

Received April 2, 2020, accepted April 16, 2020, date of publication April 27, 2020, date of current version May 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2990736

Rotation and Translation Invariant Palmprint Recognition With Biologically Inspired Transform

XIANCHENG ZHOU, KAIJUN ZHOU^{ID}, AND LIZHI SHEN

Mobile E-Business Collaborative Innovation Center of Hunan Province, Hunan University of Technology and Business, Changsha 410205, China

Corresponding author: Kaijun Zhou (alpha218@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61471170 and Grant 61976088, in part by the Natural Science Foundation of Hunan Province, China, under Grant 2018JJ2197 and Grant 2018JJ3260, and in part by the Scientific Research Foundation of Hunan Provincial University, China, under Grant 14A078, Grant 17B144, Grant 18A305, and Grant 18B338.

ABSTRACT Extracting rotation and translation invariant features is a difficult task for palmprint recognition. Traditional methods have difficulty in dealing with palmprint images degraded by those variations. Studies have shown that neurons at higher levels exhibit an increasing degree of invariance to above mentioned image variations. Moreover, primary visual cortex(V1) is believed to give stronger responses to light bars of certain directions. Based on these observations, a biologically inspired transform feature extractor, namely BIT, for palmprint recognition is proposed in this paper. BIT involves two stages, which mimics visual information processing in V1. In the first stage, we build an orientation edge detector to highlight the edges response in each direction. The orientation edge detector is primarily composed of a phase congruency based edge detector and a bipolar filter. After that, a local spatial frequency detector produces a response, converting rotation factors of orientation edges into a horizontal shifted map. In the second stage, the orientation edge detector and local spatial frequency detector are applied again, which converts shifted map into an invariant pixel in feature map. Extensive experimental results not only show that our method is robust to image variations including rotation and translation, but also illustrate the effectiveness and discriminability of the extracted invariant palmprint features in recognition problems.

INDEX TERMS Biometrics, image filter, feature extraction, palmprint recognition, biologically inspired transform.

I. INTRODUCTION

Biometrics recognition for person authentication has been extensively studied in recent years [1]–[3]. Among commonly used human biological traits, palmprint involves information that is unique to an individual, so that it can be used for identification and access control with high security. Therefore, palmprint recognition is beginning to play an important role in law enforcement and forensic applications [1], and has drawn much attention in biometrics area.

Palmprint of human beings primarily involves lines, wrinkles and ridges. Essentially, palmprint identification is a biometric authentication process through comparing those unique patterns [4], [5]. It is not difficult for human vision systems to verify a palmprint as they are able to align palmprint images automatically and thus eliminate the effect of variations such as illumination, rotation and translation.

The associate editor coordinating the review of this manuscript and approving it for publication was Yonghong Peng^{ID}.

However, it remains a very challenging task for a computer to handle these interferences simultaneously when a palmprint image is captured by a mobile terminal device [1]. For example, some financial applications in mobile phones or laptop are seeking security authorization through scanning user's palmprint. But user's identity sometimes can't be recognized correctly. The main reason is that palm's pose variation inevitably leads to rotation and translation problems [6]. In addition, palmprint image is often captured with contactless manner, and some uneven illuminations make image grayscale be not equalization [6]. Thus, degrade palm images would negatively affect the performance of palmprint recognition algorithms [7].

Since existing methods have difficulty in dealing with degraded palmprint images, we come up with a novel technique from biologically inspired methods. The main reason is that palm's pose variation in practical application inevitably leads to illumination, rotation and translation problems. However, it is not difficult for human vision systems to verify

a palmprint and eliminate the above variations. With the response mechanism of brain's visual cortex being increasingly revealed in recent years, it is argued that visual neurons at higher levels exhibit an increasing degree of invariance to above mentioned image variations [8]–[10]. Moreover, V1 can give stronger responses to light bars of certain directions. Also palmprint includes types of bar or lines. Therefore, we propose a biologically inspired transform solution, namely BIT, to mimic visual neurons' response about orientation edge and spatial frequency. Besides that, we also presented a palmprint enhancement procedure to preserve the essential elements of visual appearance for palmprint recognition. This operation is also coinciding with certain preprocessing stages founding the mammalian visual cortex. The advantage of this novel method can extract palmprint feature, and overcome the effect of illumination, rotation and translation variation. Subsequently, BIT feature is invariant, which shows promising results when matching two palmprints.

Visual information processing in human brain is quite complex. Vision signals from each eye are segregated into different neural layers in the LGN, and then are composed in V1. Neurons in V1 have particular orientation selectivity. One can mimic this mechanism with an edge detector at different orientation, then make a summation of every directional edge to simulate neurons aggregation in V1 [10], [11]. Besides orientation selectivity, neurons in V1 respond preferentially to spatial frequencies. Given an image, high spatial frequency represents edge detail information, and low spatial frequency involves object contour. These orientation selectivity, spatial frequency measurement and visual information mapping are performed with a hierarchical structure [11]. Inspired by these vision mechanisms, BIT feature extraction framework involves two stages. In the first stage, we build an orientation edge detector, and this detector involves two parts. The first part is the combination of Gabor filter banks and phase congruency based edge detectors, which mimics phase response of V1. Similar with the function of bipolar cells of visual cortex, we construct a bipolar filter that is composed of a horizontal filter and a vertical filter. This bipolar filter can detect edges in different directions and more importantly highlight the edges of corresponding directions. Then, a local spatial frequency detector is used to measure spatial frequencies at all directions and intervals. This procedure converts rotation variations of palmprint image into horizontal shifts in transform map. In the second stage, the edge detector and local spatial frequency detector are applied again, making the entire system invariant to translations. The extracted 2D feature maps are concatenated to form a final feature vector.

The whole process of BIT feature extraction simply consisting of two stages is shown in Fig.1. The first stage (shown in Fig.2) converts a rotated input image into a shifted map and then into an invariant feature map after the second stage. The translation invariance is also achieved in the first stage. In each stage, the orientation edge detector and local spatial frequency detector is conducted once. Orientation edge

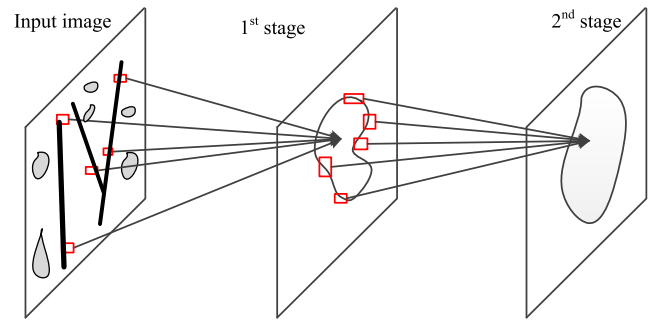


FIGURE 1. BIT feature extraction framework.

detector is utilized to highlight the object edge in a certain direction. Local spatial frequency detection is used for local spatial frequency analysis. The orientation sensitive operators are similar to V1 [8], [9], so our approach could also be understood as recognizing palmprint by simulating the early visual processing stages of primate vision.

To obtain edges of a palmprint image, a palmprint enhancement algorithm embedding Gabor filter is applied to remove illumination affection. After that, edges are detected by phase congruency algorithm. The edge is combined with the directional filter to estimate edge direction. These filters can be thought as a filter-filter structure [11], which functionally mimics the different response of each neuron.

We also build a local spatial frequency detector that measures local spatial frequencies. Then, the spatial frequencies at all directions and intervals are accumulated, resulting in a dense feature map consists frequency values.

A. RELATED WORK

Generally, palmprint recognition systems use a scanning device or a camera based equipment to acquire image data from an individual's palm and do verification according to stored features for that person [12]–[16]. During past several years, great efforts have been made to improve the performance of palmprint image acquisition systems. With various cameras equipped, these systems can obtain images of palmprints with resolution ranging from 100 to 500 ppi. Besides, some platforms can even deal with 3D palmprint images. Consequently, palmprint recognition algorithms are proposed to match images of different qualities [4], [6].

For high resolution palmprint images, most recognition algorithms typically follow the ridge and apply minutiae based matching strategy. Minutiae can be used to quality based and adaptive orientation field estimation algorithm, which can deal with a large number of creases [17]. In addition, Jain and Feng [18] proposed a latent-to-full palmprint matching algorithm and estimated the local ridge direction and frequency. In their work, ridge and minutiae features can be extracted even from palmprints of poor quality. In order to solve the distortion problem, a sequence of robust feature extraction approaches allows for reliably detecting minutiae with a local matching strategy [19]. Moreover, Liu et al. [20]

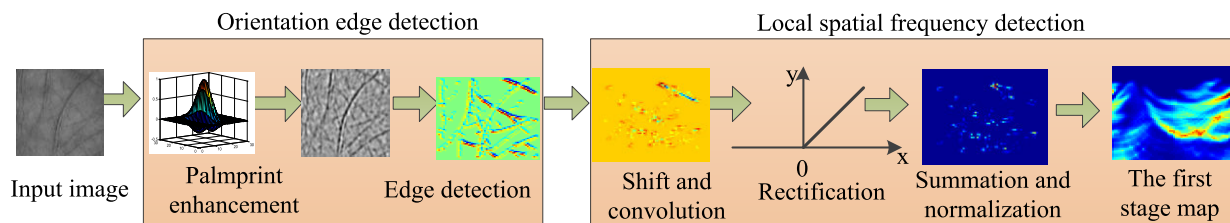


FIGURE 2. The first stage's transformation.

proposed a local feature based minutiae clustering algorithm, which divides minutiae into several groups. The coarse matching is then performed within each cluster to establish initial minutiae correspondences between two palmprints. In [21], a ridge line based matching and fusion algorithm is proposed to further handle the skin distortion, which strengthens the varying discrimination power of different palmprint regions. Although these methods achieve high recognition accuracy on some databases, feature extraction on high resolution images make less attention than low resolution images. The reason possibly lies in two aspects. One limitation is that palmprint image is captured by expensive equipment with contact manner. The other disadvantage can be attributed to minutiae's vulnerability when suffering from uneven illumination or palm deformation.

Further, Lula and Nardiello [22] built a 3D ultrasound palmprint recognition system that accounts for principal line depth. After that, palmprint is classified by an ad hoc matching criterion. Further progress is presented in [23] to make palmprint recognition systems be more usable. Since palmprint images are captured by two cameras, 3D features could be used to further refine the coarse matching results based on 2D features. In [24], Zhang *et al.* proposed a feature extraction scheme based on block-wise statistics. In this scheme, a cropped 3D palmprint ROI is divided into uniform blocks, and then a histogram of surface types within each block is calculated and concatenated to form a single feature vector. A 3D palmprint image involves more valuable information than 2D image. Therefore, 3D palmprint features are quite robust in some application fields. However, one main drawback of 3D palmprint image processing techniques is based on extra support systems, which will inevitably increase hardware cost.

From the perspective of computational cost, recognizing palmprints in low resolution images are more efficient than in high resolution or 3D images. Therefore, our approach will merely focus on palmprint recognition in low resolution images. Until recently, a number of methods have been proposed to solve this problem [12]–[14]. These approaches generally achieve palmprint recognition by making use of techniques including principal line, coding, texture, subspace learning, local descriptor methods and so on.

Conventionally, palm line is believed to serve as an important characteristic for recognition [12]. Li *et al.* [5] extracted

principal lines from palmprint image. Wu *et al.* [25] also proposed principal lines extraction approach in terms of palmprint characteristics. Since patterns of palm lines vary a lot even within a single palm, Palma *et al.* [26] conducted palmprint verification based on a dynamical system approach for principal palm lines matching.

In addition, coding methods are also proved to be effective in extracting palmprint features [27]. There are a number of coding based methods proposed to do the job, such as Double Orientation Coding [15], CompCode [28], [29], Palmprint Orientation Code [30], Hierarchical multi-feature Code [31], Robust Line Orientation Code (RLOC) [32], Ordinal Code [33], Fusion Code [34], Palm Code [35] and CR_CompCode [36]. Recently, Xu *et al.* [37] proposed a more discriminative and robust competitive code (DRCC), which uses a more accurate dominant orientation representation of palmprint images. This method essentially weights the orientation information of a neighbor area to improve the precision and stability of the dominant orientation code. Due to the importance of direction information, Jia *et al.* [38] proposed a complete direction representation method (CDR), which is a general framework for direction representation due to a comprehensive and complete way. Basically, the complete direction representation extract palmprint feature at different direction levels, scales and regions through modified finite radon transform. Fei *et al.* [39] proposed a double-layer direction extraction method for palmprint recognition. An apparent direction code is extracted by utilizing modified radon transform from the surface of a palmprint. Also, a latent direction code is exploited from the energy map layer of apparent direction. A histogram feature of palmprint is extracted by pooling the two apparent and local direction code (ALDC). Basically, most of palmprint coding features are extracted by masking Gabor filter on whole image. When a palmprint image is corrupted by illumination, rotation and translation, Gabor filters cannot remove these affections. Subsequently, the performance of coding feature will be inevitably impacted.

In order to handle image variations, palmprint texture based representation approaches have been considered as the promising methods [40], [41]. Specifically, Raghavendra and Busch [40] employed a sparse representation of features obtained from bank of binarized statistical image features (B-BSIF) to construct a texture descriptor. For subspace

learning methods, high dimensional palmprint image is mapped into a low dimensional space. Further, Zhu *et al.* [41] proposed a domain adaptive method based on low rank canonical correlation analysis, which is able to exploit some subspace feature information for palmprint images. Rida *et al.* [42] employed 2DPCA to build nearly incoherent random subspaces. Then, palmprint features are extracted in each subspace using 2D linear discriminant analysis. These techniques achieve good performance on some contact palmprint database. However, one primary limitation of texture based or subspace learning methods is that feature robustness will be worse when image is corrupted by illumination, rotation and scaling noise.

Local descriptor based methods use pixel intensity variations to encode a local representation of the image. In [43], Gaussian derivative phase pattern image and its block-wise histograms are concatenated to form a single vector referred as local descriptor. Li and Kim [44] presented a unique local microstructure tetra pattern to palmprint recognition. Based on the idea of local binary patterns(LBP), Luo *et al.* [14] proposed a local line directional patterns(LLDP) descriptor. Other methods such as sparse representations and SIFT based palmprint feature extractor also achieved quite good performance. Rida *et al.* [45] proposed a palmprint feature extraction method based on an ensemble of sparse representations through an ensemble of discriminative dictionaries. This method can reduce the sensitivity due to the limited size of the training data. Almaghtuf and Khelifi [46] presented a SIFT-based palmprint matching method, which can take into account the geometric relationship between SIFT points within the query image in comparison with the relationship of the corresponding matched points in the reference image.

Biologically inspired methods for feature extraction are numerous, so we will focus only on invariant feature extraction. In the early stage, model simulating V1 serves as an invariant feature extractor, and it is able to mimic simple and complex vision cells [47]. Each element of this model is a complex feature obtained by combining position and scale tolerant edge detectors over neighboring positions and multiple orientations. The outputs of all simple and complex cells are concatenated to form a single vector, considered as the V1 representational pattern of each image. Consequently, V1-like model is scaling and translation invariant. Similar to V1 model, there are other more complex biological invariant feature extraction methods. VisNet is a training model of visual pathway for invariant object recognition [9], [48]. This model, taking advantage of trace learning rule, can produce translation invariant representations. HMAX is also a biological vision model that involves computational units of four layers [8], [49]. The C units(max pooling layer) of this model perform a nonlinear maximum pooling operation over the units to impose translation and scale invariance. To deal with difficulties in training, Soutsov *et al.* [50] proposed another invariant feature extraction approach. In this approach, the representation is designed to be invariant to position, scaling, and rotation, so that the model could be regarded as early

visual processing stages of primate vision. However, there are two major limitations of this method. The first shortcoming is that the model is sensitive to noise, and the other is that the edge detector has poor discriminability for objects of simple structures, which may damage the overall recognition accuracy.

