Multimed Tools Appl (2018) 77:17595–17623
DOI 10.1007/s11042-017-4965-6

Fast periocular authentication in handheld devices with reduced phase intensive local pattern

Sambit Bakshi[1] · Pankaj K. Sa[1] · Haoxiang Wang[2,3] · Soubhagya Sankar Barpanda[1] · Banshidhar Majhi[1]

Received: 8 April 2017 / Revised: 7 June 2017 / Accepted: 18 June 2017 /
Published online: 4 July 2017

Abstract To ensure highest security in handheld devices, biometric authentication has emerged as a reliable methodology. Deployment of mobile biometric authentication struggles due to computational complexity. For a fast response from a mobile biometric authentication method, it is desired that the feature extraction and matching should take least time. In this article, the periocular region captured through frontal camera of a mobile device is considered under investigation for its suitability to produce a reduced feature that takes least time for feature extraction and matching. A recently developed feature Phase Intensive Local Pattern (PILP) is subjected to reduction giving birth to a feature termed as Reduced PILP (R-PILP), which yields a matching time speed-up of 1.56 times while the vector is 20% reduced without much loss in authentication accuracy. The same is supported by experiment on four publicly available databases. The performance is also compared with one global feature: Phase Intensive Global Pattern, and three local features: Scale Invariant Feature Transform, Speeded-up Robust Features, and PILP. The amount of reduction can be

✉ Sambit Bakshi
bakshisambit@nitrkl.ac.in; sambitbaksi@gmail.com

Pankaj K. Sa
pankajksa@nitrkl.ac.in

Haoxiang Wang
hw496@goperception.com

Soubhagya Sankar Barpanda
soubhagya1984@gmail.com

Banshidhar Majhi
bmajhi@nitrkl.ac.in

[1] Department of Computer Science & Engineering, National Institute of Technology, Rourkela, Odisha 769 008, India

[2] Cornell University, Ithaca, NY, USA

[3] GoPerception Laboratory, New York, NY, USA

varied with the requirement of the system. The amount of reduction and the performance of the system bears a trade-off. Proposed R-PILP attempts to make periocular suitable for mobile devices.

Keywords Fast biometric matching · Biometric on mobile device · Periocular biometric · Phase intensive local pattern · Feature reduction

1 Introduction

Biometric has gained importance over two other classical authentication paradigms: token-based and knowledge-based methods due to its less chance of being spoofed or stolen. Face is the primitive way of human authentication and hence early researches have proven the possibility of machine based recognition of human face. Further need of partial face recognition, and the complexity of iris recognition, jointly advocated the research in the direction of periocular (periphery of ocular) biometrics. Periocular biometric is observed to contain gross features (prominent in visible spectrum) as well as subtle features (prominent in near-infrared spectrum). Hence periocular region of a person can be useful for recognition either in visible spectrum (VS) or near-infrared (NIR) imaging. Many researchers have applied several global and local feature extraction techniques to evaluate performance of periocular biometrics. In general, local features exhibit better performance than global feature based recognition in case of periocular biometric. However the local features, with the benefit of high accuracy of recognition, brings the overhead of large feature vector and slow feature extraction and matching. Among the latest local features, Phase Intensive Local Pattern (PILP) [4] is observed to produce highest accuracy as it considers a combination of coarse-to-fine features for recognition. PILP generates a large feature template, which needs huge storage for a database with many subjects. The 1:1 matching of two such large templates also requires huge amount of time. Hence it becomes compulsory to reduce the feature size of PILP for authentication through handheld mobile devices. This article discusses a reduction method of PILP. The reduced feature vector is termed as R-PILP. The reduced feature is tested on two NIR spectrum constrained databases: BATH, CASIAv3, and two VS spectrum partially unconstrained databases: UBIRISv2, FERETv4. The objective of evaluation over these four datasets is to evaluate its efficacy of recognition from subtle and gross features.

2 Related work

Last decade has seen a paradigm shift of research from highly-accurate constrained biometric recognition towards achieving moderate accuracy in unconstrained scenarios, and forwarding the latter to improve performance in terms of accuracy and time. Demand of using biometric system for forensic identification and terrorism investigation is responsible for the lurch. In an unconstrained scenario, difficulty in achieving a desirable biometric template can be broadly classified into two class of reasons:

1. *Intrinsic limitation of imaging system:* distant imaging, limitation of imaging spectrum, out-of-focus blur, and motion blur
2. *External environmental factors:* non-cooperation of subjects, out-plane (non-orthogonal) imaging, and illumination variance

Periocular biometric emerged as a counterpart to the pre-established iris biometric to support the following scenarios: to use the images categorized as 'Failure to Acquire' (full or partial closed eyes occluding iris) by an iris biometric system, to find a minimal subset of face biometric so that recognition from partial face images can take place, and to achieve recognition in visible spectrum images. It has also been observed that even most advanced preprocessing techniques are insufficient to bring unconstrained templates to the standards of ideal constrained templates. Feature extraction hence plays an important role to extract in-plane orthogonal transformation (viz. scaling and rotation) invariant features from the less-informative noisy templates. Through last two decades, researchers have established iris recognition systems which are highly accurate (in order of 99% in Identification mode) for NIR iris databases like BATH, CASIAv1, ICE, etc. The journey was pioneered by experiments on constrained iris databases. Particularly it can be marked that Proenca and Alexandre, in their work [24], have reported their system to perform with an Equal Error Rate (EER) of 1.01% on CASIA while the EER increases to 2.83% on noisy UBIRIS database. Likewise the work in [28] by Vatsa et al. operates with high accuracy on constrained datasets, but the performance degrades comparatively on UBIRIS which is a partially non-cooperative and noisy dataset. The cause behind this fall in accuracy had been analysed by the researchers as: *a*. visible spectrum imaging and *b*. unconstrained noises during acquisition for UBIRIS database. Hence an advanced research has proceeded to investigate whether recognition from unconstrained VS eye image is achievable. Research began to trace the answer to the following question: *Can addition of features from the periphery region further improve the recognition accuracy already achieved solely using iris?* While researchers had been yet to establish the answer to this question, in [8], Hollingsworth et al. had proposed the existence of features in periocular region in NIR images through experiments with human subjects. In first step of his approach, some human experts are shown a periocular image of a subject. In the next step the experts are shown few periocular images and asked to recognize which image belongs to the same subject shown in the first step. Subsequently accuracy is calculated depending on whether the experts can identify a subject from periocular image. Moreover, the human experts involved in the experiments are asked to mention which feature in the NIR image best helped their recognition judgement. Further research by Hollingsworth et al. in [9] is extended to investigate the existence of features in both NIR and VS periocular images through human expert analysis and automated algorithms. The work by Miller et al. [15] achieves the highest accuracy in periocular recognition among landmark works in this domain [1, 8, 11, 14–17, 19, 20, 26, 27, 29, 30]. However, the test in his work has been experimented on FRGC face database which constitutes VS images of high resolution of the order of 1704×2272 or 1200×1600 and the number of pixels between centers of two eyes are 250. Patel et al. [21] have discussed about periocular biometrics based kinship verification. Their proposed algorithm is based on *neighborhood repulsed metric learning* (NRML). Zhao and Kumar [31] have investigated about feature extraction from periocular region by employing *semantics-assisted convolutional neural networks* (SCNN). The proposed method is claimed to have superior performance compared to contemporary works with relatively smaller training set. Ahuja et al. [2] have proposed a *convolutional neural network* (CNN) based hybrid technique for ocular smartphone based biometrics. Their developed method employs spervised and unsupervised CNN augmented with Root SIFT model. However, test of recognition from low resolution VS periocular has remained unverified.

