.35 | 92.06±0.55 | 83.22±0.27 | 80.52±0.19 | 88.96±0.33 | 80.24±0.12 | 80.43±0.14 | **94.16±0.43** |
| HW→XM | 92.23±0.21 | 92.73±0.23 | 86.34±0.22 | 96.62±0.65 | 95.07±0.24 | 95.41±0.23 | 84.94±0.27 | 84.44±0.24 | 93.03±0.27 | 83.16±0.30 | 83.36±0.17 | **98.04±0.18** |
| XM→HW | 92.43±0.22 | 92.46±0.20 | 87.09±0.30 | 96.90±0.33 | 94.87±0.22 | 95.22±1.14 | 85.10±0.13 | 82.78±0.24 | 93.26±0.14 | 84.00±0.30 | 84.16±0.15 | **98.54±0.28** |
| N→G | 89.64±0.22 | 91.90±0.21 | 83.35±0.09 | 95.44±0.41 | 86.23±0.22 | 86.75±0.34 | 83.53±0.09 | 83.53±0.05 | 85.88±0.14 | 83.37±0.04 | 83.38±0.04 | **97.46±0.13** |
| G→N | 88.59±0.16 | 92.00±0.11 | 83.33±0.14 | 95.10±0.30 | 86.05±0.31 | 86.99±0.51 | 83.17±0.11 | 83.77±0.12 | 86.63±0.21 | 83.46±0.06 | 83.36±0.03 | **96.27±0.36** |
| N→B | 86.97±0.19 | 90.46±0.27 | 83.36±0.09 | 93.20±0.40 | 84.94±0.16 | 86.25±0.25 | 83.40±0.05 | 83.62±0.05 | 85.44±0.14 | 83.40±0.08 | 83.37±0.04 | **95.66±0.15** |
| B→N | 85.12±0.14 | 90.33±0.19 | 83.35±0.18 | 92.54±0.46 | 84.80±0.32 | 85.88±0.39 | 83.86±0.15 | 83.89±0.22 | 86.65±0.19 | 83.41±0.04 | 83.37±0.03 | **95.80±0.07** |
| N→R | 99.66±0.07 | 99.76±0.10 | 84.08±0.09 | 99.95±0.04 | 97.42±0.13 | 99.44±0.19 | 97.67±0.21 | 84.13±0.11 | 98.65±0.25 | 84.50±0.12 | 83.76±0.08 | **99.96±0.03** |
| R→N | 99.33±0.26 | 99.65±0.22 | 84.13±0.15 | 99.76±0.18 | 97.02±0.11 | 99.44±0.06 | 97.88±0.23 | 87.94±0.31 | 99.44±0.24 | 84.52±0.12 | 83.80±0.08 | **99.80±0.08** |

image from "A" is compared with each image from "B" to calculate the GAR and FAR, and Fig. 4 depicts the ROC curves drawn by different methods.

Similar recognition performance to that observed in homogeneous palmprint experiments indicates that the E-MLPL performs more effectively than state-of-the-art methods, with enhanced identification and verification accuracy relative to existing approaches. This phenomenon may be explained by the fact that the palmprint images with heterogeneous modalities usually have a large feature gap, making most existing hand-crafted methods hard to obtain the modality-invariant palmprint feature descriptors with fixed feature extraction engineering. Moreover, deep learning networks such as EfficientVit and MobileNetV4 typically require large-scale labeled data to generalize effectively, which constrains their performance on small-scale palmprint datasets. In addition, most existing learning methods learn discriminative representations separately for diverse heterogeneous palmprint modalities, which makes it difficult to bridge the modality feature differences inherent in heterogeneous images. Unlike them, our proposed method first isolates and learns discriminative features specific to heterogeneous palmprint images and further jointly obtains the common and collaborative binary features. As a result, the modality gap at the final feature level can be effectively reduced, and enhanced recognition accuracy for heterogeneous data can be obtained. Such improvement is more evident on the datasets (such as MPD and PolyU_MS) exhibiting smaller cross-domain discrepancies, whereas the considerable variations in lighting conditions and imaging devices in XJTU-UP introduce more significant domain shifts, leading to a marginal reduction in recognition accuracy.