Recently, deep learning techniques have become popular in palmprint recognition [51]–[53]. Among these methods, all kinds of convolutional neural networks(CNNs) are typically used to extract palmprint features, and a distance measure or a trained classifier is then adopted to match the palmprint templates. Minaee and Wang proposed a deep scattering network presented in [54], which can process input images using a bank of fixed filters based on the scattering transform; then, an SVM [54] was used to classify the palmprint images. Besides, Genovese *et al.* [55] also proposed a CNNs based palmprint feature extractor, which was embedded in Gabor and principle component analysis(PCA) filters. The k-nearest neighbors(KNN) classifier based on the Euclidean distance was used to classify palmprint images. Further, Wang *et al.* [52] proposed a multi-weighted co-occurrence descriptor. This approach embedded co-occurrence filters among CNNs, and used large margin distribution machine to classify palmprint images. Other pretrained or trained CNN models including AlexNet, VGG-16 and VGG-19 are also widely adopted to extract palmprint features [51], [56]. Theoretically, CNNs of multi-layers could be considered as a modeling of vision systems, as they simulate the hierarchical structure of primate vision [57]. While the recognition task is complex, the layers of model are becoming deeper. Since convolutional layers are usually followed by subsampling layers that perform local pooling and subsampling operations, the feature maps generated are invariant to small input shifts. For rotation invariant image recognition, Cheng *et al.* [58], [59] proposed some novel CNNs model, which mainly introduced and learnt a new rotation-invariant layer on the basis of the existing CNN architectures. In addition, they trained rotation invariant and Fisher discriminative CNN models to further boost image recognition performance [60]. Nevertheless, traditional CNNs model suffers from the need of massive training images to achieve high recognition accuracy. Compared with large-scale image datasets, the scale of the existing palmprint datasets are very much smaller. In addition, training of CNNs is computationally expensive and the powerful graphical processing unit(GPU) is usually needed. Therefore, insufficient exploitation of CNNs by way of training or parameter tuning may limit the applications of palmprint recognition.

For vision perception, when a stimulus is appeared, preliminary processing takes place in retina where basic features are detected [8], [47]. Then, the feature signals are transmitted as neural spikes along the optic nerve. These features include visual patterns such as edges, orientations, gradient information and so on [8]–[10]. Various stages of functionalities in vision cortex are combined to form a transformation that is able to extract invariant feature regardless of

its scale, position and orientation [10], [49], [50]. Meanwhile, individual neuron recorded in the visual cortex of animals exhibits orientation tuning, i.e., they respond more vigorously to stimuli of a certain orientation [9], [61]. If an awake macaque is repeated to stimuli by using a bar with different direction, one can analyze the tuning properties of a neuron recorded in V1 [47], [48]. Therefore, it is important to select a properly designed filter for 2D edge detection. Moreover, bipolar cells are functionally crucial neurons that comprise the middle components of the vertical transduction pathway through the retina [47]. Each bipolar pathway is capable of independent image processing. In other words, bipolar cells are more commonly used in human vision systems because horizontal and vertical lines are more important in human vision. Hence, an obvious way of obtaining edges of various orientations is to combine horizontal and vertical directions.

Besides orientation selectivity, neurons in V1 also respond preferentially to stimuli with distinct spatial frequencies. In the study of visual perception, sinusoidal gratings are frequently used to probe the capabilities of the visual system. In sinusoidal gratings stimuli, spatial frequency is expressed as the number of cycles per degree of visual angle. As for an image, high spatial frequency represents image edge information, and low spatial frequency involves object contour characteristics [50], [61]. In another word, V1 operates with a code of spatial frequencies. One can build a local spatial frequency detector to probe spatial position relationships for palmprint edges and contours in the receptive field. The spatial frequency also involves palmprint structure information, which will be important to identification.

Human can recognize rotated and translated object through vision signal sensing and processing. One important factor is that there exists a mapping from receptive field to visual respond in V1. Some earliest cortex mappings come from experiments and observations by neurologists. In the present study, many researches use high-resolution functional magnetic resonance imaging (fMRI) to measure the shape and morphometry of V1 to the center of the visual receptive field. The overall relationships between visual field position and position in the cortex are described approximately by a complex logarithmic function [62], [63]. According to these observations, it is feasible to extract image features by mimicking the visual mappings from palmprint images to feature maps.

B. CONTRIBUTIONS AND ORGANIZATION

The contributions of our work are four-fold. (1) A two-stage palmprint feature extraction framework is proposed. This framework is able to mimic visual object response of simple cell in visual primary cortex, which typically makes the feature be invariant to image corruption. (2) An orientation edge detector is designed to highlight all edge responses on each orientation for palmprint image, which is able to imitate receptive fields in human vision system. (3) Palmprint local spatial frequency detector is built to measure the spatial frequency at each spatial distance and orientation, which mimics

the neuron spike to produce a response of double orientation edges appeared. (4) Extensive experiments have been conducted on several palmprint databases, including PolyU II, PolyU multispectral, CASIA, COEP and TongjiU database, and the results on corrupted images demonstrate the effectiveness of our method.

The rest of paper is organized as follows: Section 2 introduces palmprint enhancement and phase congruency based edge detectors. In section 3, palmprint spatial frequency measurement methods including orientation edge detection with bipolar filters and local spatial frequency detection detector are discussed. Section 4 explains palmprint matching methods. In section 5, experimental results including comparison and analysis in terms of effectiveness and performance are presented. Section 6 gives a summary of the previous BIT feature extraction approaches.

II. ORIENTATION EDGE DETECTION

A. PALMPRINT ENHANCEMENT

During palmprint image acquiring, some illumination variations, local shadowing or highlights are often inevitably introduced. Thus, a palmprint enhancement procedure is employed to preserve the essential elements of visual appearance for palmprint image [64]. This operation is also coinciding with certain preprocessing stages founding the mammalian visual cortex.

In order to enhance local dynamic range of palmprint image, we replace gray level I of original image with Gamma correction I^τ . The gamma value τ is an encoding parameter, which can compress or expand gray level with power-law nonlinearity. Small τ value will make image has high contrast and brightness. Conversely, large τ value will lead to low contrast and brightness. Due to human physiological structure, the contrast between palmprint lines and palm is always low in captured image. When highlighting image details for each palmprint database, we find that τ equals 0.2 is the most appropriate value. However, Gamma correction is not able to remove shading region caused by uneven illumination. Therefore, we globally rescale the palmprint image intensities to standardize a robust measure of overall contrast or intensity variation. It is important to use a robust estimator because the palmprint image typically still contains a small extreme values produced by illuminations. In order to speed up computation, we employ a rapid approximation based on a two stage process:

$$J(x, y) = \frac{I(x, y)}{(\text{mean}(|I(x, y)|^\mu))^{1/\mu}} \quad (1)$$

$$L(x, y) = \frac{J(x, y)}{(\text{mean}(\min(\delta, |J(x, y)|)^\mu))^{1/\mu}} \quad (2)$$

where, μ is a compressive exponent that reduces the influence of large values. When μ is large, all shading regions of palmprint image can be removed, but the contrast of palmprint details are reduced. Small μ value can not only keep high contrast, but also remove some highlighting regions. δ is a threshold used to truncate large values after the first phase

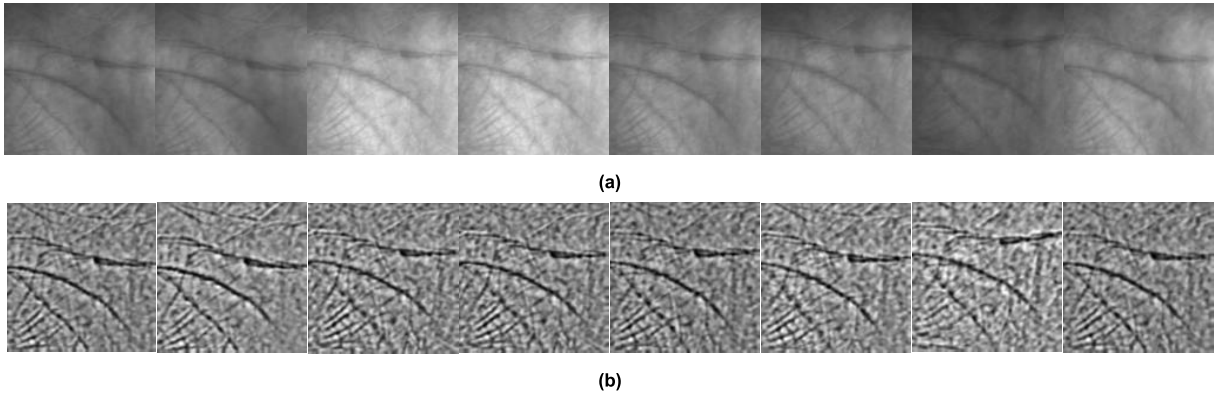


FIGURE 3. Palmprint enhancement. (a) original image with various illuminations, (b) palmprint image after enhancement.

of normalization, and the mean is over the whole image. Small δ value holds more palmprint details, but inevitably narrows image dynamic range, and vice versa. Although there are two new parameters, these parameters have little effect on palmprint feature extraction. Basically, palmprint feature information changes very little when $\mu < 0.6$ and $\delta > 7$. Consequently, we set $\mu = 0.1$ and $\delta = 10$ throughout the experiments.

To further remove extreme values, we adopt a nonlinear function (i.e. hyperbolic tangent) to compress over large values. This operation limits L to the range $(-\delta, \delta)$.

$$F(x, y) = \delta \tanh\left(\frac{L(x, y)}{\delta}\right) \quad (3)$$

Palmprint principal lines structure is dominant low pass information that is often hard to be separated from illumination. Some high pass information such as palmprint detail is also important for palmprint recognition. Gabor filter can be used for band pass filter, and has been extensively used to model the receptive fields of simple cells for decades [58]. Subsequently, we build a real Gabor filter at all orientation and wavelength to preserve principal lines and detail lines. In the spatial domain, real Gabor wave is given as follows [65].

$$G_{\theta, \sigma}(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(i\left(2\pi \frac{x'}{\lambda} + \varphi\right)\right) \quad (4)$$

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta \\ y' &= -x \sin \theta + y \cos \theta \end{aligned}$$

In above equations, x, y are the coordinates of pixel positions in the image. λ denotes the wavelength, θ represents the filter orientation, φ is the phase, σ denotes the standard deviation, and γ is the spatial aspect ratio that specifies the ellipticity of Gabor wave.

To smooth palmprint edge at all width and orientation, a Gabor filter bank is defined at each sample point corresponding to the center of a Gabor filter in the spatial domain. Let $\Theta = \{\theta_1, \dots, \theta_D\}$ be a set of D orientations. Given a sample point (x, y) and an orientation $\theta_i \in \Theta$, we define a

set of Gabor filters at the fixed frequency, which is located at (x, y) in the spatial domain, such as

$$G_{ij} = \{G_{11}, \dots, G_{KL}\} \quad (5)$$

where $G_{ij} = G_{ij}(x, y) = G(x, y, \theta_j, \lambda_i)$, $\sigma = 4$, $\varphi = 0^\circ$, $\gamma = \sqrt{2}$.

Therefore, we had 16384 (128 bandwidths \times 128 orientations) Gabor filters to smooth image edge, which consisted of a Gabor bank. Fig.4 shows some filter kernels and their filtered images with $\theta_l = \{0^\circ, 30^\circ, 60^\circ, 120^\circ, 150^\circ\}$ for a bandwidth. In this picture, the palmprint image is filtered with six real Gabor filters to generate six filtered images. Each of the filtered images highlights the prominent palmprint lines and creases in corresponding direction while suppressing background noise and structures in other directions.

B. EDGE DETECTION BASED ON PHASE CONGRUENCY

With regard to edge detector, palmprint lines should be firstly detected as many as possible. As a biologically plausible edge detector, phase congruency algorithm is invariant to image illumination or contrast variations, which lead to an excellent performance in locating palmprint edges [66].

From signal processing's point of view, any signal can be decomposed by Fourier series, and Fourier components $FC(x)$ are sine waves in phase at the point of the step. Thus, congruency of phase at any angle produces a clearly perceived feature, formulated as follows,

$$FC(x) = \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin\{(2n+1)x + \phi\} \quad (6)$$

where ϕ is the offset at which congruence of phase occurs, varying from 0 to $\pi/2$.

The phase congruency function in terms of the Fourier series expansion of a signal at some location x is defined as the following equation [60],

$$PC(x) = \max_{\bar{\phi}(x) \notin [0, 2\pi]} \frac{\sum_n A_n(x) \cos(\phi_n(x) - \bar{\phi}(x))}{\sum_n A_n(x)} \quad (7)$$

where A_n is the amplitude of the n th Fourier component, and $\phi_n(x)$ denotes the local phase of the Fourier component at

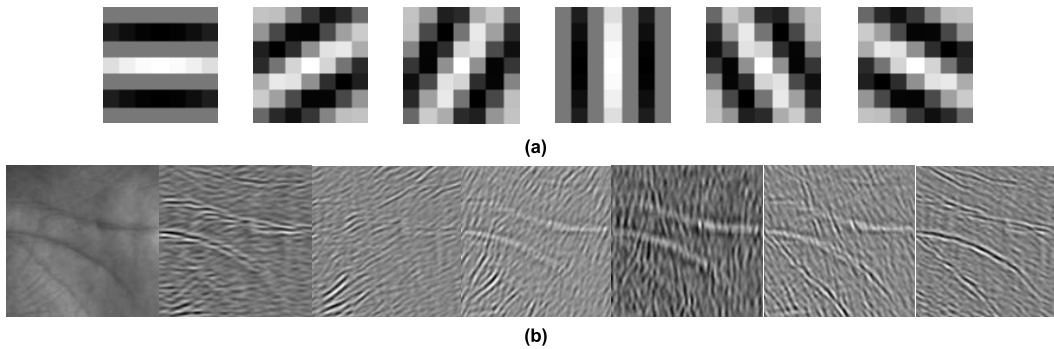


FIGURE 4. Real Gabor filters implementation with six orientations. (a) real Gabor filter kernel with six orientation, (b) palmprint and six filtered images.

position x . The value of $\phi(x)$ that maximizes this equation is the amplitude weighted mean local phase angle of all the Fourier terms at the point being considered.

The maximum phase congruency can be calculated by searching for peaks in the local energy function as follows,

$$EN(x) = \frac{\sum_n A_n(x) \cos(\phi_n(x) - \bar{\phi}(x))}{\sqrt{Z^2(x) + H^2(x)}} \quad (8)$$

where $Z(x)$ is a signal with its direct current(DC) component removed, and $H(x)$ is the Hilbert transform of $Z(x)$.

The approximations of $Z(x)$ and $H(x)$ are obtained by convolving the signal with a quadrature pair of filters. Furthermore, the energy can be transformed into phase congruency scaled by the sum of the Fourier amplitudes. Consequently, the local energy function $EN(x)$ is directly proportional to the phase congruency function.

$$EN(x) = PC(x) \sum_n A_n(x) \quad (9)$$

Phase congruency $PC(x)$ is ill-conditioned if all the Fourier amplitudes are very small. This problem can be solved by adding a small positive constant ε to avoid division by zero.

$$PC(x) = \frac{EN(x)}{\sum_n A_n(x) + \varepsilon} \quad (10)$$

In order to reduce noise effect and high frequency components in the signal, phase congruency is modified by the following formula,

$$PC(x) = \frac{W(x) |EN(x) - \zeta|}{\sum_n A_n(x) + \varepsilon} \quad (11)$$

where $W(x)$ is a phase congruency weighting function that can be constructed by applying a sigmoid function to the filter response spread value, and ζ denotes the estimated noise influence.