Local features extraction using automatic scale selection has been investigated by Chenhong and Zhaoyang [13] for normalized NIR acquired iris images. The technique deals

with first filtering the given normalized iris image by employing a bank of Laplacian of Gaussian (LoG) filters with many different scales and computes the normalized response of every filter. The maxima of normalized response over scales for each point are selected together as the optimal filter outputs of the given iris image. The iris feature is represented in a vector comprises off location and scale information, which is further binary coded for final presentation. In [3], conventional Local Binary Pattern (LBP) [18] followed by Scale Invariant Feature Transform (SIFT) has been applied to extract features from periocular region and tests the same on UBIRISv2 and FERETv4 databases. The accuracy of such feature extraction technique has been found to be limited to approximately 85%, which motivates further development of novel features that will pull up the accuracy for noisy VS periocular recognition to a higher level. Section 3 illustrates the proposed PILP technique along with its reduction to find R-PILP, which attempts to achieve periocular recognition from low resolution VS images. We have converted the VS color images to grayscale and the grayscale images (without separate color information) are subjected as input to the proposed approach. Thus we have further made the input least favourable for recognition and attempted to attain recognition through proposed R-PILP in this challenging scenario. As comparative literature, we have considered a global feature PIGP [3] and widely used local features: SIFT, SURF and PILP. SIFT and SURF are described in Sections 2.1 and 2.2 respectively. A detailed description of PILP is given in this article as the method of reducing PILP is very interlaced. PIGP is not described separately as PIGP comprises of finding feature descriptor in the same way as PILP.

2.1 Scale invariant feature transform

A local feature descriptor termed as Scale Invariant Feature Transform (SIFT) [12] is used as a comparative feature with the proposed one. SIFT provides stable set of features while being less sensitive to geometric transformations of area of interest. The feature is extracted with following steps: *a*. Keypoint Detection and *b*. Keypoint Descriptor Computation. SIFT extracts keypoints using Difference of Gaussian and a distribution of gradient orientation from a window around the interest point is described as descriptor. Finally Keypoints of one image is matched to other to find a distance between the images. If the distance is higher than a threshold, the two images are concluded to be captured from different subjects, else they are concluded to be of same subject.

2.1.1 Keypoint detection

The keypoints are detected from periocular image using cascade filtering approach. This is done to search stable features across all possible scales. To define the scale space, input periocular image (I) is convolved with Gaussian kernel $G(x, y, \sigma)$ as defined by

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{1}$$

where $*$ is the convolution operation and σ defines the width of Gaussian filter. The Difference of Gaussian (DoG) images are computed from two nearby scales differentiated by constant multiplicative factor k as:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{2}$$

DoG images are used to detect interest points with the help of local maxima and minima across different scales. Each pixel in DoG image is compared to 8 neighbors in the same scale and 9 neighbors in the scale above and below. The pixel is selected as a candidate keypoint if it is local maxima or minima in $3 \times 3 \times 3$ region.

2.1.2 Keypoint descriptor computation

Orientation is assigned to each keypoint location to achieve invariance to image rotations as descriptor can be represented relative to orientation. To determine keypoint orientation, a gradient orientation histogram is computed in the neighborhood of the keypoint. The scale of keypoint is used to select Gaussian smoothed image L. For each Gaussian smoothed image $L(x, y)$, magnitude ($m(x, y)$) and orientation ($\theta(x, y)$) are computed as given in (3).

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \tag{3}$$

$$\theta(x, y) = \tan^{-1}\left[\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right] \tag{4}$$

Orientation histogram is then formed for gradient orientation around each keypoint. The histogram has 36 bins for 360 degree range of orientations and each sample is weighted by gradient magnitude and Gaussian weighted circular window with σ of 1.5 times of scale of keypoint before adding it to histogram. Peaks in the histogram correspond to orientation and any other local peak within 80% of largest peak is used to create keypoint with the computed orientation. This is done to increase stability during matching [12].

Once orientation has been selected, the feature descriptor is computed as a set of orientation histograms on 4×4 pixel neighborhoods. These histograms contain 8 bins each and each descriptor contains an array of 16 histograms around the keypoint. This generates SIFT feature descriptor of $4 \times 4 \times 8 = 128$ elements. The descriptor vector is invariant to rotation, scaling, and illumination.

2.1.3 Keypoint matching

To match two images, corresponding feature sets are subjected to nearest neighbour matching and number of keypoints matched is considered as the parameter to interpret the degree of matching.

2.2 Speeded up robust features

A local feature descriptor termed as Speeded Up Robust Features (SURF) [6] is used as a comparative feature with the proposed one. SURF, like SIFT provides stable set of features while being less sensitive to geometric transformations of area of interest. However, the SURF descriptor is 64 dimensional while SIFT descriptor is 128 dimensional. The feature is extracted with following steps: *a*. Keypoint Detection and *b*. Keypoint Descriptor Computation. SURF extracts keypoints using Hessian matrix and a distribution of Haar wavelet responses from a window around the interest point is described as descriptor. Finally Keypoints of one image is matched to other to find a distance between the images. If the distance

is higher than a threshold, the two images are concluded to be captured from different subjects, else they are concluded to be of same subject.

2.2.1 Keypoint detection

Hessian matrix based interest point (keypoint) detection is adopted in SURF. For detection of keypoints, determinant of Hessian matrix is used for selecting location and scale. Given a point $P = (x, y)$ in an image I, the Hessian matrix $H(P, \sigma)$ in P at scale σ (where σ is the standard deviation of the Gaussian) can be found using (5).

$$H(P,\sigma) = \begin{bmatrix} L_{xx}(P,\sigma) & L_{xy}(P,\sigma) \\ L_{xy}(P,\sigma) & L_{yy}(P,\sigma) \end{bmatrix} \tag{5}$$

where $L_{xx}(P, \sigma)$ is obtained through convolution of the Gaussian second order derivative ($\frac{\sigma^2}{\sigma x^2} g(\sigma)$) with image I at point P. Likewise $L_{xy}(P, \sigma)$ and $L_{yy}(P, \sigma)$ can also be derived. D_{xx}, D_{xy}, and D_{yy} are discretized version of L_{xx}, L_{xy}, and L_{yy} respectively. The discretization is done to achieve different sized filter at different scales.

The approximation for Hessian determinant can be computed using (6).

$$Det(H_{approx}) = D_{xx}\, D_{yy} - (0.9 D_{xy})^2 \tag{6}$$

The scale space construction starts with 9×9 filter and then filters with sizes 15×15, 21×21, and 27×27 are applied. The increment in filter size is doubled for every new octave.

Keypoints are localized in scale and image space by applying a non-maximum suppression in $3 \times 3 \times 3$ neighbourhood. The local maxima found on the determinant of Hessian matrix are interpolated to image space.

2.2.2 Keypoint descriptor computation

Orientation of a circular window around every keypoint is found. The neighbourhood is split into 4×4 sub-regions. Haar wavelet responses of this circular neighbourhood of interest point is calculated for each sub-region. The size of wavelets are scale dependent. Haar wavelet responses are calculated in x and y direction separately. The feature vector v of a 4×4 sub-region can be obtained using (7).

$$v = \sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \tag{7}$$

where d_x and d_y are Haar responses obtained from the subregion in x and y direction respectively.

Concatenating the feature vector for all 4×4 sub-regions results in a descriptor with vector length of $4 \times 4 \times 4 = 64$ which represents the keypoint. For all detected keypoints, such 64-D feature vectors are extracted.