### D. Parameter Analysis

The objective functions of our proposed MLPL and E-MLPL methods, i.e., (7) and (27), mainly contain four balance parameters, including $\lambda_1$, $\lambda_2$, $\lambda_3$, and $\lambda_4$. To the best of our knowledge, directly determining the optimal values of these balancing parameters is challenging, as these values generally depend on the diverse characteristics inherent in palmprint images from various sources. In this analysis,

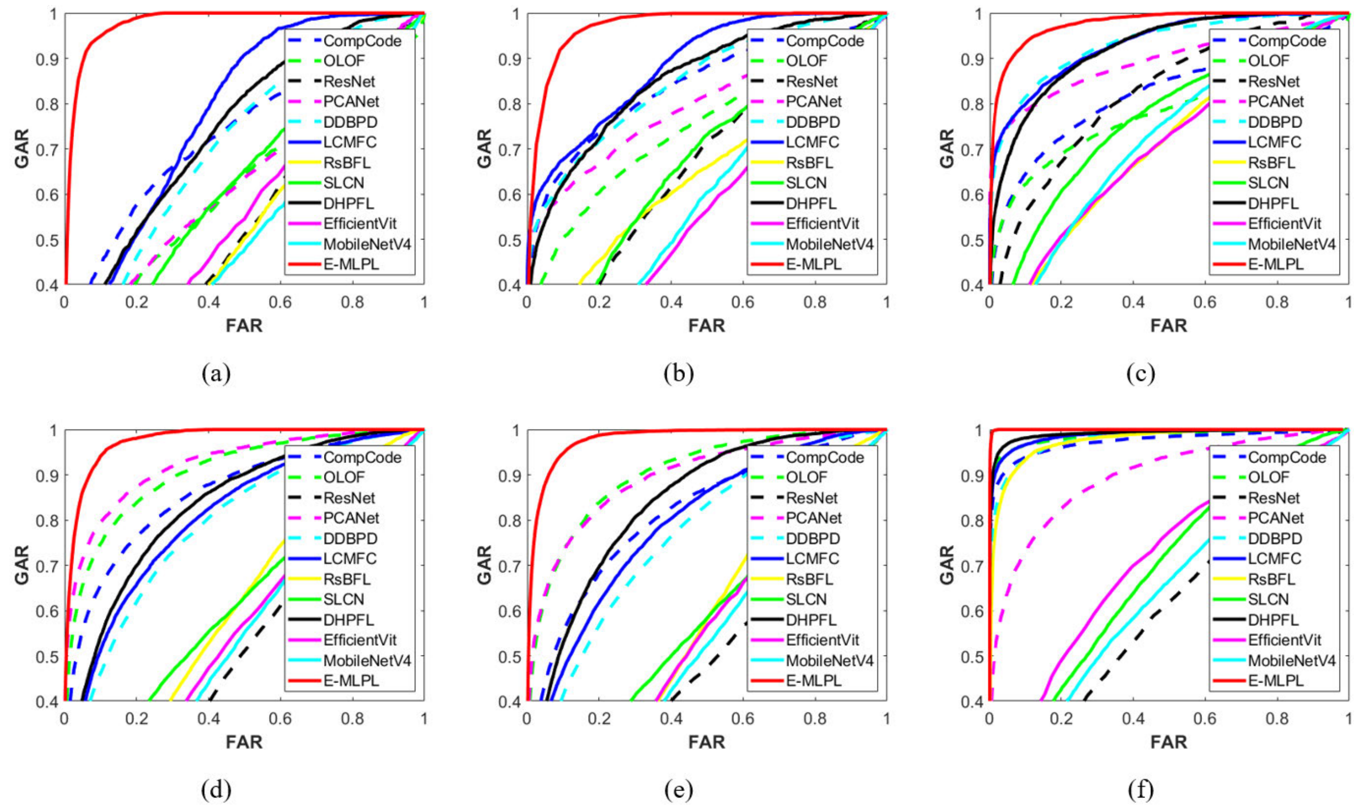



Fig. 4. ROC curves (FAR versus GAR) of heterogeneous palmprint verification drawn by different methods on (a) HN_IF, (b) HF_IN, (c) HW_XM, (d) G_N, (e) B_N, and (f) R_N datasets.

we first empirically define a candidate value set for these parameters, i.e., {0.001, 0.01, 0.1, 1, 10, 100, 1000}, and then conduct comparative verification experiments to test the proposed method. Notably, there is a large number of possible combinations (i.e., $7^4$) for the four parameters within the candidate set. For this reason, we search for the suboptimal value of one parameter while empirically fixing the possible values of the others.