To further improve localization for poor and blurred feature, one can construct a more sensitive phase deviation measure as follows.

$$\Delta\Phi(x) = \cos(\phi_n(x) - \bar{\phi}(x)) - |\sin(\phi_n(x) - \bar{\phi}(x))| \quad (12)$$

Using the measure of phase deviation $\Delta\Phi(x)$, the phase congruency can be rewritten as follows:

$$PC(x) = \frac{\sum_n W(x) |A_n(x) \Delta\Phi_n(x) - \zeta|}{\sum_n A_n(x) + \varepsilon} \quad (13)$$

where ε is a small positive constant to avoid division by zero, and ζ is the estimated noise influence.

For a 2D image, the energy $EN(x)$ could be calculated in each orientation of every location in the image, and then noise could be compensated by subtracting the estimated radius of noise circle. Meanwhile, the weighting for frequency spread is employed to form the sum over all orientations. The sum of energy terms could be normalized by dividing by the sum over all orientations and scales of the individual filter responses amplitudes at that location in the image.

C. ORIENTATION EDGE DETECTOR

In order to improve discrimination ability of feature map, orientation information of palmprint edges are quite important. The reason is that neurons in V1 have particular orientation selectivity. From the perspective of vision mechanism, bipolar cells receive inputs from a set of photoreceptor cells that define the center-surround receptive field. Center-surround receptive fields arise from a pool of photoreceptors on-center and off-center fields in retinal bipolar and ganglion cells form by pooling the response of groups of photoreceptors. The photoreceptors can either act to excite or to inhibit a downstream bipolar cell. In an on-center bipolar cell, light hitting the central photoreceptors will be excitatory and light in the surround will be inhibitory. In an off-center bipolar cell, light in the center will be inhibitory, and light in the surround will be excitatory [67]. Inspired by the above mechanism, we built bipolar filters to highlight edge responses at each orientation, which can mimic the orientation preference of lines in V1. More significantly, the orientation edges have both positive and negative parts, which is consistent with the response of on/off center bipolar cells for a stimuli line. The bipolar filters consist of 1×3 and 3×1 sub-filters, which are composed of triangle function. With the two sub-filters, edge response about an orientation can be obtained by convoluting two sub-filters with the original input image, respectively.

Horizontal filter is a filter of size 1×3 . Here we use cosine functions to construct the horizontal filter \mathbf{Hb}_θ .

$$\mathbf{Hb}_\theta = [-\cos(\theta) \quad 1 - |\cos(\theta)| \quad \cos(\theta)] \quad (14)$$

Although the horizontal filter is constructed, there are some shortcomings. For instance, the sum of \mathbf{Hb}_θ is 0 while $\theta = 0^\circ$, the pixels at the horizontal direction will be removed, which will corrupt the horizontal edges. To resolve the problems, each element in the filter is weighted by the following step function.

$$S(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (15)$$

The step function is set for the weighting factor, so the horizontal sub-filter equals the dot product of \mathbf{Hb}_θ and $S(\mathbf{Hb}_\theta)$.

$$\begin{aligned} \mathbf{F}_{\theta 1} &= \mathbf{Hb}_\theta * S(\mathbf{Hb}_\theta) \\ &= \begin{bmatrix} -\cos(\theta) \times S(-\cos(\theta)) \\ (1 - |\cos(\theta)|) \times S(1 - |\cos(\theta)|) \\ \cos(\theta) \times S(\cos(\theta)) \end{bmatrix}^T \end{aligned} \quad (16)$$

where, $*$ denotes element-by-element multiplication. Since $1 - |\cos(\theta)| \geq 0$, the filter can be further rewritten to the following equation,

$$\mathbf{F}_{\theta 1} = \begin{bmatrix} -\cos(\theta) \times S(-\cos(\theta)) \\ 1 - |\cos(\theta)| \\ \cos(\theta) \times S(\cos(\theta)) \end{bmatrix}^T \quad (17)$$

where, $\sum \mathbf{F}_{\theta 1} = 1$, while $\theta \in [0, 180^\circ)$. The horizontal sub-filter that consists of cosine functions has two advantages. Firstly, the filter can enhance horizontal direction edges while suppress vertical direction edges. Secondly, the filter is vertically symmetrical in a cycle, so we only need to detect edges whose directions are within $[0, 180^\circ)$, improving the computational efficiency.

Similarly, we utilize sine functions to build the vertical bar, which is a filter of size 3×1 .

$$\mathbf{Vb}_\theta = \begin{bmatrix} -\sin(\theta) \\ 1 - |\sin(\theta)| \\ \sin(\theta) \end{bmatrix} \quad (18)$$

With the step function, the vertical sub-filter can be represented by the dot product of \mathbf{Vb}_θ and $S(\mathbf{Vb}_\theta)$.

$$\begin{aligned} \mathbf{F}_{\theta 2} &= \mathbf{Vb}_\theta * S(\mathbf{Vb}_\theta) \\ &= \begin{bmatrix} -\sin(\theta) \times S(-\sin(\theta)) \\ (1 - |\sin(\theta)|) \times S(1 - |\sin(\theta)|) \\ \sin(\theta) \times S(\sin(\theta)) \end{bmatrix} \end{aligned} \quad (19)$$

where, $*$ denotes element-by-element multiplication.

Since $1 - |\sin(\theta)| \geq 0$, the filter could be further expressed by the following equation.

$$\mathbf{F}_{\theta 2} = \begin{bmatrix} -\sin(\theta) \times S(-\sin(\theta)) \\ 1 - |\sin(\theta)| \\ \sin(\theta) \times S(\sin(\theta)) \end{bmatrix} \quad (20)$$

where $\sum \mathbf{F}_{\theta 2} = 1$, and $\theta \in [0, 180^\circ)$.

Obviously, equation (17) is horizontally symmetrical in a cycle, so we also only need to detect the edge of $[0, 180^\circ)$ in a period.

Since $\mathbf{F}_{\theta 1}$ and $\mathbf{F}_{\theta 2}$ are both vectors, the convolution is equal to the product of them. Thus, the size of 3×3 orientation edge filter is composed by the bipolar filter.

$$\begin{aligned} \mathbf{F}_\theta &= \mathbf{F}_{\theta 1} \otimes \mathbf{F}_{\theta 2} = \mathbf{F}_{\theta 2} \times \mathbf{F}_{\theta 1} \\ &= \begin{bmatrix} f_{11}(\theta) & f_{12}(\theta) & f_{13}(\theta) \\ f_{21}(\theta) & f_{22}(\theta) & f_{23}(\theta) \\ f_{31}(\theta) & f_{32}(\theta) & f_{33}(\theta) \end{bmatrix} \end{aligned} \quad (21)$$

where,

$$\begin{cases} f_{11}(\theta) = \sin(\theta) \times \cos(\theta) \times S(-\sin(\theta)) \times S(-\cos(\theta)) \\ f_{12}(\theta) = \sin(\theta) \times |\cos(\theta)| \times S(-\sin(\theta)) - \sin(\theta) \times S(-\sin(\theta)) \\ f_{13}(\theta) = -\sin(\theta) \times \cos(\theta) \times S(-\sin(\theta)) \times S(\cos(\theta)) \\ f_{21}(\theta) = -\cos(\theta) \times S(-\cos(\theta)) + |\sin(\theta)| \times \cos(\theta) \times S(-\cos(\theta)) \\ f_{22}(\theta) = 1 - |\sin(\theta)| - |\cos(\theta)| + |\sin(\theta) \times \cos(\theta)| \\ f_{23}(\theta) = \cos(\theta) \times S(\cos(\theta)) - |\sin(\theta)| \times \cos(\theta) \times S(\cos(\theta)) \\ f_{31}(\theta) = -\sin(\theta) \times \cos(\theta) \times S(\sin(\theta)) \times S(-\cos(\theta)) \\ f_{32}(\theta) = \sin(\theta) \times S(\sin(\theta)) - \sin(\theta) \times |\cos(\theta)| \times S(\sin(\theta)) \\ f_{33}(\theta) = \sin(\theta) \times \cos(\theta) \times S(\sin(\theta)) \times S(\cos(\theta)) \end{cases}$$

In above equation, all elements are determined by θ . Consequently, the elements in the first row are always zero when $\theta \in [0, 180^\circ)$. Similarly, the elements in the last row are also zero when $\theta \in [180^\circ, 360^\circ)$. It is not difficult to see that the composited filter is centrosymmetric in a cycle, so the range of $[0, 180^\circ]$ for θ is enough to detect all directional edges. The orientation edge detector can be built by a convolution between phase congruency edge and bipolar filter.

$$\mathbf{E} = \mathbf{PC} \otimes \mathbf{F}_\theta \quad (22)$$

where \otimes denotes convolution operation.

The outputs of the edge detection on some orientations are shown in Fig. 5. In these edge maps, red color represents positive values and blue stands for negative value. Deeper color means larger absolute values.

III. LOCAL SPATIAL FREQUENCY DETECTION

The next procedure in the first stage is applying a local spatial frequency detector looking for orientation edges separated by interval I at the same orientation. As shown in equation (23), the local spatial frequency detector R is built to output a neuron spike if the orientation edge map E for a given orientation θ has two edges separated by an interval I . Given an interval I and an angle θ , the orientation edge map E is shifted by I at angle $\theta + 90^\circ$ and multiplied by itself. This multiplication operation ensures that there is no spike if only

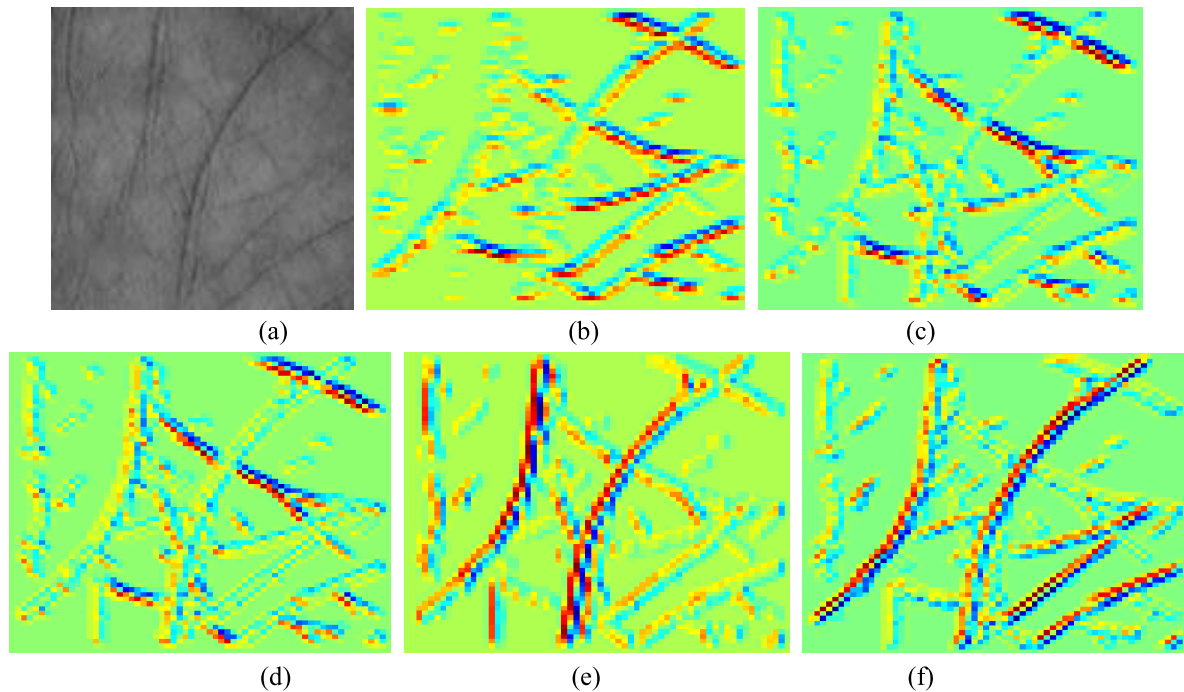


FIGURE 5. Palmprint edge of different orientation. (a) original image, (b) 0°, (c) 30°, (d) 45°, (e) 90°, and (f) 135°.

a single edge has appeared. All pixels in the image after multiplication are then accumulated. The accumulated values are then normalized by the squared sum of orientation edge map E ,

$$\mathbf{R}_{\theta,I} = \frac{\sum \mathbf{E} \times \mathbf{E}_{\theta+90,I}}{\|\mathbf{E}\|} \quad (23)$$

where I is interval value, and $\theta \in [0, 180^\circ)$.

Since the output of a neuron is a spike with non-negative rates [50], a half-wave rectification is applied to set negative value to zero. Zero values are generated by the convolution of positive and negative edges. Here, the output map of local spatial frequency is rectified using the following Heaviside function.

$$\mathbf{T}_{\theta,I} = \text{Rect}(\mathbf{R}_{\theta,I}) = \begin{cases} \mathbf{R}_{\theta,I}, & \text{if } \mathbf{R}_{\theta,I} > 0 \\ 0, & \text{if } \mathbf{R}_{\theta,I} \leq 0 \end{cases} \quad (24)$$

When image shift operation produces fractional pixel coordination, the new pixel must be generated by analyzing the surrounding pixels. Here, the bilinear interpolation algorithm is utilized to deal with this problem.

The local spatial frequency detector is applied to all positions in the subregion, and the outputs over this subregion are summed up. These summation values, for a range of orientations and intervals, are concatenated in a map of orientation and log interval. Three local spatial frequency detection processes with different directions and intervals are illustrated in Fig 6. Fig 6.(a) reveals an edge map and shifted edges, and shift interval values set to 15 and θ set to 135° . Obviously, some regions are positive, and some are negative at different direction. The sum of superposition maps is normalized by

input edge, which is correspondent with a pixel of the first transform map. Fig.6(b) and Fig.6(c) depicts the other two local spatial frequency detection process. The interval values is 30, and θ equals to 135° in Fig.6(b). The interval value is 15, and θ equals to 30° in Fig.6(c). In these feature maps, red color regions denote large feature values, and the largest value is located in dark red regions. By contrast, blue indicates small value, and dark blue of the background is zero.

During the transformation in the second stage, orientation edge detection and local spatial frequency detection are conducted again and the left and right small part in the image is exchanged since the first map is periodic. Compared to interval values in the first stage, the interval range in the second stage varies from 15 to 85 percent of the first feature map size. This scope ensures that most of vision information in receptive fields will be processed. The range of direction angle is still the same as the first stage. After the second local spatial frequency detection, the spike intensity is normalized to $[0,1]$.

Fig.7 shows the second local spatial frequency detection process with the orientation of 45° and interval value set to 15. Fig.7(a) is the feature map from the first stage. Fig.7(b) shows the overlap of the original edge and shifted edge, and the product of two edge maps is shown in Fig.7(c). In Fig.7(d), the palmprint image is transformed into a 64×64 feature map, in which the red color represents high intensity, and red colored region contains the primary image feature. On the contrary, blue color represents low intensity where there is little feature information, and dark blue of the background is zero.