2.2.3 Keypoint matching

The keypoint matching is same as followed by SIFT. To match two images, corresponding feature sets are subjected to nearest neighbour matching and number of keypoints matched is considered as the parameter to interpret the degree of matching.

3 Phase intensive local pattern

This section illustrates the Phase Intensive Local Pattern (PILP) which is formed to match fine-to-coarse features in periocular VS image. A matching technique illustration follows the feature extraction section which explains a matching technique used for matching two patterns obtained through feature extraction.

3.1 Extraction of PILP feature

The feature extraction technique to achieve the final feature vector representing a periocular image comprises of four sequential steps: *a.* Keypoint detection through Phase Intensive Patterns, *b.* Edge feature removal, *c.* Oriented Histogram computation, and *d.* Feature vector formation. These steps are elaborated hereafter.

3.1.1 Keypoint detection through phase intensive patterns

The first step of a local feature extraction technique is to choose from the periocular image few points which holds the important features and are sufficient to uniquely describe the image and make the image identifiable. Such points are termed as keypoints. First part of keypoint generation (pattern generation) is exactly same as generation of patterns by PIGP [3]. Still we elaborate the process in this article for completeness of readability.

We vary our scale (Δ) for feature detection from 3 to 9 with an increment of 2. Subsequently, we will use a filter to trace the pattern in each scale which will vary in size from 3×3 to 9×9. The justification of choosing this range of the scale lies as the variation in distance of subject from camera is approximately three times of the minimum distance of camera and subject. At a given scale Δ, the phase-intensive global pattern (PIGP) at a pixel (x_c, y_c) with respect to its $\Delta^2 - 1$ neighbors considering a phase-tilt ϕ can be derived using (8). At said pixel (x_c, y_c), this equation convolves s-function and assigns a weight to each neighbouring pixel depending on its spatial location and the phase-tilt ϕ inclined to which

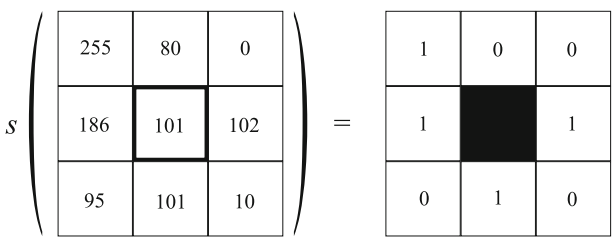


Fig. 1 An example how s-function works for PILP filter

we are aiming to extract the pattern. For example, when $\Delta = 3$, there are $3^2 - 1 = 8$ neighbors to any pixel. To find the PIGP at a pixel (x_c, y_c) at angle $\phi = \frac{\pi}{4}$, the 3×3 neighbors around the pixel are operated with corresponding pixel and result is obtained. Figure 1

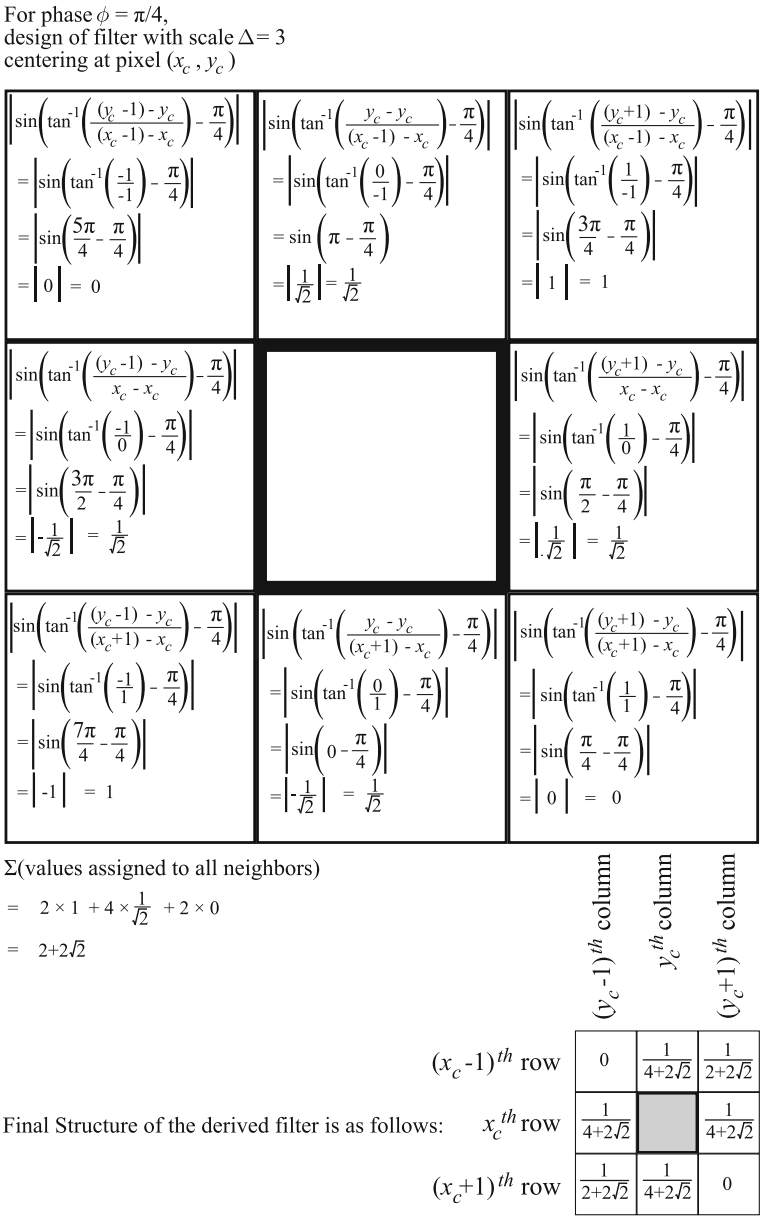


Fig. 2 Filter formation for PILP

illustrates the working of s-function for $\Delta = 3$, where the neighbors of center pixel are labeled as 0 if they are less than the center pixel, else labeled as 1.

$$PIGP(x_c, y_c, \Delta, \phi) = \frac{\sum_{n=1}^{\Delta^2-1} s(i_n, i_c).2^{\sin\left(\tan^{-1}\left(\frac{y_n-y_c}{x_n-x_c}\right)-\phi\right)}}{\sum_{n=1}^{\Delta^2-1} 2^{\sin\left(\tan^{-1}\left(\frac{y_n-y_c}{x_n-x_c}\right)-\phi\right)}}$$

$$= \sum_{n=1}^{\Delta^2-1} \left(s(i_n - i_c). \left(\frac{2^{\sin\left(\tan^{-1}\left(\frac{y_n-y_c}{x_n-x_c}\right)-\phi\right)}}{\sum_{n=1}^{\Delta^2-1} 2^{\sin\left(\tan^{-1}\left(\frac{y_n-y_c}{x_n-x_c}\right)-\phi\right)}} \right) \right) \tag{8}$$

where,

$$s(i_n, i_c) = \begin{cases} 1, & \text{if } i_n - i_c \geq 0 \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

This process can be re-represented as a convolution with filter formed with the process shown in Fig. 2. For each Δ the demonstrated process is repeated eight times varying the value of ϕ from 0 to $\frac{7\pi}{4}$ with an interval of $\frac{\pi}{4}$. Figure 3 demonstrates the eight filters formed in scale $\Delta = 3$. We consider four unique filters out of these eight, as every two filters with phase difference of π are having same structure (can be observed in Fig. 3). Figure 4 shows four 3×3 filters in filter bank for $\Delta = 3$. Similar filter banks are formed for other values of Δ. Finally we obtain four filter banks (corresponding to $\Delta = 3, 5, 7, 9$), each having four filters within the bank. The spatial size of all four filters for scale Δ is $\Delta \times \Delta$.