Specifically, we empirically set $\lambda_2$, $\lambda_3$, and $\lambda_4$ to 0.01, and conduct palmprint verification with the MLPL on the CASIA dataset and the E-MLPL on the HN_IF dataset to compute the equal error rate (EER) versus different values of $\lambda_1$, as shown in Fig. 5(a) and (d). The results indicate that our method consistently achieves superior performance when $\lambda_1 \in \{0.1, 1, 10, 100, 1000\}$. Then, we empirically set $\lambda_1$ to 1, and calculate the EERs of the MLPL and E-MLPL methods for different values of $\lambda_2$ and $\lambda_3$, as shown in Fig. 5(b) and (e), respectively. It is observed that the MLPL method achieves the lowest EER when $\lambda_2$ is set to 1000 and $\lambda_3 \in \{0.001, 0.01, 100, 1000\}$, while the E-MLPL method performs optimally when $\lambda_2, \lambda_3 \in \{100, 1000\}$. Finally, we empirically set both $\lambda_2$ and $\lambda_3$ to 1000 and test the EER changes of the MLPL and the E-MLPL versus different values of $\lambda_4$, as reported in Fig. 5(c) and (f), respectively. The proposed method usually achieves the best recognition performance when $\lambda_4$ is set to a value from the set $\{1, 10, 100, 1000\}$. As a result, we empirically set $\lambda_1$, $\lambda_2$, $\lambda_3$, and $\lambda_4$ to 1, 1000, 1000, and 10, respectively, for both the MLPL and E-MLPL methods.

### E. Convergence Analysis

To analyze the convergence of our proposed MLPL and E-MLPL, we calculate the values of the MLPL and E-MLPL objective functions, i.e., (7) and (27), versus varying numbers of iterations, respectively, as shown in Fig. 6. It can be seen that both the MLPL and E-MLPL converge in approximately six iterations on all datasets, demonstrating the rapid convergence speed of both methods.

### F. Computational Efficiency and Memory Usage Analysis

The computational costs of the proposed MLPL and E-MLPL frameworks primarily stem from matrix inversion and singular value thresholding, while memory usage is determined by the size of the involved variables. Specifically, for Algorithm 1, the most time-consuming operations are the matrix inversion and the SVT. Updating $Z$ requires inversion of an $N \times N$ matrix with complexity $O(N^3)$, while updating $H$ involves SVT on an $N \times N$ matrix with complexity $O(\tau_1 N^3)$, where $\tau_1$ is the number of SVT iterations. Updating $P$ and $W$ involves $d \times d$ matrix inversion and SVT, with complexities $O(d^3)$ and $O(\tau_2 d^3)$, respectively, where $\tau_2$ denotes the iteration number for $W$. Other updates mainly involve basic matrix operations with negligible cost. Consequently, the overall complexity of Algorithm 1 can be approximated as $O((\tau_1+1)N^3 + (\tau_2+1)d^3)$. For the multimodal extension, the E-MLPL method, as presented in Algorithm 2, performs a joint optimization of the modality-specific variables $Z^m$ and $H^m$ together with the shared projection mappings $W$ and $P$