The detailed procedure of our solution is described in Algorithm 1. In our implementation, most of operations including Gabor convolution, directional edge detection and

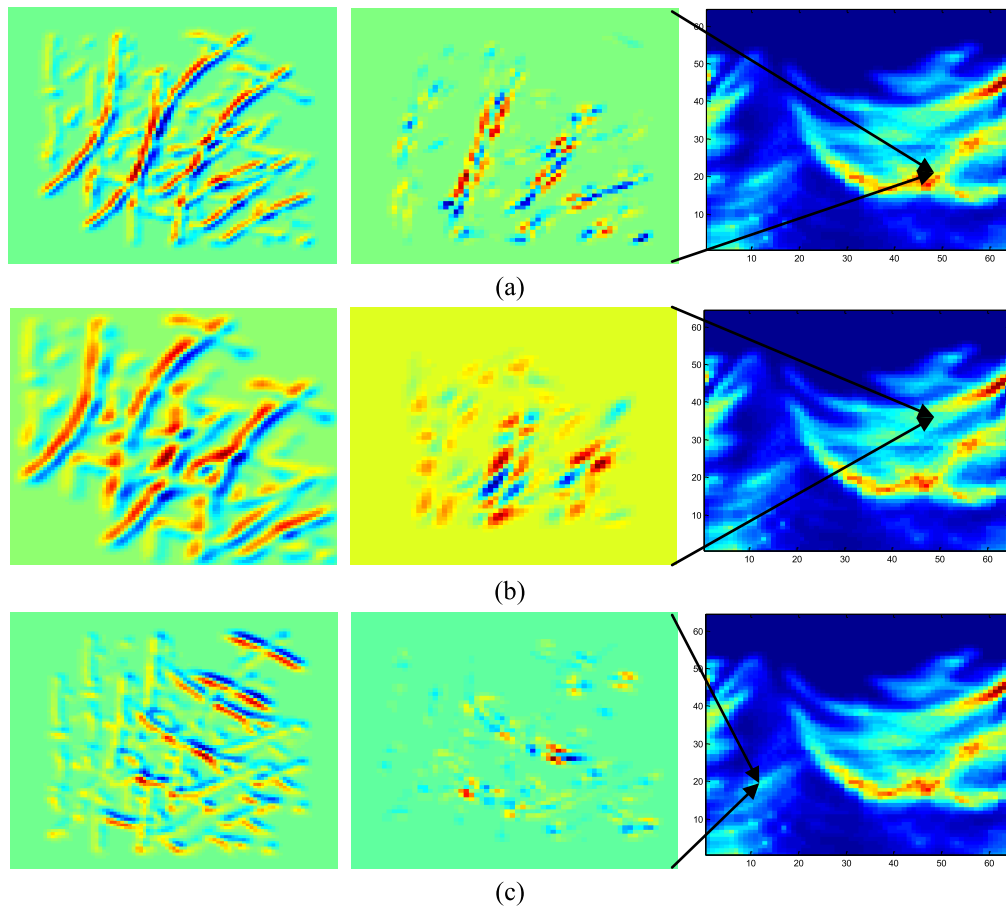


FIGURE 6. Local spatial frequency detection at different directions and intervals in the first stage. (a) Interval values is 15, θ is 135° , (b) interval values is 30, θ is 135° , and (c) interval values is 15, θ is 30° .

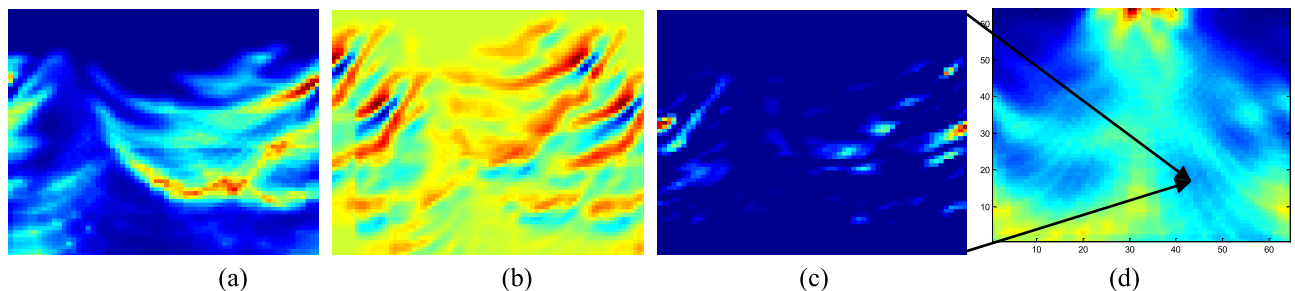


FIGURE 7. The second stage's transform. (a) The first stage's map, (b) overlap of original and shifted edge, (c) production of edge maps, and (d) the second stage's map.

local spatial frequency detection are identical in two stages. Edge detection based on phase congruency is only performed in the first stage.

IV. PALMPRINT MATCHING

There are several typical palmprint matching algorithms such as support vector machine (SVM) [54], collaborative representations classification (CRC) [57], Chi-square distance [39] and k-nearest neighbor (KNN) [55]. SVM [54] is a discriminative classifier formally defined by a separating hyperplane. Given labeled training feature, the algorithm outputs an optimal hyperplane which categorizes new samples. There are some parameters such as kernel, regularization,

gamma and margin that need to be selected. CRC [57] is an effective technique for classification of palmprint images. Usually, CRC [57] uses sparse representation and learned redundant dictionaries to classify image. Therefore, SVM and CRC [57] methods must be trained with a supervised procedure, and each database must be trained separately. Unlike with SVM [54] and CRC [57], Chi-square distance is one of the distance measures that can be used as a measure of dissimilarity between two feature vectors and has been widely used in palmprint recognition. KNN [55] is also a popular classification algorithm that stores all available templates and classifies new sample based on a distance measure.

Algorithm 1 Feature Extraction using Biologically Inspired Transform (BIT)

```

im ← palmprint image;
in ← interval number;
on ← orientation number;
for i: = 1 to in-1
  for j: = 1 to on-1
    g ← palmprint image enhancement(i,j,im);
    p ← edge detection(i,j,g);
    oe ← directional edge detection(i,j,p);
    the first map(i,j) ← local spatial frequency
    detection(i,j,oe);
  end for
end for
for i: = 1 to in-1
  for j: = 1 to on-1
    g ← Gabor convolution(i,j,the first map);
    oe ← directional edge detection(i,j,g);
    the second map(i,j) ← local spatial frequency
    detection(i,j,oe);
  end for
end for
return the second map;

```

In general, palmprint orientation and local spatial frequency are very obvious features. In this framework, we thoroughly exploited each palmprint orientation and local spatial frequency through numbers of convolutions with orientation edge filter and spatial frequency filter. Further, there are rectification, sum pooling and normalization operations as deep learning. Comparing with other deep learning methods, our framework has fewer layers. Theoretically, the extracted features from biologically inspired transform are belonging to deep features, which have high degree of discrimination. In addition, the invariant property of palmprint feature can be achieved by two stage's transform. Consequently, the invariant feature is insensitive to illumination, rotation and translation variations. Thus, palmprint feature can be separated by a KNN classifier.

When matching palmprint feature with KNN [55], we search for the best value of k between the 1 and 10 using an inverse weighting method($1/\text{distance}$). Intuitively, it seems like it would provide more robust results as each neighbor has less influence from the training sample. A distance function is employed to calculate between the new sample and each neighbor, and then the new palmprint feature is assigned to the class of k closest neighbor. With regards to the convenience of validating the effectiveness of the extracted invariant features, we first reshape each 2D feature map to a vector of size 1×4096 .

For KNN, the Euclidean distance is typically used to measure absolute distance differences [55]. Therefore, it measures distance between corresponding elements of two feature vectors without adjustment for differences in scale. Instead of

Euclidean distance dealing with palmprint feature, a correlation distance is employed to measure the trend or similarity between two feature maps. It is invariant under admissible feature transformations, e.g. little changing in scale. Let U be a new feature vector and V a template respectively. $U = \{u_i | i = 1, \dots, N_i\}$, $V^k = \{v_i^k | i = 1, \dots, N_j\}$, and $k = 1, \dots, K$, $N_i, N_j \in \mathfrak{R}$. The correlation distance between two feature vectors is calculated by the difference between the new feature and the template, which could be written in equation (25).

$$d(U, V^k) = 1 - \frac{(U - \bar{U})(V^k - \bar{V}^k)}{\sqrt{(U - \bar{U})(U - \bar{U})'} \sqrt{(V^k - \bar{V}^k)(V^k - \bar{V}^k)'}} \quad (25)$$

The size of BIT feature is constant regardless of the size of the image, in our case the dimensionality of U and V^k is fixed to be 4096. This property helps our method outperform other approaches when the input image is large. For the purpose of comparison, $\min(N_i, N_j)$ is adopted in our experiments.

For verification, the similarity score between the i th and the j th samples is measured as,

$$\text{Score}(i, j) = 1 - d_k \quad (26)$$

According to the above equation, the perfect matching score is 1 when corresponding feature maps of two palmprints are the same. On the other hand, the maximal score 0 will occur when two palmprints are totally different. In this case, there is no matched pixel in two feature maps.

For palmprint identification, the feature vector U is classified as belonging to the subject k to whose reference feature that has the minimal distance. Notably, one can produce several subjects by ranking candidates.

$$O = \arg \min_{k \in K} (d_k) \quad (27)$$

V. EXPERIMENTAL RESULTS**A. PALMPRINT DATABASES AND EXPERIMENTAL ENVIRONMENT**

In this section, a series of experiments are conducted on five popular palmprint databases, including PolyU II, PolyU multispectral, CASIA, COEP and TongjiU database, to test the performance of the proposed approach.

All palmprint images are preprocessed before feature extraction. This procedure extracts the central region of the palm. We use the most representative method proposed in [68] to extract the region of interest(ROI). This method uses the valley between fingers as reference points to determine the ROI. Firstly, the input palmprint image is convoluted with a low-pass filter and then converted into a binary image. Secondly, the boundaries and valley points are obtained using a boundary tracking algorithm, where the valley points are at the bottom of gaps between index and middle fingers

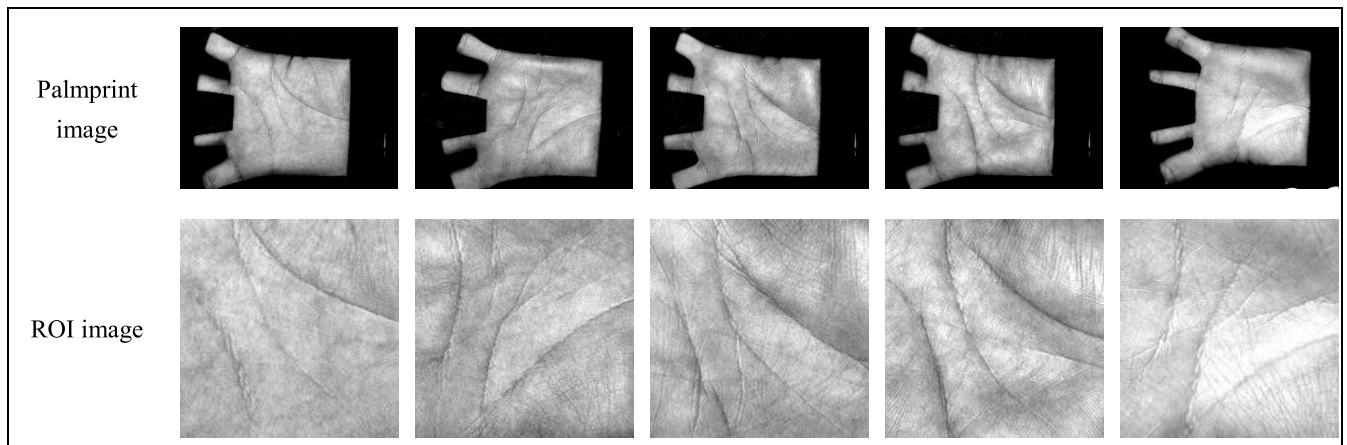


FIGURE 8. Palmprint images and ROI of PolyU II database.

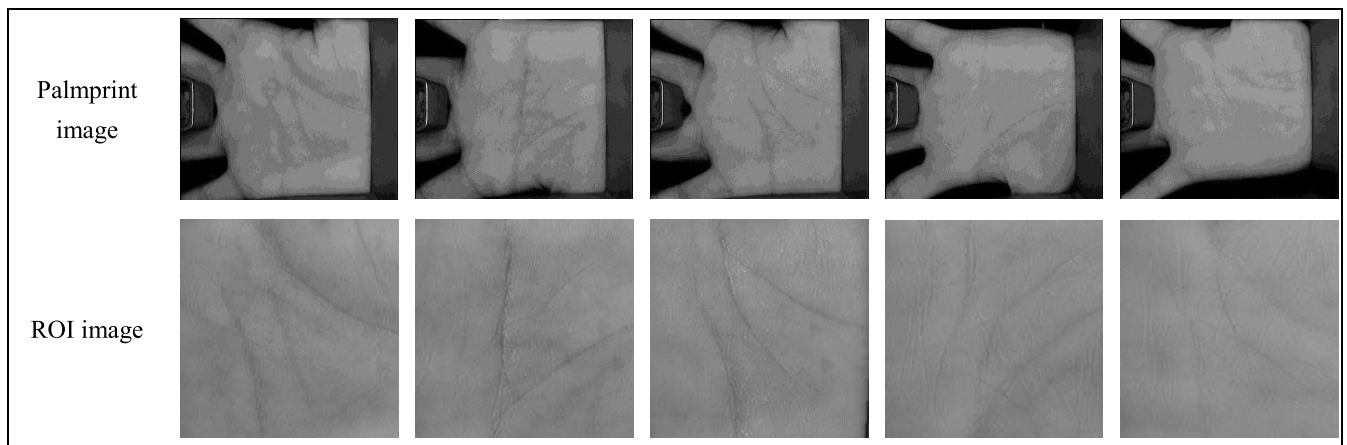


FIGURE 9. Palmprint images and ROI of PolyU multispectral database.

and between ring and little fingers. Thirdly, we locate the perpendicular bisector of the line segment between two valley points to determine the centroid of the palmprint region. Finally, we crop the sub-image as the ROI, which is located at the central area of a palmprint and used for the palmprint feature extraction. Since the distance between the palm and the camera is not constant, instead of cropping a sub-image of fixed size, we crop a ROI whose size is determined by the length of line segment between two valley points.

PolyU II palmprint database is collected by Hong Kong Polytechnic University [35], [69]. This database contains 7752 images of 386 different palms. On average, 20 images from each palm are collected in two sessions, where 10 samples are captured in the first session and the rest in the second session. Fig.8 shows several examples of palmprint images and the corresponding ROI region extracted by pre-processing. The original image size is 352×288 , and the size of ROI is about 190×190 .

The PolyU multispectral database includes four independent spectral palmprint databases [34], [70]. Each spectral database is collected from 250 volunteers including 195 males and 55 females. Each subject provides two palms.

Therefore, there are 500 different palms under a single illumination condition. Each palm includes 12 images taken under the Red, Green, Blue and near-infrared(NIR) illuminations, respectively. Therefore, PolyU database totally contains 6,000 palmprint images for one spectral. Fig.9 depicts palmprint images and their corresponding ROI. The original image size is 384×284 , and the size of ROI is about 160×160 .

CASIA palmprint database is built by Chinese Academy of Science [33], [71]. This database contains 5,502 palmprint images captured from 312 subjects. Both palms of each subject are collected. Fig.10 depicts palmprint images and ROI images. The original image size is 640×480 , and ROI size is about 220×220 .

COEP palmprint database is collected by College of Engineering, Pune [72]. The database consists of 8 different images of single palm, and consists of total 1344 images from 168 individuals. Fig.11 shows several palmprint and their ROI images. The size of palmprint image is 1600×1200 , and ROI size varies from 290×290 to 330×330 .

TongjiU dataset is collected by Tongji University with contactless manner [36], [73]. In this database, images were collected from 300 volunteers with two separate sessions.