The overall flow of keypoint detection methodology is as follows. When the aforementioned four filters of a filter bank are convolved with the original image, four pattern images

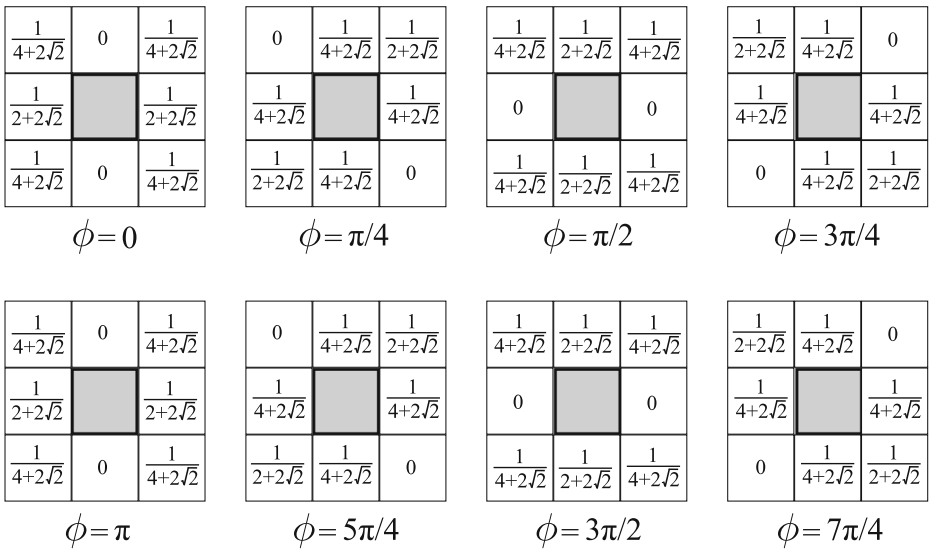


Fig. 3 PILP filter bank

are found. These four images are subjected to extrema detection, where each pixel of a pattern image at phase ϕ is compared with its 27 neighbours (including the pixel itself) in same phase and phase images of $[(\phi+\pi/4) \mod 2\pi]$ and $[(\phi-\pi/4) \mod 2\pi]$ (refer to Fig. 5). These extrema are claimed as potential keypoints containing features.

To understand the whole feature extraction process mathematically, the following symbols are introduced:

I:	Original image
$f_{\Delta,i\pi/4}$	Filter of size $\Delta \times \Delta$ corresponding to phase ϕ varying as $i\pi/4$ ($i = 0, 1, 2, 3$)
F_Δ	Filter bank comprising four $\Delta \times \Delta$ filters, i.e., $F_\Delta \equiv \{f_{\Delta,i\pi/4} \mid i = 0, 1, 2, 3\}$
$I_{\Delta,i\pi/4}$	Result of convolution of I with $f_{\Delta,i\pi/4}$ and s-function
$k_{\Delta,i\pi/4}$	keypoints found from local minima in $I_{\Delta,[i \mod 4]\pi/4}$ with respect to itself and neighbouring pixels in $I_{\Delta,[i \mod 4]\pi/4}$, and $I_{\Delta,[i \mod 4]\pi/4}$
K_Δ	Set of all keypoints in $\Delta \times \Delta$ scale, i.e., $K_\Delta \equiv \{k_{\Delta,i\pi/4} \mid i = 0, 1, 2, 3\}$
K	Set of all keypoints from all scales, i.e., $K \equiv \bigcup_{\Delta=3,5,7,9} K_\Delta$

The whole keypoint extraction procedure is given in Algorithm 1 which is further schematically presented in Fig. 6. The algorithm is computation intensive as it involves convolution operation and pixel-wise extrema detection. However, the computation is worth as it yields sufficient number of keypoints.

Algorithm 1 Keypoint_Extraction_PILP

Require: I: Original image; F_3, F_5, F_7, F_9: Filter banks
Ensure: K: Extracted keypoints from I

for $\Delta := 3$ to 9 **do**
  $K_\Delta \equiv \Phi$
  for $i := 0$ to $3\pi/4$ **do**
    $I_{\Delta,i\pi/4} \leftarrow I \oplus f_{\Delta,i\pi/4} \oplus s$
    $i \leftarrow i + \pi/4$
  end for
  for $i := 0$ to $3\pi/4$ **do**
    for each pixel (x, y) in $I_{\Delta,[i \mod 4]\pi/4}$ **do**
      if $I_{\Delta,[i \mod 4]\pi/4}(x, y)$ is extrema among its neighbours in $I_{\Delta,[i \mod 4]\pi/4}$, $I_{\Delta,[(i-1) \mod 4]\pi/4}$, and $I_{\Delta,[(i+1) \mod 4]\pi/4}$ **then**
        $K_\Delta \leftarrow K_\Delta \cup (x, y)$
        $i \leftarrow i + \pi/4$
      end if
    end for
  end for
  $\Delta \leftarrow \Delta + 2$
end for
$K \leftarrow K_3 \cup K_5 \cup K_7 \cup K_9$
return K

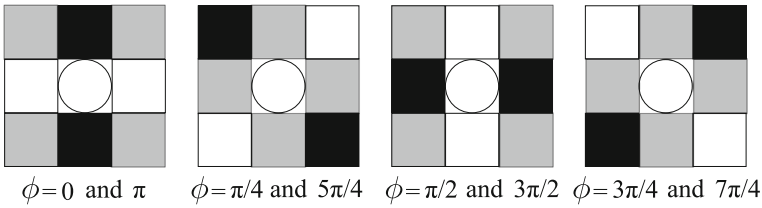


Fig. 4 Intensity representation of PILP orthogonal filter bank

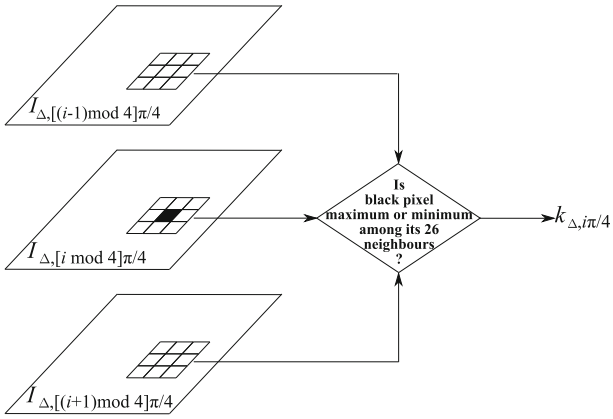


Fig. 5 Extrema detection method

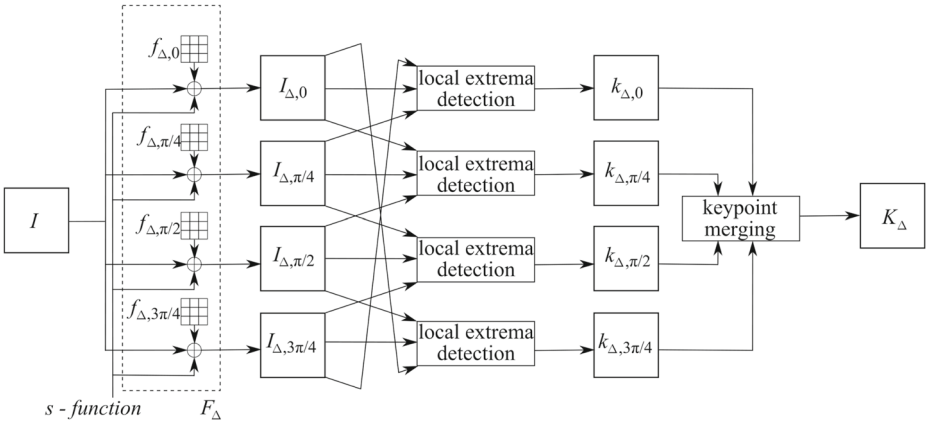


Fig. 6 Keypoint extraction method

3.1.2 Edge feature removal

It is necessary to determine keypoints belonging to edges and to remove them without further processing. For a keypoint at (x, y) in image $I_{\Delta,i\pi/4}$ where $i = 0, 1, 2, 3$, the magnitude $m_{\Delta,i\pi/4}(x, y)$ is computed as given in (10). The value of $m_{\Delta,i\pi/4}(x, y)$ will be high if there is any edge at (x, y), otherwise the value will be low.