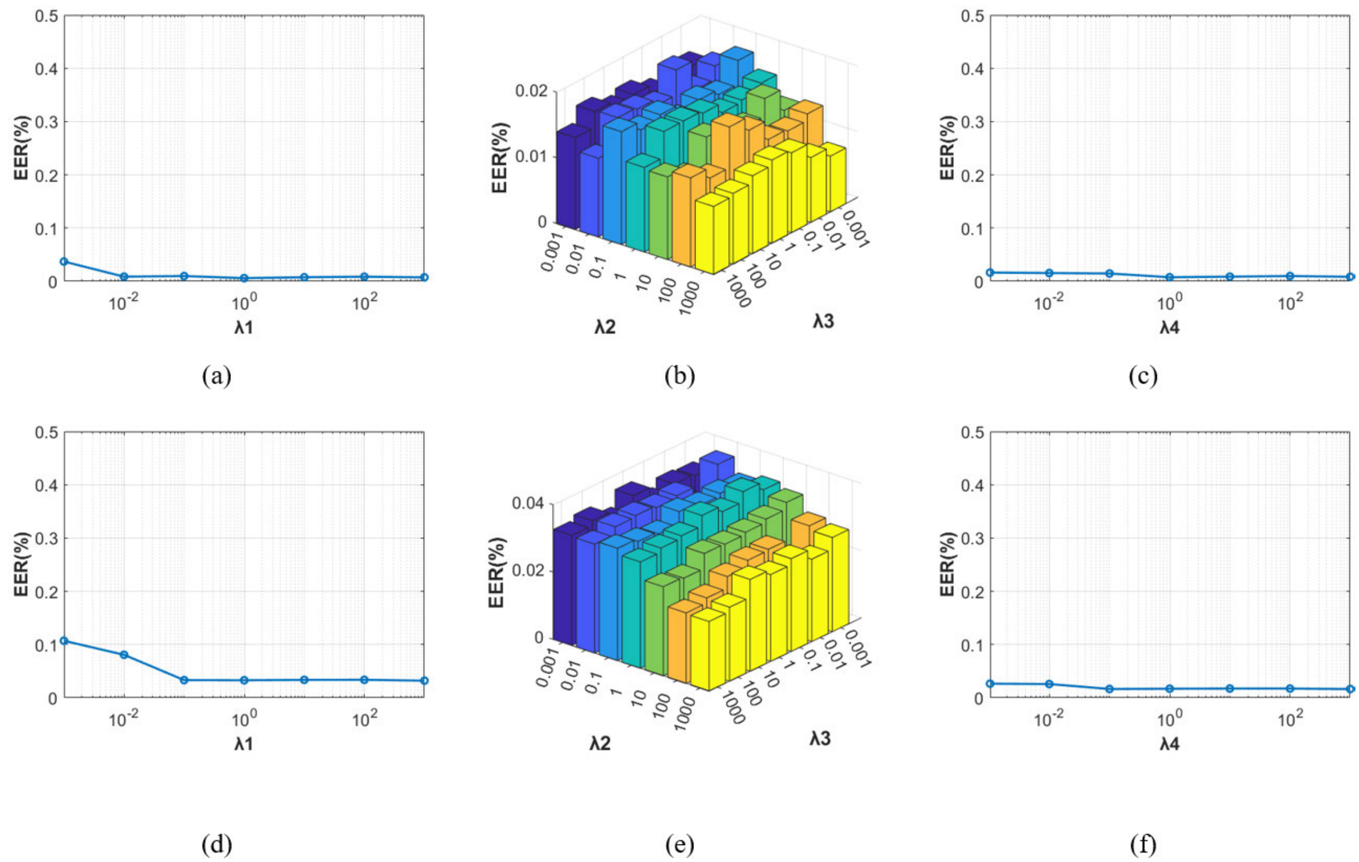



Fig. 5. EER variations of the MLPL and the E-MLPL methods versus the values of $\lambda_1$, $\lambda_2$, $\lambda_3$, and $\lambda_4$ on the (a)–(c) CASIA and (d)–(f) HN_IF databases, respectively.

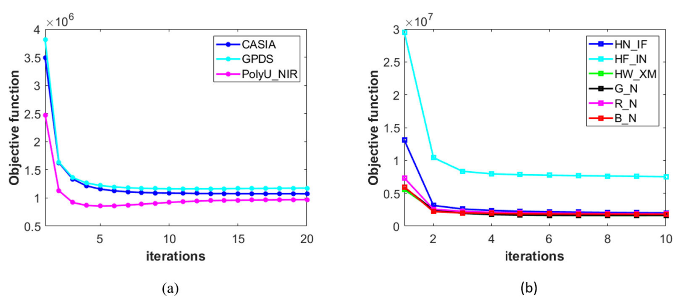



Fig. 6. Objective function value of (a) MLPL and (b) E-MLPL on different databases.