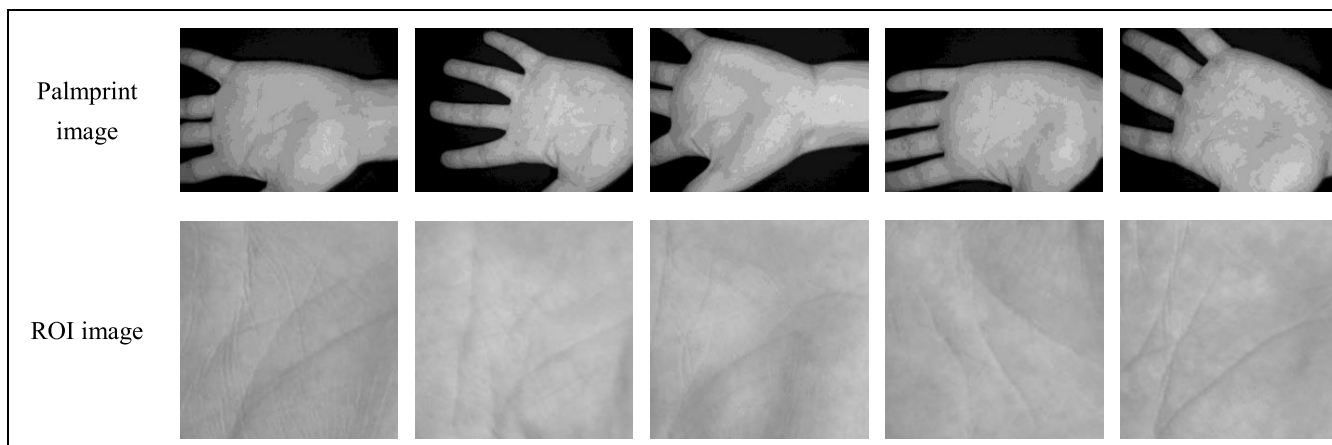


FIGURE 10. Palmprint images and ROI of CASIA database.

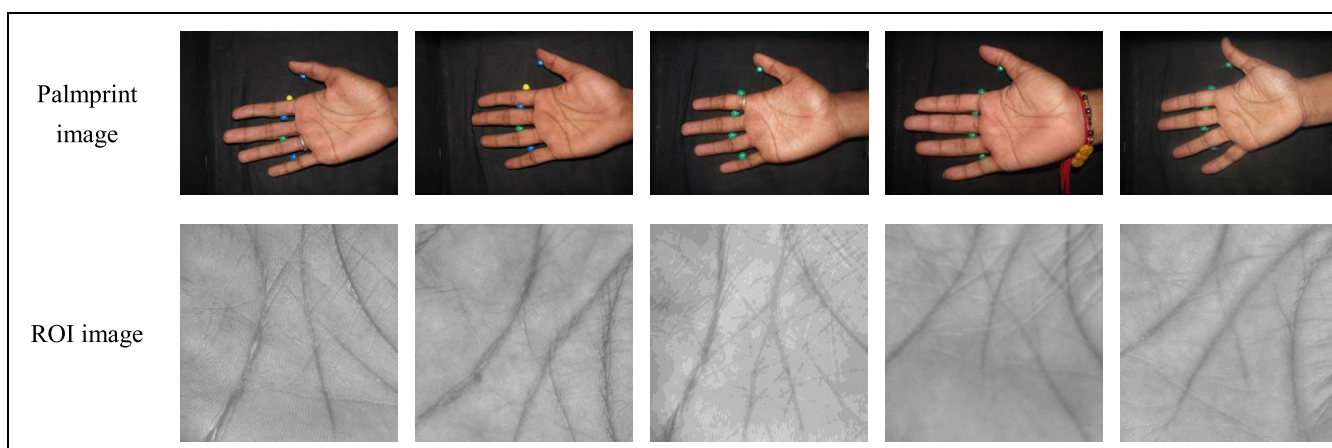


FIGURE 11. Palmprint images and ROI of COEP database.

Each session includes 10 images for each palm. Therefore, 40 images from 2 palms were collected from each subject. In total, the database contains 12,000 images captured from 600 different palms. Due to the free distance between palm and camera, the illumination of every palm is almost uneven, which leads to obvious contrast variations in many images. Fig.12 shows several palmprint and their ROI images. The size of palmprint image is 800×600 , and ROI size varies from 112×112 to 150×150 .

For each database, we randomly select one sample from the one session as a training set, and the remaining samples are used to evaluate the performance. Several state-of-the-art methods are implemented to serve as benchmark algorithms.

All experiments are conducted using MATLAB 2010 on a PC with Intel Core i5-2450M@2.50GHz CPU and RAM of 4GB. The operating system is Windows 10.

B. INVARIANCE ANALYSIS

In practical applications, images captured from the same palm may still differ by rotation, scaling, and translation

caused by different camera settings. Moreover, slight changes in illumination conditions may also affect the quality of images acquired. In order to comprehensively evaluate the invariance of BIT features, we synthesize corrupted palmprint images by rotation, scaling and translation as well as noise interferences. In our experiments, a 2D bilinear interpolation algorithm is employed to obtain the rotated and scaled images. Two stage's transform maps are shown in Fig. 13, in which Fig.13(a) is the original image. For comparison considerations, we show the first and second transform map of all test images. Fig.13(b) (c) are the feature maps after first and second transformation, respectively.

When an input image is rotated by 135° as shown in Fig.13(d), the first stage transform map moves to the right, and the shift angel is 45° on the horizontal axis, as shown in Fig.13(e). However, the second transform map that shown in Fig.13(f) is invariant. During spatial frequency processing, the local spatial frequency detector measures the superimposed edges in all direction, and the value of summation on overlay regions is placed on horizontal axis. Thus, when the edges of the image are rotated, the first stage's map will

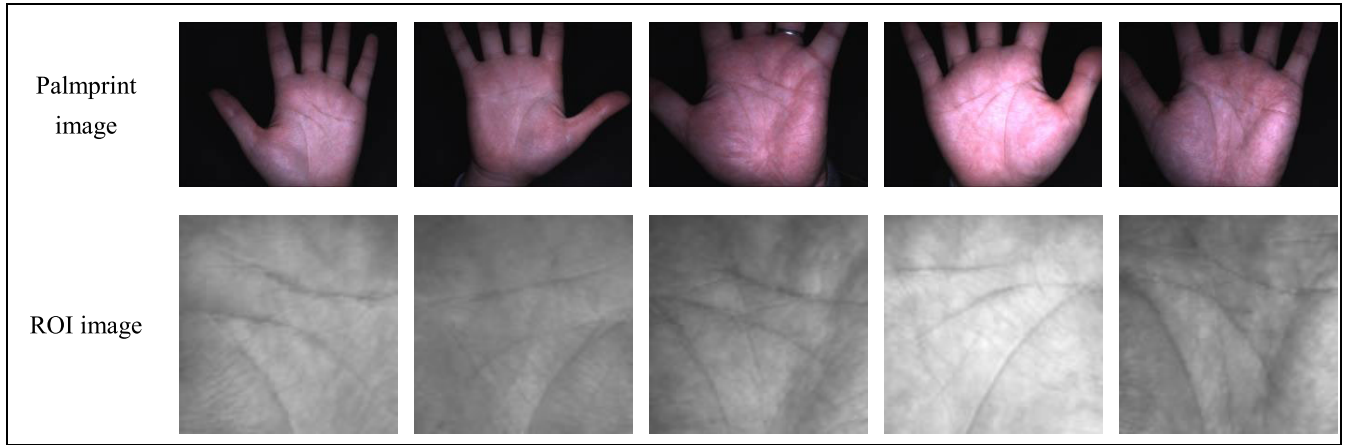


FIGURE 12. Palmprint images and ROI of TongjiU database.

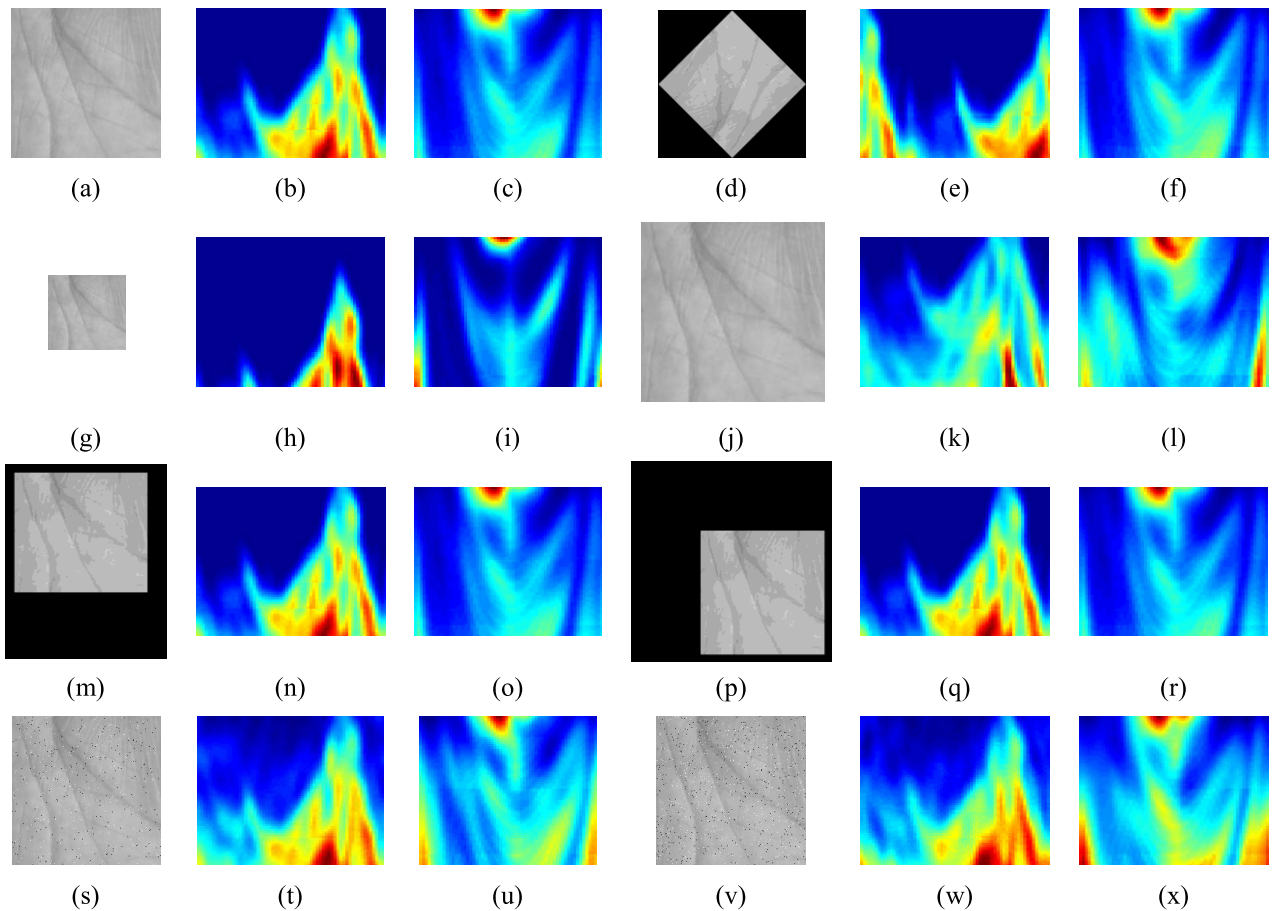


FIGURE 13. Palmprint feature extraction for scaled, rotated, shifted and noised image. (a) is an original image, and (b) (c) are the first and second feature maps respectively. (d) is an image rotated by 135° , and (e) (f) are the first and second feature maps respectively. (g)(j) are two scaled images with the scaling factor set to 0.5 and 1.2, (h)(k) are their first stage's feature map, (i)(l) are the second stage's feature maps. (m) is an image translated by $(-50, -50)$, and (n) (o) are the first and second feature maps respectively. (p) is an image translated by $(50, 50)$, and (q) (r) are the first and second feature maps respectively. (s) and (v) are noisy images, which are generated by adding pepper noise of intensity 0.01 and 0.02, respectively. Their first stage's feature maps are (t) and (w), and (u)(x) are the second feature maps.

shift left or right periodically. This shifted edge map has no effect on local spatial frequency detection in the second stage, which suggests that the proposed method is able to achieve rotation invariance.

Fig.13(g)(j) are two scaled images with the scaling factor set to 0.5 and 1.2, respectively. Fig.13(h)(k) depict their first stage's transform map. Clearly, when the object is scaled down, the first stage transform image is shifted downward.

On the contrary, if the object is scaled up, the first stage transform map is shifted upward. Fig.13(i)(l) are the second stage transform maps which are similar to the previous ones. The reason is that the interval between two shifted edges of an image changes when the image is scaled, and the local spatial frequency detector could detect the overlaid edges at all interval values and the summation of detection is placed on vertical axis. Thus, when the edge of the image is scaled, the first stage transform will shift up or down. In the second stage, there is no scaling in the input image except for a vertical shift. For the shifted image, its edges are shifted simultaneously. Since there is no impact on local spatial frequency detection, the feature map is invariant under scaling condition.

To demonstrate translation invariance, the palmprint image is translated by $(-50, -50)$ to upper left direction (Fig.13(m)), and the image is translated by $(50,50)$ to lower right direction (Fig.13(p)). Fig.13(n) and Fig.13(q) are the first stage map, and Fig.13(o) and Fig.13(r) are the second transform map. Clearly, these transform maps agree with that of the unshifted image. The reason that our method can achieve translation invariance is consistent with the reason for rotation or scaling invariance. The edge contours of shifted image are invariant under translation. In the first stage, the local spatial frequency detector can detect the superimposed edges at all directions and intervals, and the summation of overlay regions is invariant. Thus, the first stage transform map, even though the edge of the image is shifted, will not change. In the second stage, there is no difference between their input images. Therefore, their edges are identical as well as the local spatial frequency detection, which leads to translation invariance.

Fig.13(s) and Fig.13(v) are noisy images, which are generated by adding pepper noise of intensity 0.01 and 0.02, respectively. Their first stage's transform maps are shown in Fig.13(t) and Fig.13(w). Obviously, higher noise intensity leads to heavier interference in the first transform map. Consequently, these interferences will be brought into the second stage's transform map as shown in Fig.13(u) and Fig.13(x). Comparing with Fig.13(c), we can see that some regions in the feature map of noisy image are enhanced. However, the contours of high level regions are highly similar, which demonstrates our method is able to tolerate noise to some degree.

From the above experiments, the contours of feature map usually remain invariant when image is rotated, scaled and translated as well as corrupted by noise. Similarly, the proposed approach achieves the same invariance on other palmprint images, which implies that the proposed BIT feature is RST-invariant.

C. PALMPRINT VERIFICATION

Palmprint verification is a one-to-one comparison of a new palmprint with a specific template stored in the database to verify the individual is the person they claim to be. In this stage, each palm image is matched with all other palm images

in the same database. A match is genuine if the matched palm sample belongs to the same subject, otherwise it is considered as impostor. To prevent impostor palmprint feature (in this case all palmprint features of persons not known by the system) from being regarded as valid, the matching score must exceed a certain level, otherwise the palm is rejected.

For PolyU II palm database, it has 7,752 samples from 386 palms, but there are not always 10 images for each palm. For the purpose of evaluating performance, we select 3,600 images from 360 palms as the test database. In this way, there are $3600 \times 3599.2 = 6,478,200$ matches in total, where 16,200 matching are genuine, and 6,462,000 matching are impostor. In PolyU multispectral database, one spectral contains 6,000 palmprint images, so there are 17,997,000 matches. Each session contains 12 images, so there are 66 genuine. Totally, there are 33,000 genuine matches, and 17,964,000 impostor matches. The CASIA database contains 312 subjects. However, the number with the same type of palm is also not even. We set up a database by choosing 4,800 images. These images are captured from 300 individuals including 8 pairs of palm images. Thus, there are 11,517,600 matches in total, with 16,800 matches being genuine and 11,500,800 matches being impostor. Similarly, COEP database has 902,496 matches including 47,264 genuine matches and 85,5232 impostor matches respectively. All images in TongjiU dataset produce 71,994,000 matches, and there are 114,000 genuine matches and 71,880,000 impostor matches.