$$m_{\Delta,i\pi/4}(x, y) = \sqrt{(I_{\Delta,i\pi/4}(x+1, y) - I_{\Delta,i\pi/4}(x-1, y))^2 + (I_{\Delta,i\pi/4}(x, y+1) - I_{\Delta,i\pi/4}(x, y-1))^2} \tag{10}$$

A high value of m denotes a keypoint to be on edge and those keypoints are discarded through proper thresholding and not considered in further processing [12].

3.1.3 Keypoint descriptor computation

Orientation is assigned to each valid keypoint location to achieve invariance to image rotations as descriptor can be represented relative to orientation. To determine keypoint orientation, a gradient orientation histogram is computed in the neighborhood of the keypoint. For a keypoint at (x, y) in image $I_{\Delta,i\pi/4}$ where $i = 0, 1, 2, 3$, the orientation $\theta_{\Delta,i\pi/4}(x, y)$ is computed as given in (11).

$$\theta_{\Delta,i\pi/4}(x, y) = \tan^{-1}\left[\frac{I_{\Delta,i\pi/4}(x, y+1) - I_{\Delta,i\pi/4}(x, y-1)}{I_{\Delta,i\pi/4}(x+1, y) - I_{\Delta,i\pi/4}(x-1, y)}\right] \tag{11}$$

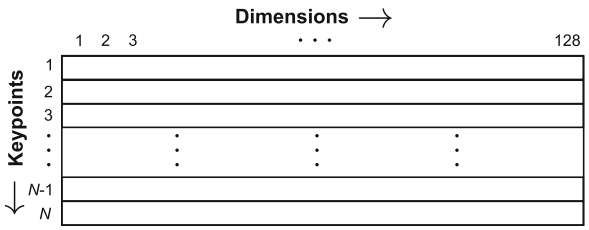
For every detected keypoint (x, y) in $I_{\Delta,i\pi/4}$, the orientation histogram is formed from a 16×16 pixel block oriented at $\theta_{\Delta,i\pi/4}(x, y)$ around each keypoint. The 16×16 pixel block under consideration is further divided into 16 sub-blocks, each having 4×4 pixels. For each sub-block, a histogram is formed which has 36 bins for 2π range of orientations. Peaks in the histograms correspond to peak within 80% of largest peak. Each histogram is represented with 8 of its most significant directional values (peaks). Hence 16 sub-blocks will produce together $16 \times 8 = 128$ peaks which represents the orientation of the 16×16 window around the keypoint (x, y). These 128 dimensional value is used to represent the detected keypoint in the feature vector.

As described, we obtain 128 values corresponding to each keypoint. Hence, we obtain a vector of size $128 \times N$ where N is the number of detected keypoints. This PILP feature vector is claimed to represent each identification uniquely and capable of matching periocular images.

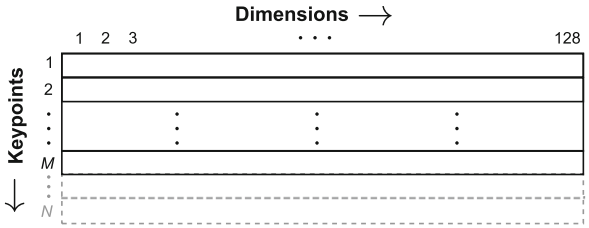
3.1.4 Why reduction is needed?

It is evident from Section 3.1.3, if N keypoints are detected through PILP in a periocular image, the corresponding feature vector size becomes $N \times 128$. This feature size is large when N is large, and feature extraction is time consuming. Furthermore, the feature matching becomes time intensive. To reduce the feature size, we can adopt three mechanisms: *a*. to prune number of keypoints (N) *b*. to reduce the number of keypoints by applying clustering technique and/or *c*. to reduce the dimension of features from 128 to some lower value.

Figure 7 shows the schematic diagram of the proposed mechanism. In the proposed mechanism, if an image yields $N \times 128$ PILP feature vector as shown in Fig. 7a (where N is the number of keypoints detected), some keypoints among N will be pruned and finally



(a) Block diagram of feature vector of PILP



(b) Block diagram of feature vector of R-PILP

Fig. 7 Feature reduction of PILP to obtain R-PILP feature set

Algorithm 2 Feature_Reduction_R-PILP

Require: F: Feature vector of size $N \times 128$ from PILP
Ensure: F': Feature vector of size $0.8N \times 128$ from R-PILP

$F' \leftarrow \phi$
$Redundancy \leftarrow$ A zero-matrix with N rows and 1 column
for $i := 1$ to N **do**
  $sum_i \leftarrow 0$
  $k \leftarrow 1$
  for j := k to 128 **do**
    if $F(i, j : j + 8 - 1)$ is a monotonic sequence **then**
      Differentiate $F(i, j : j + 8 - 1)$ until it yields 4 zeros
      $sum_i \leftarrow sum_i +$ number of differentiations
    end if
    if $F(i, j : j + 8 - 1)$ is not a monotonic sequence **then**
      $sum_i \leftarrow sum_i + 0$
    end if
    $k \leftarrow (k + 8)$
  end for
  $Redundancy(i) \leftarrow sum_i$
end for
$F' \leftarrow F' \cup$ rows of F corresponding to zero$Redundancy$
Sort F according to inverse of non-zero $Redundancy$
$F' \leftarrow F' \cup (F -$ lowest $(20\% \times N)$ rows of $F)$
return F'

M (where $M < N$) is retained in the reduced PILP feature set to yield a $M \times 128$ feature vector as shown in Fig. 7b.

There is an undoubted need of sequencing the features from high relevance to low relevance in order to ensure pruning of least significant features. Pruning of least significant features cannot hamper the recognition rate, whereas removal of any important feature can lead to compromising final recognition accuracy.

3.1.5 Proposed methodology: reduced phase intensive local pattern (R-PILP)

Each SIFT and PILP descriptor contains an array of 16 histograms (each having 8 values) around the keypoint, and hence each keypoint descriptor is of size $16 \times 8 = 128$. Hence it is clear that set of every 8 values within 128 values are independent. Each set of 8 values do have cohesion among themselves but such two sets are purely uncorrelated. The proposed method aims towards checking if every of these 8-value sequences are monotonic (consistently increasing or decreasing in nature). If an 8-value sequence is non-monotonic its differentiation can never produce zero values. However, if an 8-value sequence is monotonic, it will certainly produce 50% zero values after finite number of differentiations. A 128-D feature with most of its parts as monotonic is of less significant than another 128-D feature with most of its parts as non-monotonic. We thereby sequence the features and remove least-significant features. The overall process is given in Algorithm 2.