TABLE III

Average Time (ms) Taken by Different Methods to Perform Feature Extraction and Matching for Processing a Palmprint Image

| Methods | Feature extraction | Matching | Total |
|---|---|---|---|
| PCANet | 173.2012 | 69.312 | 242.513 |
| LMDP | 35.967 | 6.915 | 42.882 |
| LDDBP | 83.347 | 14.344 | 97.691 |
| DDBPD | 27.883 | 10.777 | 38.660 |
| SLCN | 44.539 | 12.862 | 57.401 |
| LSIR | 33.864 | 83.840 | 117.704 |
| **MLPL** | **27.707** | **6.525** | **34.232** |

across $m = 1, 2, \ldots, M$ modalities. This results in a total complexity of $O(\sum_{m=1}^{M} (\tau_1 + 1)N_m^3 + (\tau_2 + 1)d^3)$.

To further verify the theoretical complexity analysis, we compared the actual computational efficiency of the proposed method with the state-of-the-art methods. For a fair comparison, we mainly implemented the traditional CPU-based palmprint descriptors, such as PCANet, LMDP, and DDBPD, for efficiency comparison, while the typical deep learning networks like VGG-F and MobileNetV4 were not compared, because they heavily rely on GPU accelerators. In addition, since the MLPL and the E-MLPL methods share similar feature learning, extraction, and matching procedures, we focus primarily on calculating the time cost of homogeneous palmprint identification with the proposed MLPL. Table III presents the average time (ms) for feature extraction and matching of palmprint images across different methods. It shows that our proposed method usually requires less time than the state-of-the-art approaches, demonstrating its outstanding efficiency.

Moreover, the memory usage of the MLPL is primarily determined by the matrices stored during optimization, including the feature dimension $d$, the number of DRD vectors per modality $N$, and the total number of heterogeneous modalities $M$. Consequently, the proposed method is most suitable for medium-scale datasets, while mini-batch strategies can be adopted to reduce memory requirements for very large-scale datasets. Overall, the proposed algorithms can achieve a favorable balance between computational efficiency and memory consumption, making them applicable for real-world applications.

## VI. Conclusion

In this article, we propose a three-layer feature projection learning method for palmprint recognition, which first learns an LRR mapping to decompose noise information from raw palmprint images, then learns a discriminative feature projection to capture the personal-invariant features from corrected palmprint information, and finally learns a feature rotation mapping for discriminative feature encoding while minimizing the quantization loss. Moreover, we extend the MLPL by reducing the domain gaps of heterogeneous palmprint images, making our proposed method applicable for both homogeneous and heterogeneous palmprint recognition. Compared with state-of-the-art approaches, the proposed method achieves superior recognition performance, as evidenced by experiments on various types of palmprint image databases. Its lightweight architecture, high robustness, and accurate feature representation make the framework well-suited for practical applications, such as mobile payment and access control systems, while the MLPL and E-MLPL may exhibit performance degradation under extreme cross-spectral or cross-device conditions. In future work, we will focus on addressing such limitations to improve model stability under diverse acquisition scenarios. How to extend our proposed method to other palmprint-related biometric modalities, such as finger vein and hand geometry recognition, constitutes another promising direction.

## REFERENCES


[1] A. K. Jain, D. Deb, and J. J. Engelsma, "Biometrics: Trust, but verify," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 4, no. 3, pp. 303–323, Jul. 2022.

[2] S. Zhao, W. Nie, and B. Zhang, "Multi-feature fusion using collaborative residual for hyperspectral palmprint recognition," in *Proc. IEEE 4th Int. Conf. Comput. Commun. (ICCC)*, Dec. 2018, pp. 1402–1406.

[3] L. Fei, G. Lu, W. Jia, J. Wen, and D. Zhang, "Complete binary representation for 3-D palmprint recognition," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 12, pp. 2761–2771, Dec. 2018.

[4] S. A. Grosz, A. Godbole, and A. K. Jain, "Mobile contactless palmprint recognition: Use of multiscale, multimodel embeddings," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 8428–8440, 2024.

[5] L. Su, L. Fei, B. Zhang, S. Zhao, J. Wen, and Y. Xu, "Complete region of interest for unconstrained palmprint recognition," *IEEE Trans. Image Process.*, vol. 33, pp. 3662–3675, 2024.