To measure the verification performance, we built a training and test set for all databases. In each database, the first three palmprint images from each session are used for training. The rest in each session and the palmprints from other sessions are used for test. Therefore, the numbers of images for training and test are 1,158 and 6,594 respectively in PolyU II database. In PolyU multispectral database, each band includes 1,500 training images and 4,500 test images. There are 936 training images and 4,566 test images in CASIA database. COEP database involves 504 training images and 840 test images respectively. TongjiU database has 3,600 training images and 8,400 test images. Experiments are conducted on each database respectively. After all test images matched with all training images, the statistical values of false accept rate (FAR), genuine acceptance rate (GAR) and equal error rate (EER) are calculated. In our experiments, the receiver operating characteristics (ROC) curve is employed to measure the variation of FAR with GAR. Fig.14 shows the ROC curve of the proposed approach and other methods including DRCC [37], ALDC [39], CompCode [28], LLDP [14], LTMrP [44], CR_CompCode [36], VGG-16 [56], AlexNet [56] and PalmNet [55]. VGG-16 [56] and AlexNet [56] are pre-trained CNN models that will extract palmprint feature. From these ROC curves, it is not difficult to see that our approach can usually achieve high GAR at the same FAR level.

The EERs obtained from different methods are summarized in Table 1. The EERs of BIT are apparently lower

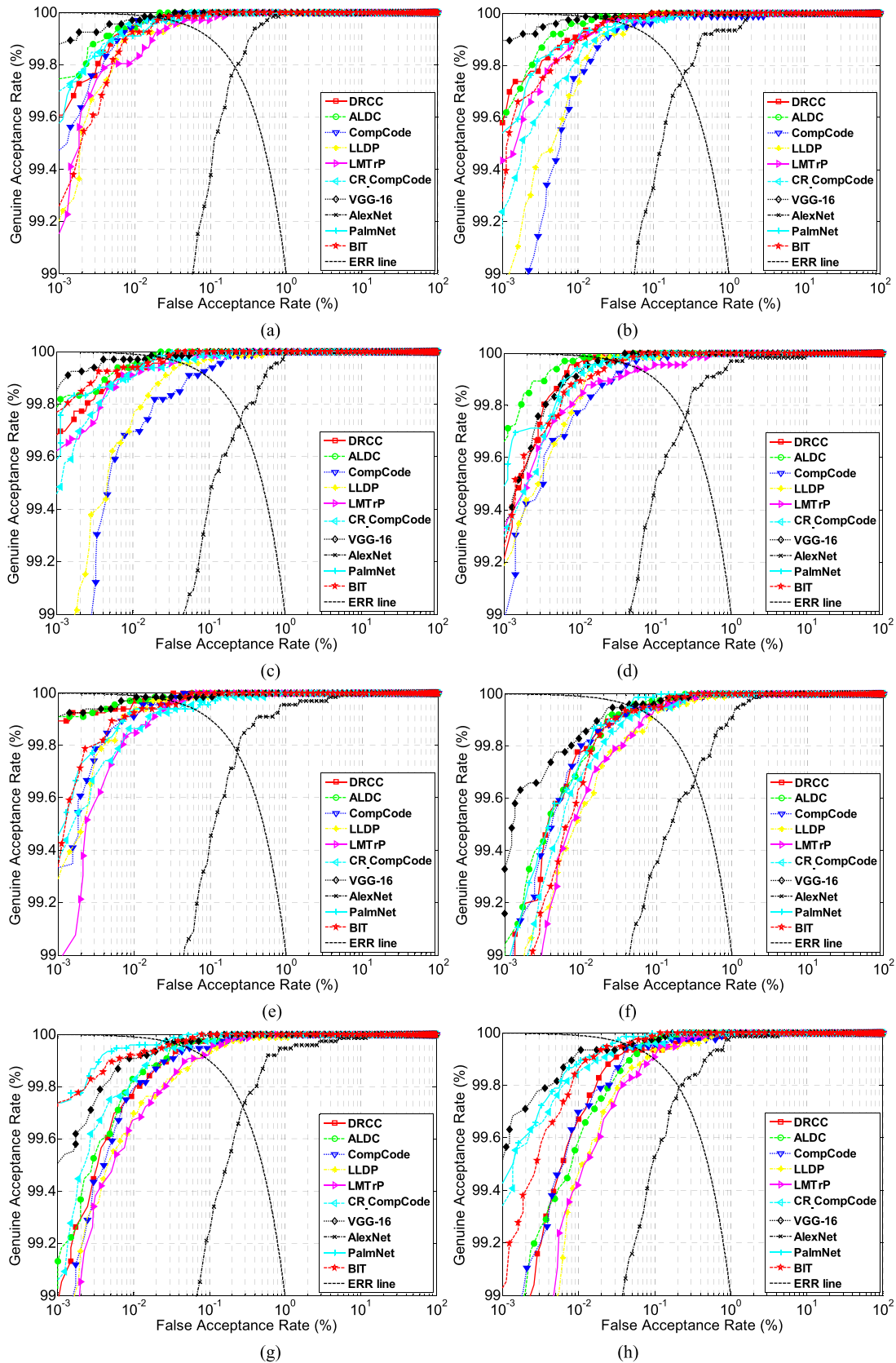


FIGURE 14. ROC curves of different methods on normal images of each database. (a) Red, (b) Green, (c) Blue, (d) NIR spectral, (e) PolyU II, (f) CASIA, (g) COEP and (h) TongjiU database.

TABLE 1. Verification EER(%) on each database with different methods.

EERs (%)	DRCC[37]	ALDC[39]	CompCode [28]	LLDP[14]	LMTrP [44]	CR_Comp Code[36]	VGG-16 [56]	AlexNet [56]	PalmNet [55]	BIT
PolyU Red spectral	0.0227	0.0154	0.0290	0.0430	0.0408	0.0303	0.0152	0.2116	0.0304	0.0265
	0.0565	0.0409	0.1061	0.0956	0.0811	0.0909	0.0606	2.0794	0.1212	0.0455
PolyU Green spectral	0.0263	0.0164	0.0507	0.0484	0.0263	0.0424	0.0132	0.2454	0.0367	0.0263
	0.0969	0.0909	0.1316	0.1757	0.1049	0.1061	0.0517	2.4909	0.1373	0.0407
PolyU Blue spectral	0.0281	0.0172	0.0777	0.0455	0.0388	0.0322	0.0154	0.2616	0.0204	0.0274
	0.0612	0.0457	0.1061	0.1578	0.0645	0.0723	0.0658	2.6087	0.1599	0.0427
PolyU NIR spectral	0.0160	0.0152	0.0424	0.0190	0.0606	0.0328	0.0295	0.2488	0.0250	0.0303
	0.0962	0.0766	0.1404	0.1102	0.0744	0.0694	0.0606	2.6405	0.1721	0.0470
PolyU II	0.0282	0.0171	0.0326	0.0303	0.0303	0.0455	0.0152	0.2263	0.0606	0.0381
	0.0931	0.0846	0.1667	0.1917	0.1045	0.1216	0.0934	2.5254	0.1349	0.0455
CASIA	0.0561	0.0500	0.0536	0.0861	0.0789	0.0658	0.0429	0.3275	0.0453	0.0623
	0.1507	0.1214	0.2296	0.2891	0.0993	0.0822	0.0710	2.4941	0.1298	0.0644
COEP	0.0449	0.0388	0.0526	0.0801	0.0828	0.0462	0.0263	0.2774	0.0263	0.0355
	0.1818	0.0979	0.1515	0.1818	0.0706	0.1042	0.0758	2.4803	0.1597	0.0480
TongjiU	0.0658	0.0619	0.0526	0.0842	0.0956	0.0526	0.0469	0.2068	0.0321	0.0450
	0.1345	0.0909	0.2358	0.1485	0.1319	0.0944	0.0152	1.5931	0.1965	0.0758

than CompCode [28], CR_CompCode [36] an AlexNet [56] on PolyU multispectral database, and even lower than PalmNet [55] on Red and Green spectral databases. With regard to PolyU II database, the EERs of BIT is 0.0381%, which is superior than CR_CompCode [36], AlexNet [56] and PalmNet [55]. In COEP and TongjiU database, BIT has better verification performance than most of methods. Specifically, the EER of BIT approach is 0.0355% on COEP database, which is lower than most of EERs among the listed methods except for VGG-16 [56], and PalmNet [55]. In these methods PalmNet [55] and VGG-16 [56] achieve the lowest EER that are both 0.0263%. DRCC [37] is 0.0449%, and ALDC [39] is 0.0388%. The EER of AlexNet [56] is 0.2774%, which is the highest. Also, The EER of BIT is 0.0450% on TongjiU database. Although the EER of BIT is higher than PalmNet [55], BIT outperforms the existing methods, i.e. VGG-16[56] achieves 0.0469%. The above experiments reveal that BIT can achieve relatively low EERs in less constraint palmprint database.

To further evaluate the robustness of verification on corrupted images, some experiments are conducted on noisy ROI images. We randomly corrupt test images with one of the following variations: noise intensity: 0-1, rotate angle: 0-360°, scaling factor: 0.1-2, and translation range: 0-20 pixels. The size of training and test dataset is the same as normal image experiments. ROC curves of all methods on Red, Green, Blue, NIR spectral, PolyU II, CASIA, COEP and TongjiU database are presented in Fig.15. According to these pictures, the EER of BIT is obviously lower than others methods for all GAR and FAR points in most of databases.

Compared with EERs on normal images, EERs on corrupted images are slightly higher, but the gap between BIT approach and other methods increases, suggesting that our method is robust to corruptions including rotation, scaling, translation and noise. More detailed verification EERs of each method on all databases are also presented with grey color in Table 1. For PolyU II database, the EER of BIT is 0.0455%, which is lower than the lowest EER among other methods, i.e. 0.0846% for ALDC [39]. BIT achieve higher EER than ALDC [39] on Red spectral database. The EER of BIT is 0.0455%, and ALDC [39] is 0.0409%. For Green, Blue and NIR spectral database, BIT achieves the lowest EERs that are 0.0407%, 0.0427% and 0.0470%, and it is not difficult to see that these results are still the lowest EER among these methods. The EER of BIT is 0.0644% on CASIA database, which is much better than other methods, i.e. the lowest EER of VGG-16 [56] is 0.0710%. For COEP database, the EER is similar to that of PolyU multi-spectral and PolyU II databases. In terms of TongjiU database, the EER of BIT is 0.0758%, which is higher than other databases, the reason is that this contactless database has many more corruptions or noises. Nevertheless, it is lowest when comparing with other methods.

D. PALMPRINT IDENTIFICATION

Palmprint identification is a one-to-many comparison against a palmprint database. For each person, we randomly select an image as template in training stage. Then, the biometric template is calculated by DRCC [37], ALDC [39], CompCode [28], LLDP [14], LMTrP [44], CR_CompCode[36],

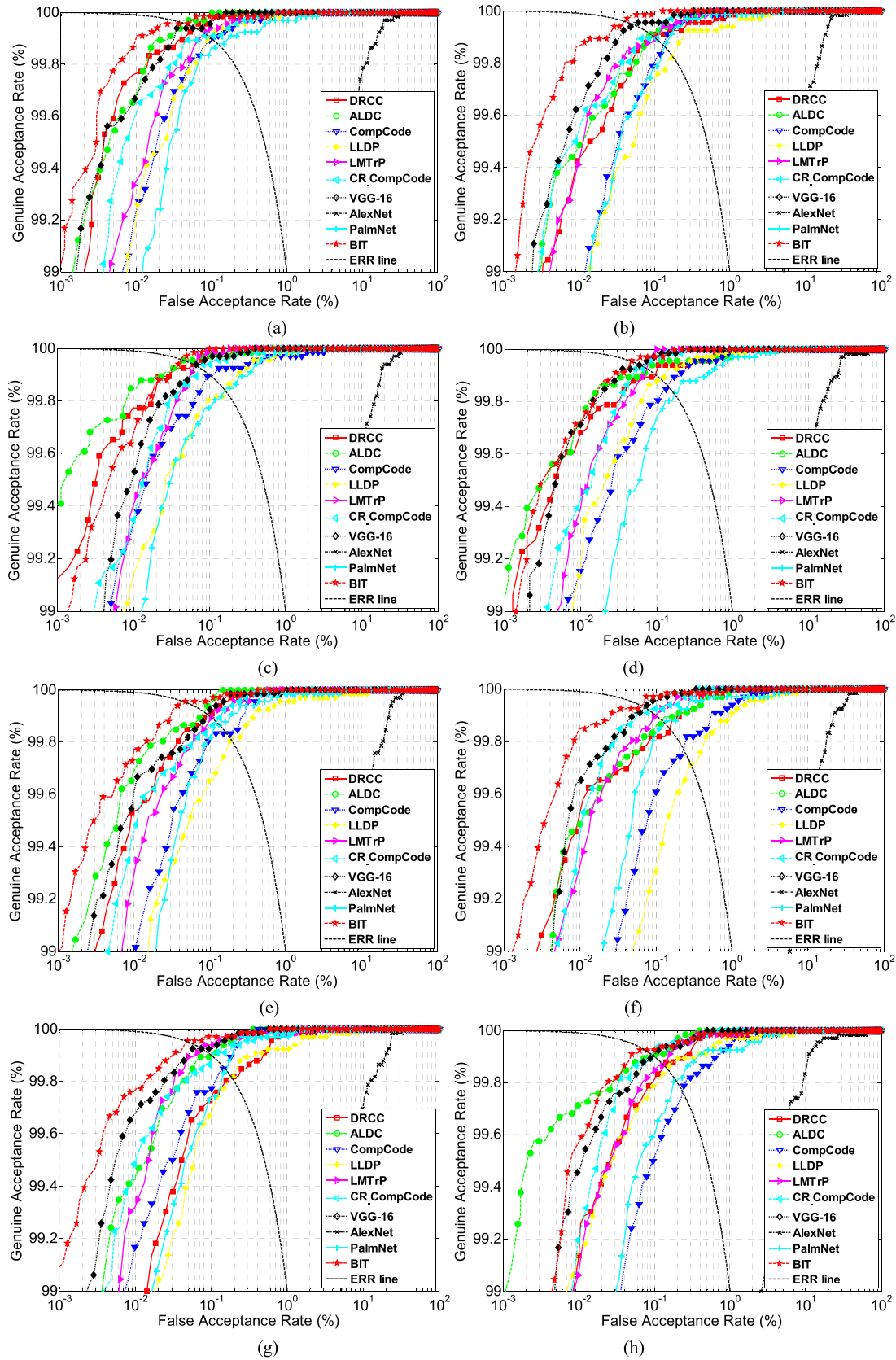


FIGURE 15. ROC curves of different methods on corrupted images of each database. (a) Red, (b) Green, (c) Blue, (d) NIR spectral, (e) PolyU II, (f) CASIA, (g) COEP and (h) TongjiU database.

TABLE 2. The rank-1 recognition accuracy(%) of various methods on corrupted PolyU II database.