One important issue to ensure in this reduction process is not to loose any significant feature as that will harm the recognition accuracy. Hence we have experimented to ensure with how much reduction we can achieve almost similar accuracy as PILP. It is empirically found that upto 20% reduction in PILP feature does not affect the recognition accuracy. However, going beyond this limit and pruning more features degrade the recognition immediately.

An example of the reduction technique is demonstrated on a small region of an UBIRISv2 image in Fig. 8. The proposed reduction technique can also be applied to SIFT as well as SURF due to their similarity in nature of orientation histogram-based feature extraction. The technique is scalable in terms of choosing number of features or degree of reduction. Compromising on accuracy, degree of reduction can be tuned as per demand of system to be developed. If a system is to be developed in which security is prime issue and

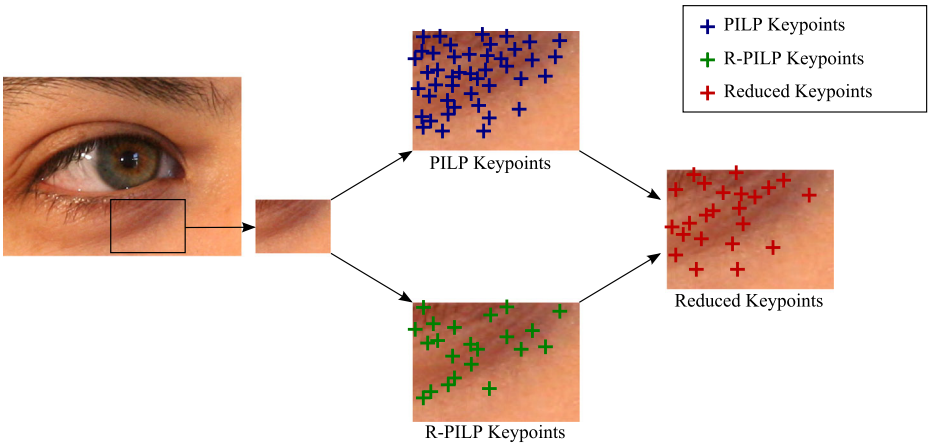


Fig. 8 Example of feature reduction on a sample UBIRISv2 periocular image

no compromise on accuracy can be made, the reduction can be limited to very small portion of total number of features. On the other hand, if a system is required to be developed that needs minimal template storage but agree to compromise to recognition accuracy, then this reduction technique can be applied to prune a high portion of input feature set.

3.2 Matching of proposed R-PILP feature

The matching algorithm plays a significant role in any biometric system. In local feature matching, the total number of paired keypoints (similarity score) is considered to find the authenticity of an individual. Let I be the set of all images available in the database. For understanding, I_m be a gallery iris image and I_n be a probe image where $I_m, I_n \epsilon\ I$. Let K_m be the set of p keypoints found in I_m and K_n be the set of q keypoints found in I_n by applying R-PILP local feature detector. Let D_m and D_n denote the set containing keypoint descriptors for each keypoint in K_m and K_n respectively. For each element in D_m the Euclidean distance is found with every element in D_n. $|D_m|$ denotes number of keypoint found in I_m and $|D_n|$ denotes number of keypoint found in I_n. The nearest neighbor approach pairs the i^{th} element in D_m with j^{th} element in D_n, iff the descriptor distance between the two (after multiplying with a threshold) is minimum. The details of the algorithm is given in Algorithm 3.

Algorithm 3 Nearest_Neighbour_Matching

Require: D_m: Keypoint descriptors for I_m, D_n: Keypoint descriptors for I_n
Ensure: R: Matched keypoint pairs

$\omega \leftarrow$ Threshold for pairing
while $i \leq |D_m|$ and $|D_n| \neq 0$ **do**
  for $j := 1$ to $|D_n|$ **do**
    $E_j \leftarrow \sqrt{\sum_{k=1}^{128} [(D_{m_i})_k - (D_{n_j})_k]^2}$
  end for
  $[min\_value\ index] \leftarrow$ minimum(E)
  $E_{index} \leftarrow min\_value \times \omega$
  $[min\_value\ new\_index] \leftarrow$ minimum(E)
  if $index = new\_index$ **then**
    $R_i \leftarrow (i, index)$
    $D_m \leftarrow D_m - D_{m_i}$
    $D_n \leftarrow D_n - D_{n_{index}}$
  end if
end while
return R

4 Experimental results

The performance of the proposed R-PILP feature is compared with two features developed successively: PIGP and PILP. As R-PILP is a successor to these features, comparing with them truly reflects the improvement achieved due to reduction. The comparison is made with respect to accuracy yielded, template size (proportional to number of keypoints), and feature extraction and matching time. Two landmark works SIFT and SURF are also used for comparison.

4.1 Performance measures

The following measures are employed to evaluate the performance of the proposed biometric system:

1. False Acceptance Rate (FAR): FAR is the frequency of fraudulent access to imposter claiming identity. This statistic is used to measure biometric performance when operating in the verification mode. A false accept occurs when the query template of an individual is incorrectly matched to existing biometric template of another individual.
2. False Rejection Rate (FRR): FRR is the frequency of rejections relative to people who should be correctly verified. This statistics is used to measure biometric performance when operating in the verification mode. A false reject occurs when an individual is not matched correctly to his/her own existing biometric template.
3. Accuracy (Acc): Accuracy of a biometric system is defined as the rate of true acceptance and true rejection. Particularly accuracy of a biometric system can be defined as illustrated in (12).
$$Acc = \left(100 - \frac{FAR + FRR}{2}\right)\% \tag{12}$$
4. Receiver Operating Characteristic (ROC): ROC curve depicts the dependence of FRR with GAR [Genuine Acceptance Rate (GAR) $= 1 -$ FRR] for change in the value of threshold. The curve is plotted using linear, logarithmic or semi-logarithmic scales. ROC can also be represented by plotting FRR against FAR for change in the threshold value.
5. Cumulative Match Characteristic (CMC): The rank-k identification indicates the number of correct identification that occur in top k matches. Let R_k denote the number of elements of probe set in top k, then the probability of identification is given by $I = R_k/N$. CMC curve represents the probability of identification (I) at various ranks k.
6. Decidability index or d' index: d' index [10] measures the separation between the arithmetic means of the genuine and imposter probability distribution in standard deviation units as defined in (13).
$$d' = \frac{\sqrt{2}\,|\mu_{genuine} - \mu_{imposter}|}{\sqrt{\sigma^2_{genuine} + \sigma^2_{imposter}}} \tag{13}$$

4.2 Databases used

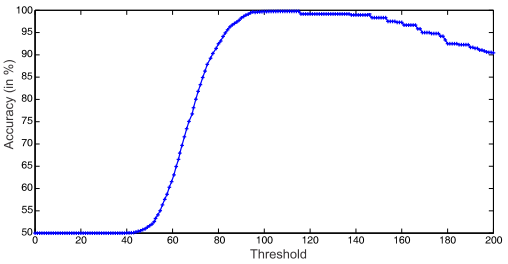
The proposed R-PILP is tested on publicly available BATH and CASIAv3 database to demonstrate its accuracy on subtle features available in NIR images. To evaluate the performance of R-PILP on gross features, noisy unconstrained images of UBIRISv2 and periocular region cropped from FERETv4 are used. A detail of these databases can be found in Table 1.