[6] D.-S. Huang, W. Jia, and D. Zhang, "Palmprint verification based on principal lines," *Pattern Recognit.*, vol. 41, no. 4, pp. 1316–1328, Apr. 2008.

[7] X. Wang, L. Lei, and M. Wang, "Palmprint verification based on 2D – Gabor wavelet and pulse-coupled neural network," *Knowl.-Based Syst.*, vol. 27, pp. 451–455, Mar. 2012.

[8] D. Zhang, Z. Guo, G. Lu, L. Zhang, and W. Zuo, "An online system of multispectral palmprint verification," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 2, pp. 480–490, Feb. 2010.

[9] Y. Xu, L. Fei, J. Wen, and D. Zhang, "Discriminative and robust competitive code for palmprint recognition," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 2, pp. 232–241, Feb. 2018.

[10] L. Fei, J. Wen, Z. Zhang, K. Yan, and Z. Zhong, "Local multiple directional pattern of palmprint image," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 3013–3018.

[11] L. Fei, B. Zhang, W. Zhang, and S. Teng, "Local apparent and latent direction extraction for palmprint recognition," *Inf. Sci.*, vol. 473, pp. 59–72, Jan. 2019.

[12] M. D. Bounneche, L. Boubchir, A. Bouridane, B. Nekhoul, and A. Ali-Chérif, "Multi-spectral palmprint recognition based on oriented multiscale log-Gabor filters," *Neurocomputing*, vol. 205, pp. 274–286, Sep. 2016.

[13] S. Zhao and B. Zhang, "Learning salient and discriminative descriptor for palmprint feature extraction and identification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 12, pp. 5219–5230, Dec. 2020.

[14] L. Fei, S. Zhao, W. Jia, B. Zhang, J. Wen, and Y. Xu, "Toward efficient palmprint feature extraction by learning a single-layer convolution network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 12, pp. 9783–9794, Dec. 2023.

[15] S. Li, J. Zhou, B. Zhang, L. Wu, and M. Jian, "Discriminative multiview learning for robust palmprint feature representation and recognition," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 6, no. 3, pp. 304–313, Jul. 2024.

[16] S. Zhao, L. Fei, B. Zhang, J. Wen, and J. Cui, "Toward mobile palmprint recognition via multi-view hierarchical graph learning," *IEEE Trans. Inf. Forensics Security*, vol. 20, pp. 101–113, 2025.

[17] D. Fan, X. Liang, C. Zhang, W. Jia, and D. Zhang, "AMGNet: Aligned multilevel Gabor convolution network for palmprint recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 6, pp. 4175–4189, Jun. 2024.

[18] L. Fei, L. Su, B. Zhang, S. Zhao, J. Wen, and X. Li, "Learning frequency-aware common feature for VIS-NIR heterogeneous palmprint recognition," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 7604–7618, 2024.

[19] Y. Zhong, W. Chai, L. Wang, D. Zheng, H. Fu, and H. Ma, "RegPalm: Toward large-scale open-set palmprint recognition by reducing pattern variance," *IEEE Trans. Inf. Forensics Security*, vol. 20, pp. 8541–8554, 2025.

[20] T. Chai, X. Wang, R. Li, W. Jia, and X. Wu, "Joint finger valley points-free ROI detection and recurrent layer aggregation for palmprint recognition in open environment," *IEEE Trans. Inf. Forensics Security*, vol. 20, pp. 421–435, 2025.

[21] C. Gao, Z. Yang, W. Jia, L. Leng, B. Zhang, and A. B. J. Teoh, "Deep learning in palmprint recognition—A comprehensive survey," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2025, pp. 1–19.

[22] H. Shao, S. Shi, X. Du, D. Zeng, and D. Zhong, "Robust palmprint recognition via multi-stage noisy label selection and correction," *IEEE Trans. Image Process.*, vol. 34, pp. 4591–4601, 2025.

[23] *PolyU Palmprint Database*. Accessed: Jan. 1, 2004. [Online]. Available: http://www.comp.polyu.edu.hk/~biometrics/

[24] A. Kumar and S. Shekhar, "Personal identification using multibiometrics rank-level fusion," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 41, no. 5, pp. 743–752, Sep. 2011.