Rank-1(%)	Red	Green	Blue	NIR	PolyU II	CASIA	COEP	TongjiU
DRCC[37]	96.51	95.71	97.68	96.82	96.11	96.36	95.34	94.12
ALDC[39]	97.15	96.71	98.32	96.93	97.25	96.57	95.64	94.37
CompCode[28]	92.31	91.27	93.00	92.34	90.87	91.35	91.52	89.95
LLDP[14]	91.59	90.22	90.97	90.34	89.16	87.15	91.27	87.31
LMTTrP[44]	91.23	91.41	90.65	91.33	88.54	88.61	91.65	87.69
CR_CompCode[36]	91.24	92.32	92.54	90.32	91.21	90.35	92.67	91.85
VGG-16[56]	92.65	91.39	92.21	89.14	92.38	91.02	94.27	90.41
AlexNet[56]	80.32	87.64	79.48	93.12	87.74	90.33	91.31	89.87
PalmNet[55]	95.67	96.38	94.69	95.93	95.02	94.55	96.89	95.75
BIT	96.11	95.67	96.89	96.73	95.88	94.91	95.86	94.69

VGG-16 [56], AlexNet [56] and PalmNet [55] and BIT. A palmprint feature that is going to be identified is matched against every known template, yielding a distance describing the similarity between the feature and the template. We assign the pattern to the identity of the individual with the most similar biometric template. Rank-1 recognition accuracy is used on all databases in this experiment.

Table 2 lists the rank-1 recognition accuracy of various methods on each database. Although the recognition accuracy of BIT is lower than DRCC [37], ALDC [39] and PalmNet [55] in some database, BIT achieves higher recognition accuracy than other methods. More importantly, BIT is superior to DRCC [37] and ALDC [39] on COEP and TongjiU database. Specifically, the recognition accuracy of BIT is 95.88% on PolyU II database, which is slightly lower than DRCC [37] and ALDC [39], but it is obviously higher than all other methods. The recognition accuracy of BIT is 96.11% for Red spectral, 95.67% for Green spectral, 96.89% for Blue spectral, 96.73% for NIR spectral, and 94.91% for CASIA database, showing that the recognition accuracy of BIT is slightly higher than most of methods. Also, the recognition accuracies of BIT on these databases are close to DRCC[37], ALDC [39], and always outperforms PalmNet [55]. For COEP and TongjiU database, BIT achieves 95.86% and 94.69% recognition accuracy respectively. These recognition accuracies on before mentioned databases are higher than ALDC [39] that are 95.64% and 94.37%, and are close to the best results given by PalmNet [55] that are 96.89% and 95.75% respectively.

To further verify the robustness of BIT in identification tasks, we artificially corrupt palmprint images on PolyU II database by adding various level of rotation, scaling, translation and noise. Here, we use one image from the first session for training, and all images from the second session for test. During the experiment, the rank-1 recognition accuracy is adopted to measure identification accuracy at all interference points in sequence.

More details about the recognition accuracy are shown in Fig.16. Fig.16(a) shows the relationship between the degree of rotation and recognition accuracy. As the degree of rotation rises from 0 to 90°, the recognition accuracy of DRCC [37], ALDC [39], CompCode [28], LLDP [14], LMTTrP [44], CR_CompCode [36] decreases dramatically. This observation tells that these methods are sensitive to rotations of input images thus damage the robustness of recognition. The recognition accuracies of deep learning based methods such as VGG-16 [56], AlexNet [56] and PalmNet [55], though moderately, drop down with image rotation. However, without much surprise, rotation almost has no impact on recognition accuracy when the features are extracted by BIT.

Fig.16(b) shows how recognition accuracies changes when the scaling factor goes from 0.1 to 2. From this figure, we can see that the recognition accuracies of the state-of-the-art methods are lower than that of BIT at most scaling points. Note that unlike rotation, scaling does clearly affect the performance of BIT in recognition tasks due to limited interval ranges, yet BIT still outperforms other methods. Nevertheless, the performance of BIT is stable and satisfying with small scaling factors from about 0.8 to 1.1, and decreases significantly outside this range.

Fig.16(c) shows the impact of translation on palmprint recognition accuracy. Clearly, BIT is not affected by translation at all, achieving an accuracy of almost 100% consistently, while the performance of other methods drops drastically when the translation becomes larger. It must point out that VGG-16[56], AlexNet[56] and PalmNet[55] maintain relatively high recognition accuracies when image translation occurs. The high recognition accuracy of deep learning methods may result from pooling or optimal filter selection operations.

In the experiment examining the influence of noise, we increase the noise intensity from 0 to 1 with a step size of 0.1. It is expected that blurred image features will

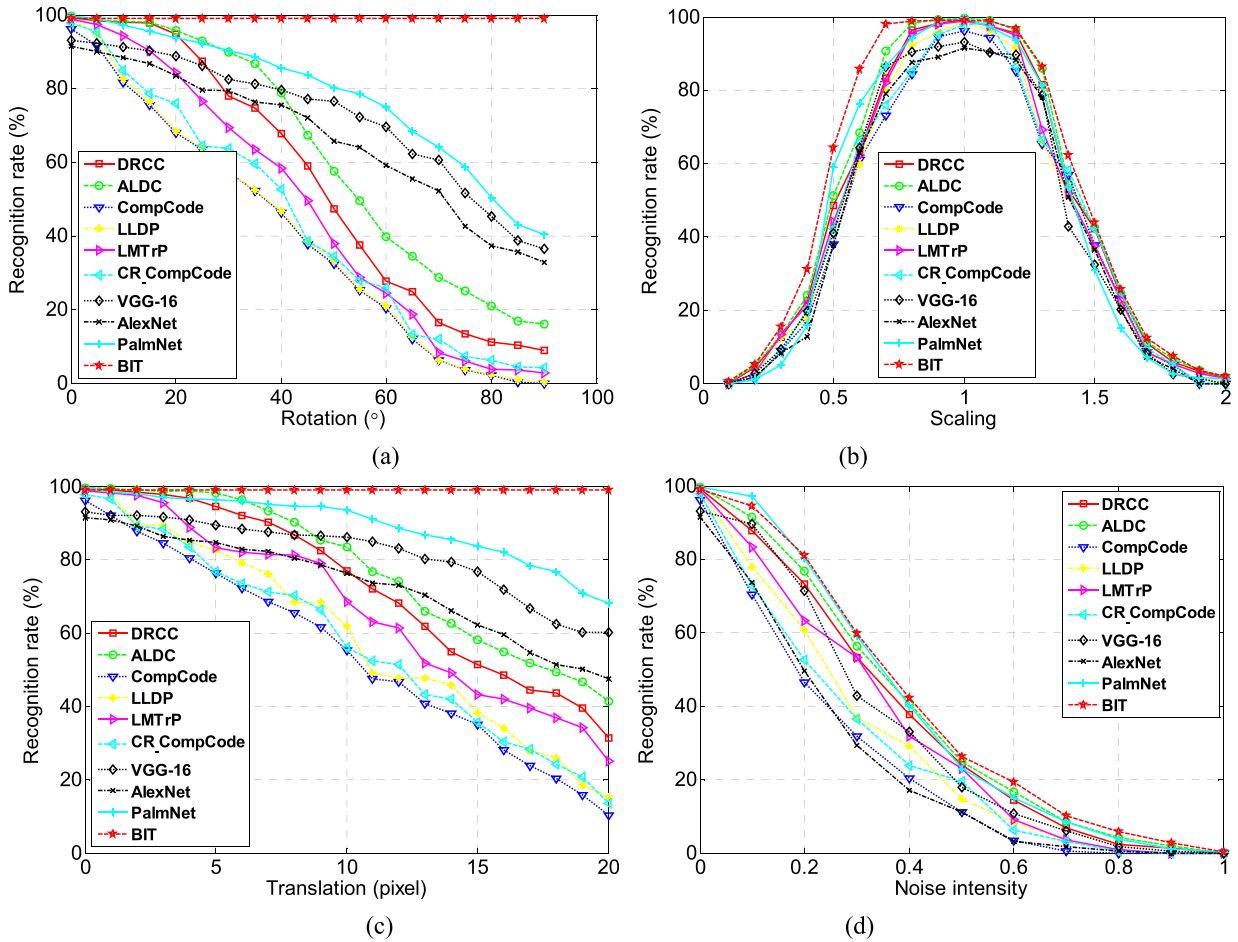


FIGURE 16. The rank-1 recognition accuracy(%) under rotation, scaling, translation, and noise on PolyU II database. (a) Rotation, (b) scaling, (c) translation, and (d) noise.

be generated from the images corrupted by random noise. Fig.16(d) describes the recognition accuracy with different noise intensities. From this picture, the recognition accuracies of all methods decline when increasing noise intensity. However, BIT remains higher recognition accuracy than other methods consistently.

The results of this experiment have demonstrated that the invariant feature extracted by BIT has better discriminability than other methods.

E. THE IMPACT OF DIFFERENT EDGE DETECTOR AND MATCHING ALGORITHM

Since edge detections have significantly impact on recognition accuracy, we compared the proposed orientation edge detector with the existing algorithm. Basically, there are some edge detectors for palmprint recognition such as Canny detector, Sobel detector and phase symmetry algorithm. We perform these edge detectors followed by local spatial frequency detector for two stage’s transform. Palmprint image enhancement operations are also performed. The comparison experiments are conducted by five times on PolyUII database, and the averaged recognition accuracies are dspicted in Fig.17.

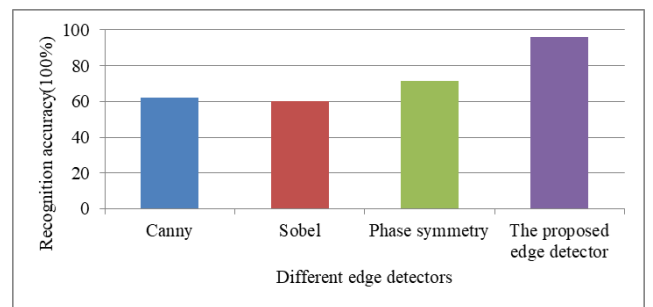


FIGURE 17. Recognition accuracy with different edge detector.

From Fig. 17, the recognition accuracies of Canny and Sobel algorithms are low, which are 62.37% and 60.19% respectively. Phase symmetry algorithm achieved higher recognition accuracy than Canny and Sobel algorithm. The reason is that phase symmetry algorithm can capture many more palmprint details. The proposed edge detector achieved the highest recognition accuracy among these algorithms that is 95.88%. The results imply that the direction of palmprint edge is quite important for feature extraction.

For matching algorithms, some typical classification methods including SVM [54], CRC [57], Chi-square distance [39] and KNN [55] are popular. To compare with the performance of these algorithms, we replace KNN [55] with the above mentioned algorithms for palmprint matching after extracting palmprint feature with the proposed framework. The averaged recognition accuracies are illustrated in Fig. 18. From this picture, it is not difficult to find that SVM [54] and CRC [57] achieved similar recognition accuracies that are 97.18% and 97.64% respectively. Due to absolute distance measuring, the recognition accuracy of Chi-square distance [39] is 93.65%, which is obviously lower than that of SVM [54], CRC [57] and KNN [55] with correlation distance algorithms. The reason may lie in the facts that SVM [54] and CRC [57] have better generalization than distance based matching algorithms. Nevertheless, KNN with correlation distance algorithm is able to achieve comparable recognition accuracy without data training and parameters tuning.

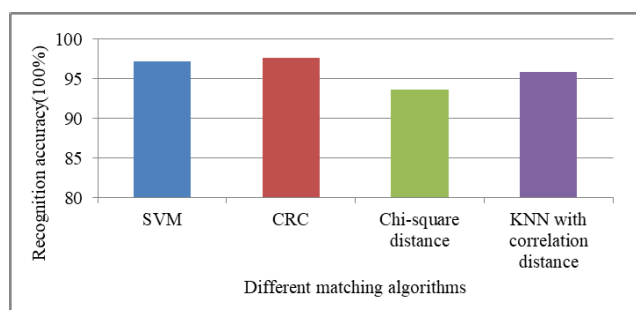


FIGURE 18. Recognition accuracy with different matching algorithms.

F. ALGORITHM COMPLEXITY

To further evaluate the algorithm complexity, we compare the computation cost of the proposed approach with the state-of-the-art methods including DRCC [37], ALDC [39], CompCode [28], LLDP [14], LTMrP [44] and CR_CompCode [36], PalmNet, VGG-16 [56], AlexNet [56] and PalmNet [55]. Suppose that the size of a palmprint image is $m \times n$, $m, n \in \mathbb{Z}$, the computational complexity of the proposed approach on edge features, orientation features and spatial frequency features extraction can be written as $O(d)$, where $d = m \times n$. For a fair comparison, all images are normalized to 128×128 . Computation costs averaged on several trials are reported.

The computation costs of each method are listed in Table 3. It can be seen that the lowest computation time is achieved by CompCode [28] with 0.011s. DRCC [37] and CR_CompCode [36] are 0.044s and 0.016s respectively. Since there are many convolution operations in BIT, the cost of feature extraction is highest among the listed methods.

Feature matching is quite fast after the features are already extracted. The time of feature matching of CompCode [28] is 0.093ms, which is the shortest among these approaches. The time of feature matching for BIT is 0.812ms, which is higher than DRCC [37], ALDC [39], CompCode[28],

TABLE 3. Computation cost comparison of ten different methods.

Methods	Feature	Feature	Feature size (Bytes)
	extraction time(s)	matching time(ms)	
DRCC[37]	0.044	0.328	278,528
ALDC[39]	0.086	0.247	28,152
CompCode[28]	0.011	0.093	6,144
LLDP[14]	0.236	1.204	65,024
LTMrP[44]	0.279	1.136	19,200
CR_CompCode[36]	0.016	25.717	3,888
VGG-16[56]	0.125	0.536	8,000
AlexNet[56]	0.107	0.629	8,000
PalmNet[55]	2.417	12.253	1,222,576
BIT	2.894	0.812	32,768

VGG-16 [56] and AlexNet [56]. The most important reason is that Hamming distance remains high efficiency when matching binary palmprint feature. Since SVM [54] is employed to classify the VGG-16 and AlexNet feature vector. Thus, the matching times of VGG-16 [55] and AlexNet [55] are also faster than BIT. However, matching time of BIT is less than LLDP [14], LTMrP [39] and CR_CompCode [36]. It should be noted that the time cost in feature extraction stage of BIT could be significantly reduced by employing specific optimization techniques (e.g., C/GPU-based) to accelerate convolution operations.

The fourth column in Table 3 summaries the feature size of each compared methods. The size of BIT feature is 64×64 . Thus, the template size of BIT feature to be stored is 32,768 Bytes, which is the larger than VGG-16 [56] and AlexNet [55]. The feature size of CR_CompCode[36] is 3,888 Bytes, which is the lowest among those methods. CompCode [28] has feature size of 128×128 . Since each element in feature map is coded with three bits, the total feature size is 6,144 Bytes, which is obviously higher than CR_CompCode [36]. The highest of feature size occurs in PalmNet [55] that is 1,222,576 Bytes. Notably, many data types of storage feature are double precision, which means that one element of feature vector has 64 bit. This will lead to a little high memory cost. Although the feature size of our approach is not the least among these methods, the feature size is independent of input image size.

VI. CONCLUSIONS AND FUTURE WORKS

In this paper, we proposed a robust BIT feature extractor for palmprint recognition. Different from the existing works, our framework has no parameters learned through back propagation or gradient descent, and it uses orientation edge detector and local spatial detector. Specifically, we handled illumination variation through a palmprint enhancement algorithm,

and built orientation edge detector and local spatial frequency detector to measure edge response on all orientations and spatial frequencies, which can simulate the visual perception mechanism of simple cell. Also, BIT feature extraction process could be divided into two stages. The translation problem can be solved in the first stage, and the rotation and scaling problems are handled in the second stage. An invariant feature is extracted to meet the need of modern contactless palmprint recognition applications. Compared with some existing methods, BIT is more robust in handling degraded images as it achieves higher recognition accuracy.