4.3 Experiment 1: evaluation of performance on subtle features

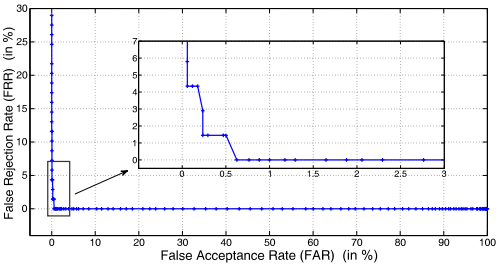
In this experiment, the proposed reduction approach is applied to PILP features obtained from a high resolution NIR image. Testing of this experiment is made on BATH and CASIAv3 databases which contains high resolution intensity-variation compensated NIR images of ocular region captured in orthogonal view. The performance results (Accuracy

Table 1 Detail of publicly available testing databases used for evaluation of proposed approach

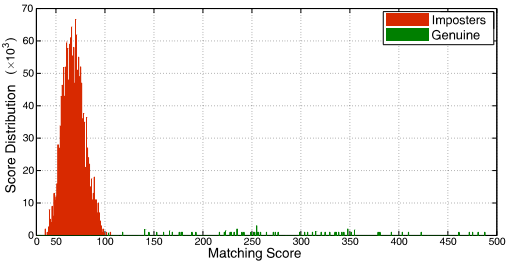
Database	Developer	Version	Images	Subjects	Resolution	Color Model	Spectrum
BATH [5]	University of Bath, Bath, United Kingdom	Iris DB 400	8000	200	1280 × 960	Grayscale	NIR
		Iris DB 800	16000	400			
		Iris DB 1600	32000	800			
CASIA [7]	Institute of Automation, Chinese Academy of Sciences, Beijing, China	IRISv3-Lamp	16212	411	640 × 480	Grayscale	NIR
UBIRIS	Soft Computing and Image Analysis (SOCIA) Group, Department of Computer Science, University of Beira Interior, Portugal	v1 [23]	1877	241	800 × 600	RGB	VS
		v2 [25]	11102	261	400 × 300	sRGB	
FERET [22]	National Institute of Standards and Technology (NIST), Gaithersburg, Maryland	v4	14126	1191	768 × 512 384 × 256 192 × 128	RGB	VS



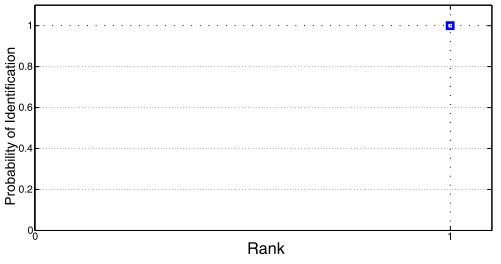
(a) Accuracy curve: Accuracy 99.82%, FAR 0.11%, FRR 0.25%, d' index 3.18



(b) ROC curve

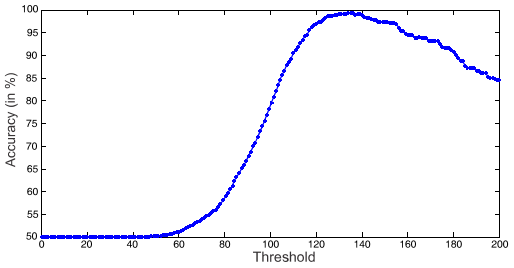


(c) Score distribution

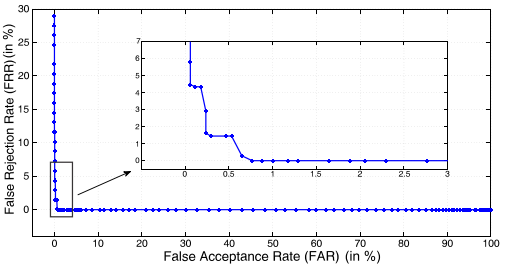


(d) CMC curve

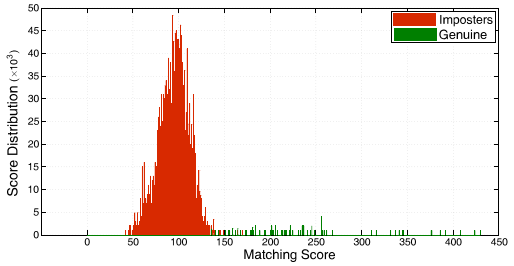
Fig. 9 Performance of R-PILP with NN matching technique on BATH database



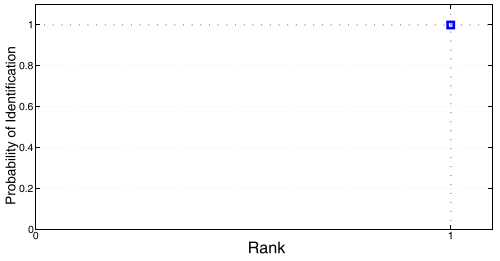
(a) Accuracy curve: Accuracy 99.47%, FAR 0.38%, FRR 0.68%, d' index 2.67



(b) ROC curve



(c) Score distribution



(d) CMC curve

Fig. 10 Performance of R-PILP with NN matching technique on CASIAv3 database

curve, ROC curve, Score distribution, and CMC curve) of R-PILP on BATH and CASIAv3 databases are shown in Figs. 9 and 10 respectively. Further quantitative results can be found in Table 2, which justifies that R-PILP can perform equally well for NIR databases as PILP.

4.4 Experiment 2: evaluation of performance on gross features

In this experiment, the proposed reduction approach is applied to PILP features obtained from a low resolution VS image. Testing of this experiment is made on UBIRISv2 and FERETv4 databases which contains low resolution color VS images of ocular region captured in unconstrained view from a distance.

The performance results (Accuracy curve, ROC curve, Score distribution, and CMC curve) of R-PILP on UBIRISv2 and FERETv4 databases are shown in Figs. 11 and 12 respectively. Further quantitative results can be found in Table 2, which justifies that R-PILP can perform equally well for VS databases as PILP.

4.5 Comparative analysis of R-PILP with PIGP and PILP

The discussed post-reduction technique takes as input feature vector F of size $N \times 128$ from PILP and reduces its keypoints to $0.8N$ so that the reduced feature vector F' has size $0.8N \times 128$. In case of PILP, if two images (one gallery and one probe) with feature vectors of size $N_1 \times 128$ and $N_2 \times 128$ are to be matched, Nearest Neighbour (NN) matching technique will operate $N_1 \times N_2$ distance calculations. If one 128-D disctance calculation costs t time, then total NN distance calculation will consume $t_{PILP} = (N_1 \times N_2) \times t$ time.

However, if the same process is executed on the same images with the proposed reduction technique, the gallery and probe image will have feature vectors of size $0.8N_1 \times 128$ and $0.8N_2 \times 128$ after reduction. Hence NN matching needs only $0.8N_1 \times 0.8N_2$ distance calculations. So, the total NN distance calculation in this case will consume $t_{R-PILP} = (0.8N_1 \times 0.8N_2) \times t$ time.

Hence, obtained speed-up of R-PILP with respect to PILP (Speed-up $^{R\text{-}PILP}_{PILP}$) achieved through the reduction can be calculated as shown in (14).