[25] L. Zhang, L. Li, A. Yang, Y. Shen, and M. Yang, "Towards contactless palmprint recognition: A novel device, a new benchmark, and a collaborative representation based identification approach," *Pattern Recognit.*, vol. 69, pp. 199–212, Sep. 2017.

[26] H. Shao, D. Zhong, and Y. Li, "PalmGAN for cross-domain palmprint recognition," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2019, pp. 1390–1395.

[27] S. Li, J. Hu, B. Zhang, X. Ning, and L. Wu, "Dynamic personalized federated learning for cross-spectral palmprint recognition," *IEEE Trans. Image Process.*, vol. 34, pp. 4885–4895, 2025.

[28] H. Shao and D. Zhong, "Learning with partners to improve the multi-source cross-dataset palmprint recognition," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 5182–5194, 2021.

[29] H. Shao, D. Zhong, and X. Du, "Deep distillation hashing for unconstrained palmprint recognition," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 2505613.

[30] H. Shao, Y. Zou, C. Liu, Q. Guo, and D. Zhong, "Learning to generalize unseen dataset for cross-dataset palmprint recognition," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 3788–3799, 2024.

[31] L. Fei, G. Lu, W. Jia, S. Teng, and D. Zhang, "Feature extraction methods for palmprint recognition: A survey and evaluation," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 2, pp. 346–363, Feb. 2019.

[32] L. Fei, B. Zhang, Y. Xu, Z. Guo, J. Wen, and W. Jia, "Learning discriminant direction binary palmprint descriptor," *IEEE Trans. Image Process.*, vol. 28, no. 8, pp. 3808–3820, Aug. 2019.

[33] G. Liu and S. Yan, "Latent low-rank representation for subspace segmentation and feature extraction," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 1615–1622.

[34] Y. Fu, J. Gao, D. Tien, Z. Lin, and X. Hong, "Tensor LRR and sparse coding-based subspace clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 10, pp. 2120–2133, Oct. 2016.

[35] W. Yang, Y. Shi, Y. Gao, L. Wang, and M. Yang, "Incomplete-data oriented multiview dimension reduction via sparse low-rank representation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 12, pp. 6276–6291, Dec. 2018.

[36] N. He, J.-B. Wang, L.-L. Zhang, and K. Lu, “Convex optimization based low-rank matrix decomposition for image restoration,” *Neurocomputing*, vol. 172, pp. 253–261, Jan. 2016.

[37] S. Li, R. Ma, J. Zhou, B. Zhang, and L. Wu, “Joint discriminative analysis with low-rank projection for finger vein feature extraction,” *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 959–969, 2024.

[38] K. Jing et al., “Prototype correction via contrastive augmentation for few-shot unconstrained palmprint recognition,” *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 5431–5446, 2023.

[39] S. Zhao and B. Zhang, “Learning complete and discriminative direction pattern for robust palmprint recognition,” *IEEE Trans. Image Process.*, vol. 30, pp. 1001–1014, 2021.

[40] L. Fei, W. K. Wong, S. Zhao, J. Wen, J. Zhu, and Y. Xu, “Learning spectrum-invariance representation for cross-spectral palmprint recognition,” *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 6, pp. 3868–3879, Jun. 2023.

[41] F. Nie, R. Zhang, and X. Li, “A generalized power iteration method for solving quadratic problem on the Stiefel manifold,” *Sci. China Inf. Sci.*, vol. 60, no. 11, May 2017, Art. no. 112101.

[42] Z. Sun, T. Tan, Y. Wang, and S. Z. Li, “Ordinal palmprint represention for personal identification,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR05)*, vol. 1, Jun. 2005, pp. 279–284.

[43] (May 2018). *GPDS Palmprint Database*. [Online]. Available: http://www.gpds.ulpgc.es

[44] Y. Zhang, L. Zhang, R. Zhang, S. Li, J. Li, and F. Huang, “Towards palmprint verification on smartphones,” 2020, *arXiv:2003.13266*.

[45] L. Fei, B. Zhang, Y. Xu, D. Huang, W. Jia, and J. Wen, “Local discriminant direction binary pattern for palmprint representation and recognition,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 2, pp. 468–481, Feb. 2020.