Future work will be dedicated to resolving the following issues: (1) High complexity of BIT. A number of convolution operations are involved in computing BIT features, which inevitable lead to high computational cost. (2) Limited applicable scenarios. In this paper we have shown that BIT is rotation and translation invariant, yet it can not be directly applied to palmprint recognition when there are affine distortions. Thus, extending BIT to be affine-invariant will surely enlarge its applicability.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

ACKNOWLEDGMENT

The authors would like to thank Hong Kong Polytechnic University (PolyU), Chinese Academy of Sciences' Institute of Automation (CASIA), College of Engineering, Pune-411005 (COEP), and Tongji University (TongjiU) for providing palmprint database. They would also appreciate anonymous reviewers for their helpful comments.

REFERENCES

- A. K. Jain, K. Nandakumar, and A. Ross, "50 years of biometric research: Accomplishments, challenges, and opportunities," *Pattern Recognit. Lett.*, vol. 79, pp. 80–105, Aug. 2016.
- D. Hong, W. Liu, X. Wu, Z. Pan, and J. Su, "Robust palmprint recognition based on the fast variation Vese–Osher model," *Neurocomputing*, vol. 174, pp. 999–1012, Jan. 2016.
- L. Fei, Y. Xu, and D. Zhang, "Half-orientation extraction of palmprint features," *Pattern Recognit. Lett.*, vol. 69, pp. 35–41, Jan. 2016.
- D. Zhong, X. Du, and K. Zhong, "Decade progress of palmprint recognition: A brief survey," *Neurocomputing*, vol. 328, pp. 16–28, Feb. 2019.
- W. Li, B. Zhang, L. Zhang, and J. Yan, "Principal line-based alignment refinement for palmprint recognition," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1491–1499, Nov. 2012.
- 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.
- A. Kumar, "Toward more accurate matching of contactless palmprint images under less constrained environments," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 1, pp. 34–47, Jan. 2019.
- M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," *Nature Neurosci.*, vol. 2, no. 11, pp. 1019–1025, Nov. 1999.
- D. H. Hubel, "Exploration of the primary visual cortex, 1955–78," *Nature*, vol. 299, no. 5883, pp. 515–524, Oct. 1982.
- L. Robinson and E. T. Rolls, "Invariant visual object recognition: Biologically plausible approaches," *Biol. Cybern.*, vol. 109, nos. 4–5, pp. 505–535, Oct. 2015.
- K. Zhou, X. Zhou, L. Yu, L. Shen, and S. Yu, "Double biologically inspired transform network for robust palmprint recognition," *Neurocomputing*, vol. 337, pp. 24–45, Apr. 2019.
- Y. Xu, L. Fei, and D. Zhang, "Combining left and right palmprint images for more accurate personal identification," *IEEE Trans. Image Process.*, vol. 24, no. 2, pp. 549–559, Feb. 2015.
- W. Wu, S. J. Elliott, S. Lin, S. Sun, and Y. Tang, "Review of palm vein recognition," *IET Biometrics*, vol. 9, no. 1, pp. 1–10, Jan. 2020.
- Y.-T. Luo, L.-Y. Zhao, B. Zhang, W. Jia, F. Xue, J.-T. Lu, Y.-H. Zhu, and B.-Q. Xu, "Local line directional pattern for palmprint recognition," *Pattern Recognit.*, vol. 50, pp. 26–44, Feb. 2016.
- L. Fei, Y. Xu, W. Tang, and D. Zhang, "Double-orientation code and nonlinear matching scheme for palmprint recognition," *Pattern Recognit.*, vol. 49, pp. 89–101, Jan. 2016.
- A. Kong, D. Zhang, and M. Kamel, "A survey of palmprint recognition," *Pattern Recognit.*, vol. 42, no. 7, pp. 1408–1418, Jul. 2009.
- J. Dai and J. Zhou, "Multifeature-based high-resolution palmprint recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 945–957, May 2011.
- A. K. Jain and J. Feng, "Latent palmprint matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 6, pp. 1032–1047, Jun. 2009.
- R. Cappelli, M. Ferrara, and D. Maio, "A fast and accurate palmprint recognition system based on minutiae," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 3, pp. 956–962, Jun. 2012.
- E. Liu, A. K. Jain, and J. Tian, "A coarse to fine minutiae-based latent palmprint matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 10, pp. 2307–2322, Oct. 2013.
- J. Dai, J. Feng, and J. Zhou, "Robust and efficient ridge-based palmprint matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1618–1632, Aug. 2012.
- A. Iula and D. Nardiello, "3-D ultrasound palmprint recognition system based on principal lines extracted at several under skin depths," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 12, pp. 4653–4662, Dec. 2019.
- A. Genovese, V. Piuri, and F. Scotti, "Touchless palmprint recognition systems," in *Advances in Information Security*. Cham, Switzerland: Springer, 2014, pp. 49–109.
- L. Zhang, Y. Shen, H. Li, and J. Lu, "3D palmprint identification using block-wise features and collaborative representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 8, pp. 1730–1736, Aug. 2015.
- X. Wu, D. Zhang, K. Wang, and B. Huang, "Palmprint classification using principal lines," *Pattern Recognit.*, vol. 37, no. 10, pp. 1987–1998, Oct. 2004.
- D. Palma, P. L. Montessoro, G. Giordano, and F. Blanchini, "Biometric palmprint verification: A dynamical system approach," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 12, pp. 2676–2687, Dec. 2019.
- M. D. Bounneche, L. Boubchir, A. Bouridane, B. Nekhou, and A. Ali-Chérif, "Multi-spectral palmprint recognition based on oriented multiscale log-Gabor filters," *Neurocomputing*, vol. 205, pp. 274–286, Sep. 2016.
- A. W.-K. Kong and D. Zhang, "Competitive coding scheme for palmprint verification," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, Cambridge, U.K., 2004, pp. 520–523.
- W. Zuo, F. Yue, K. Wang, and D. Zhang, "Multiscale competitive code for efficient palmprint recognition," in *Proc. 19th Int. Conf. Pattern Recognit.*, Tampa, FL, USA, Dec. 2008, pp. 1–4, doi: [10.1109/ICPR.2008.4761868](https://doi.org/10.1109/ICPR.2008.4761868).
- Z. Guo, W. Zuo, L. Zhang, and D. Zhang, "A unified distance measurement for orientation coding in palmprint verification," *Neurocomputing*, vol. 73, nos. 4–6, pp. 944–950, Jan. 2010.
- J. You, W.-K. Kong, D. Zhang, and K. H. Cheung, "On hierarchical palmprint coding with multiple features for personal identification in large databases," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 2, pp. 234–243, Feb. 2004.
- W. Jia, D.-S. Huang, and D. Zhang, "Palmprint verification based on robust line orientation code," *Pattern Recognit.*, vol. 41, no. 5, pp. 1504–1513, May 2008.
- Z. Sun, T. Tan, Y. Wang, and S. Z. Li, "Ordinal palmprint representation for personal identification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, San Diego, CA, USA, Jun. 2005, pp. 279–284.
- A. Kong, D. Zhang, and M. Kamel, "Palmprint identification using feature-level fusion," *Pattern Recognit.*, vol. 39, no. 3, pp. 478–487, Mar. 2006.
- D. Zhang, W.-K. Kong, J. You, and M. Wong, "Online palmprint identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1041–1050, Sep. 2003.
- 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.

- [37] 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.
- [38] W. Jia, B. Zhang, J. Lu, Y. Zhu, Y. Zhao, W. Zuo, and H. Ling, "Palmprint recognition based on complete direction representation," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4483–4498, Sep. 2017.
- [39] 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.
- [40] R. Raghavendra and C. Busch, "Texture based features for robust palmprint recognition: A comparative study," *EURASIP J. Inf. Secur.*, vol. 2015, no. 1, pp. 1–9, Dec. 2015.
- [41] Q. Zhu, N. Xu, Z. Zhang, D. Guan, R. Wang, and D. Zhang, "Cross-spectral palmprint recognition with low-rank canonical correlation analysis," *Multimedia Tools Appl.*, pp. 1–22, Dec. 2019.
- [42] I. Rida, R. Herault, G. L. Marcialis, and G. Gasso, "Palmprint recognition with an efficient data driven ensemble classifier," *Pattern Recognit. Lett.*, vol. 126, pp. 21–30, Sep. 2019.
- [43] D. Tamrakar and P. Khanna, "Kernel discriminant analysis of block-wise Gaussian derivative phase pattern histogram for palmprint recognition," *J. Vis. Commun. Image Represent.*, vol. 40, pp. 432–448, Oct. 2016.
- [44] G. Li and J. Kim, "Palmprint recognition with local micro-structure tetra pattern," *Pattern Recognit.*, vol. 61, pp. 29–46, Jan. 2017.
- [45] I. Rida, S. Al-Maadeed, A. Mahmood, A. Bouridane, and S. Bakshi, "Palmprint identification using an ensemble of sparse representations," *IEEE Access*, vol. 6, pp. 3241–3248, Feb. 2018.
- [46] J. Almaghtuf and F. Khelifi, "Self-geometric relationship filter for efficient SIFT key-points matching in full and partial palmprint recognition," *IET Biometrics*, vol. 7, no. 4, pp. 296–304, Jul. 2018.
- [47] N. V. Graham, "Beyond multiple pattern analyzers modeled as linear filters (as classical v1 simple cells): Useful additions of the last 25 years," *Vis. Res.*, vol. 51, no. 13, pp. 1397–1430, Jul. 2011.
- [48] E. T. Rolls and T. J. Webb, "Finding and recognizing objects in natural scenes: Complementary computations in the dorsal and ventral visual systems," *Frontiers Comput. Neurosci.*, vol. 8, pp. 1–19, Aug. 2014.
- [49] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 411–426, Mar. 2007.
- [50] P. Sountsov, D. M. Santucci, and J. E. Lisman, "A biologically plausible transform for visual recognition that is invariant to translation, scale, and rotation," *Frontiers Comput. Neurosci.*, vol. 5, pp. 1–7, Nov. 2011.
- [51] S. Minaee, A. Abdolrashidi, H. Su, M. Bannamoun, and D. Zhang, "Biometric recognition using deep learning: A survey," 2019, *arXiv:1912.00271*. [Online]. Available: <http://arxiv.org/abs/1912.00271>
- [52] G. Wang, C. Sun, and A. Sowmya, "Multi-weighted co-occurrence descriptor encoding for vein recognition," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 375–390, Jan. 2020.
- [53] J. D. Cheng, Q. L. Sun, and J. X. Zhang, "Supervised hashing with deep convolutional features for palmprint recognition," in *Proc. 12th Chin. Conf. Biometric Recognit. (CCBR)*, Shenzhen, China, 2017, pp. 259–268.
- [54] S. Minaee and Y. Wang, "Palmprint recognition using deep scattering network," arXiv preprint 2016, *arXiv:1603.09027v1*. [Online]. Available: <https://arxiv.org/abs/1603.09027>
- [55] A. Genovese, V. Piuri, K. N. Plataniotis, and F. Scotti, "PalmNet: Gabor-PCA convolutional networks for touchless palmprint recognition," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 12, pp. 3160–3174, Dec. 2019.
- [56] A. S. Tarawneh, D. Chetverikov, and A. B. Hassanat, "Pilot comparative study of different deep features for palmprint identification in low-quality images," in *Proc. 9th Hung. Conf. Comput. Graph. Geometry*, Budapest, Hungary, 2018, pp. 1–6.
- [57] S. Zhao and B. Zhang, "Deep discriminative representation for generic palmprint recognition," *Pattern Recognit.*, vol. 98, Feb. 2020, Art. no. 107071.
- [58] G. Cheng, P. Zhou, and J. Han, "Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7405–7415, Dec. 2016.
- [59] G. Cheng, P. Zhou, and J. Han, "RIFD-CNN: Rotation-invariant and Fisher discriminative convolutional neural networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2884–2892.
- [60] G. Cheng, J. Han, P. Zhou, and D. Xu, "Learning rotation-invariant and Fisher discriminative convolutional neural networks for object detection," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 265–278, Jan. 2019.
- [61] R. L. T. Goris, E. P. Simoncelli, and J. A. Movshon, "Origin and function of tuning diversity in macaque visual cortex," *Neuron*, vol. 88, no. 4, pp. 819–831, Nov. 2015.
- [62] L. Kauffmann, S. Ramanoël, and C. Peyrin, "The neural bases of spatial frequency processing during scene perception," *Frontiers Integrative Neurosci.*, vol. 8, pp. 1–14, May 2014.
- [63] V. Talebi and C. L. Baker, "Categorically distinct types of receptive fields in early visual cortex," *J. Neurophysiol.*, vol. 115, no. 5, pp. 2556–2576, May 2016.
- [64] O. Arandjelovi, "Gradient edge map features for frontal face recognition under extreme illumination changes," in *Proc. Brit. Mach. Vis. Conf.*, Guildford, U.K., 2012, pp. 1–11.
- [65] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *J. Neurophysiol.*, vol. 58, no. 6, pp. 1233–1258, Dec. 1987.
- [66] P. Kovesi, "Phase congruency: A low-level image invariant," *Psychol. Res.*, vol. 64, no. 2, pp. 136–148, Dec. 2000.
- [67] D. Dacey, O. S. Packer, L. Diller, D. Brainard, B. Peterson, and B. Lee, "Center surround receptive field structure of cone bipolar cells in primate retina," *Vis. Res.*, vol. 40, no. 14, pp. 1801–1811, Jun. 2000.
- [68] D. Zhang, Z. H. Guo, L. Zhang, W. Zuo, and G. M. Lu, "An online system of multispectral palmprint verification," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 2, pp. 480–490, Mar. 2010.
- [69] *PolyU Palmprint Database*. Accessed: Aug. 2016. [Online]. Available: <https://www4.comp.polyu.edu.hk/~biometrics/>
- [70] *PolyU Multispectral Palmprint Database*. Accessed: Sep. 2013. [Online]. Available: <http://www4.comp.polyu.edu.hk/~biometrics/MultispectralPalmprint/MSP.htm>
- [71] *CASIA Palmprint Image Database*. Accessed: Sep. 2013. [Online]. Available: <http://biometrics.idealtest.org/dbDetailForUser.do?id=5>
- [72] *COEP Palmprint Database*. Accessed: Jul. 2017. [Online]. Available: <http://www.coep.org.in/resources/oeppalmprintdatabase.htm>
- [73] *Tongjiu Dataset*. Accessed: Aug. 2018. [Online]. Available: <http://sse.tongji.edu.cn/linzhang/contactlesspalm/index.htm>



XIANCHENG ZHOU received the Ph.D. degree from Central South University, in 2008. He was a Visiting Scholar with Rutgers, The State University of New Jersey, in 2012. He is currently a Professor with the Hunan University of Technology and Business. He is the Dean of the Graduate School and serves as the Vice-Chairman of the Fifth Academic Committee of the Hunan University of Technology and Business. He is the Executive Director of the Mobile E-Business Collaborative Innovation Center of Hunan Province. His research interests include image processing and biometrics.



KAIJUN ZHOU received the Ph.D. degree from Central South University, in 2009. He was a Visiting Scholar at the Volen National Center for Complex Systems, Brandeis University, in 2016. He is currently a Professor with the Hunan University of Technology and Business. His research interests are biologically inspired vision modeling and image recognition.



LIZHI SHEN received the master's degree from the Changsha University of Science and Technology, in 2011. He is currently pursuing the Ph.D. degree with Central South University. He is also a Lecturer with the Hunan University of Technology and Business. His research interest is palmprint recognition.

...