[46] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, “PCANet: A simple deep learning baseline for image classification?” *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5017–5032, Dec. 2015.

[47] L. Fei, B. Zhang, L. Zhang, W. Jia, J. Wen, and J. Wu, “Learning compact multifeature codes for palmprint recognition from a single training image per palm,” *IEEE Trans. Multimedia*, vol. 23, pp. 2930–2942, 2021.

[48] S. Li, R. Ma, J. Zhou, and B. Zhang, “Row-sparsity binary feature learning for open-set palmprint recognition,” in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2022, pp. 1–8.

[49] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, *arXiv:1409.1556*.

[50] D. Qin et al., “Mobilenetv4-universal models for the mobile ecosystem,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Sep. 2024, pp. 78–96.

[51] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[52] X. Liu, H. Peng, N. Zheng, Y. Yang, H. Hu, and Y. Yuan, “EfficientViT: Memory efficient vision transformer with cascaded group attention,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 14420–14430.

[53] L. Fei, B. Zhang, Y. Xu, C. Tian, I. Rida, and D. Zhang, “Jointly heterogeneous palmprint discriminant feature learning,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 9, pp. 4979–4990, Sep. 2022.

**Lunke Fei** (Senior Member, IEEE) received the Ph.D. degree in computer science and technology from Harbin Institute of Technology, Shenzhen, China, in 2016.

He was a Post-Doctoral Fellow with the Department of Computer and Information Science, University of Macau, Macau, China. He is currently a Professor with the School of Computer Science and Technology, Guangdong University of Technology, Guangzhou, China. His research interests include pattern recognition, biometrics, image processing, and machine learning.

**Kaiting Huang** is currently pursuing the M.S. degree in computer technology with the School of Computer Science and Technology, Guangdong University of Technology, Guangzhou, China.

Her current research interests include pattern recognition, biometrics, and machine learning.

**Shuping Zhao** (Member, IEEE) received the M.S. degree in information science from South China Normal University, Guangzhou, China, in 2016, and the Ph.D. degree in computer science from the Department of Computer and Information Science, University of Macau, Taipa, Macau, in 2020.

He is currently working with the School of Computer Science, Guangdong University of Technology, Guangzhou. He has published a number of technical papers in prestigious international journals, i.e., IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, and international conferences. His research interests include biometrics, machine learning, pattern recognition, and image processing.

**Qi Zhu** received the B.S., M.S., and Ph.D. degrees in computer science and technology from Harbin Institute of Technology, Harbin, China, in 2007, 2010, and 2014, respectively.

He is currently a Professor with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China. He has published over 100 scientific articles in refereed international journals and conference proceedings. His interests include pattern recognition, medical image analysis, and biometrics.

**Bob Zhang** (Senior Member, IEEE) received the B.A. degree in computer science from York University, Toronto, ON, Canada, in 2006, the M.A.Sc. degree in information systems security from Concordia University, Montreal, QC, Canada, in 2007, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2011.

He was with the Center for Pattern Recognition and Machine Intelligence, University of Waterloo. He was a Post-Doctoral Researcher with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. He is currently an Associate Professor with the Department of Computer and Information Science, University of Macau, Taipa, Macau. His research interests focus on biometrics, pattern recognition, and image processing.

Dr. Zhang is a Technical Committee Member of the IEEE Systems, Man, and Cybernetics Society and an Editorial Board Member of the *International Journal of Information*. He is an Associate Editor of the *International Journal of Image and Graphics*.

**Wei Jia** (Member, IEEE) received the B.Sc. degree in informatics from Central China Normal University, Wuhan, China, in 1998, the M.Sc. degree in computer science from Hefei University of Technology, Hefei, China, in 2004, and the Ph.D. degree in pattern recognition and intelligence system from the University of Science and Technology of China, Hefei, in 2008.

He was a Research Assistant and an Associate Professor at Hefei Institutes of Physical Science, Chinese Academy of Sciences, Beijing, China, from 2008 to 2016. He is currently a Research Associate Professor with the School of Computer and Information, Hefei University of Technology. His research interests include computer vision, biometrics, pattern recognition, image processing, and machine learning.