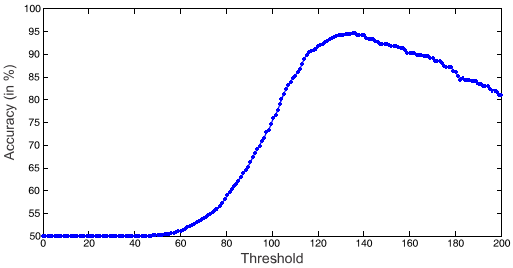
$$\begin{aligned} \text{Speed-up}^{R-PILP}_{PILP} &= \frac{1/t_{R-PILP}}{1/t_{PILP}} \\ &= \frac{t_{PILP}}{t_{R-PILP}} \\ &= \frac{(N_1 \times N_2) \times t}{(0.8N_1 \times 0.8N_2) \times t} \\ &= \frac{1}{0.8 \times 0.8} \\ &= 1.5625 \end{aligned} \tag{14}$$

Though this reduction technique consumes time for reducing the PILP feature, but that process takes place only once. The reduced feature R-PILP speeds up matching process with a factor of 1.56 with respect to PILP which is significant as matching process executes every time a live query comes. Moreover, reduction of these 20% keypoints does not affect in performance of R-PILP comparing with its predecessor PILP.

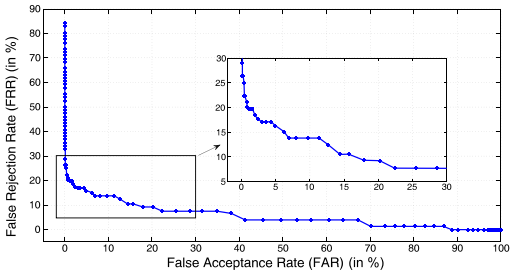
To advocate this theoretical analysis, Table 3 presents the number of keypoints found in an image through SIFT, SURF, PILP, and R-PILP. The table presents the number of keypoints as a range in $[Q1$ - $Q3]$, where, $N/4$ images yield $< Q1$ keypoints and $N/4$

Table 2 Comparison of performance of PIGP, PILP, and R-PILP on BATH, CASIAv3, UBIRISv2, and FERETv4 databases

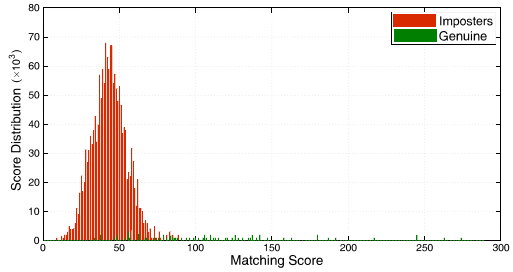
		PIGP			SIFT			SURF			PILP			R-PILP		
Approaches → Databases ↓		ACC (%)	d' index	Rank-1 RR(%)	ACC (%)	d' index	Rank-1 RR(%)	ACC (%)	d' index	Rank-1 RR(%)	ACC (%)	d' index	Rank-1 RR(%)	ACC (%)	d' index	Rank-1 RR(%)
NIR Databases	BATH	99.40	3.19	99.75	100	3.36	100	98.46	3.31	100	99.87	3.31	100	99.82	3.18	100
	CASIAv3	98.23	2.89	99.25	99.68	2.73	100	99.99	2.90	98.30	99.62	2.90	100	99.47	2.67	100
VS Databases	UBIRISv2	87.31	1.472	82.86	90.27	1.78	84.97	83.13	1.58	81.99	95.51	2.29	87.62	94.95	2.12	86.43
	FERETv4	81.22	1.04	72.37	87.33	1.53	83.20	81.86	0.93	76.33	92.47	2.01	85.80	92.04	1.88	85.64



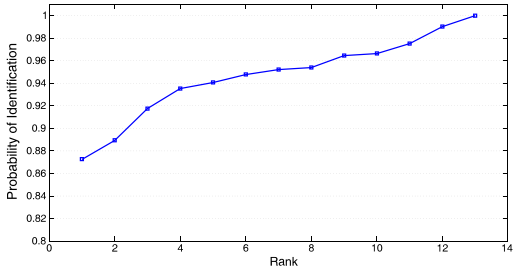
(a) Accuracy curve: Accuracy 94.95%, FAR 3.33%, FRR 6.77%, d' index 2.12



(b) ROC curve

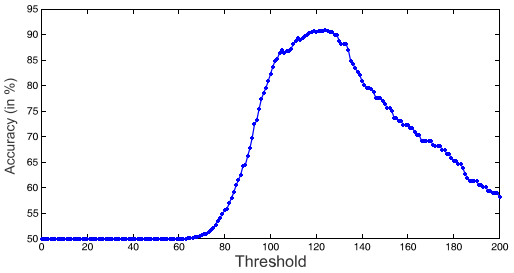


(c) Score distribution

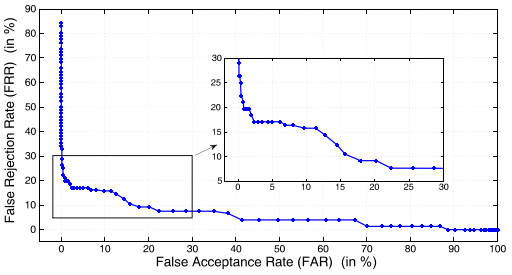


(d) CMC curve

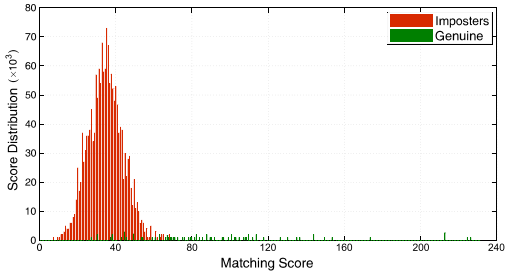
Fig. 11 Performance of R-PILP with NN matching technique on UBIRISv2 database



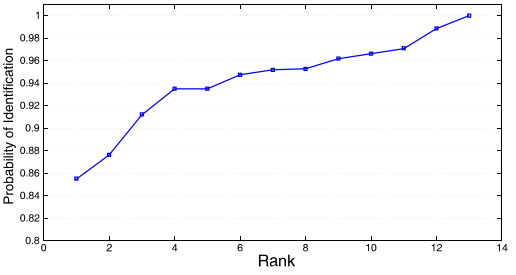
(a) Accuracy curve: Accuracy 92.04%, FAR 6.69%, FRR 9.23%, d' index 1.88



(b) ROC curve



(c) Score distribution



(d) CMC curve

Fig. 12 Performance of R-PILP with NN matching technique on FERETv4 database

Table 3 Comparison of time of feature extraction by SIFT, SURF, PILP, and proposed R-PILP on BATH, CASIAv3, UBIRISv2, and FERETv4 databases

Approaches →		SIFT	SURF	PILP	Proposed R-PILP
Databases ↓		Number of key-points/image	Number of key-points/image	Number of key-points/image	Number of key-points/image
NIR Databases	BATH	[75 – 250]	[50 – 226]	[525 – 825]	[400 – 675]
	CASIAv3	[575 – 900]	[325 – 650]	[1375 – 2150]	[1100 – 1750]
VS Databases	UBIRISv2	[500 – 725]	[200 – 450]	[1250 – 2100]	[975 – 1675]
	FERETv4	[425 – 650]	[150 – 325]	[1050 – 1850]	[825 – 1500]

Table 4 Comparison of number of keypoints extracted by SIFT, SURF, PILP, and proposed R-PILP on BATH, CASIAv3, UBIRISv2, and FERETv4 databases

Approaches →		SIFT	SURF	PILP	Proposed R-PILP
Databases ↓		Average time for feature extraction (s)	Average time for feature extraction (s)	Average time for feature extraction (s)	Average time for feature extraction (s)
NIR Databases	BATH	0.54	0.28	1.97	4.27
	CASIAv3	1.52	0.49	6.40	14.33
VS Databases	UBIRISv2	1.39	0.46	5.99	14.28
	FERETv4	1.37	0.44	5.78	13.04

Table 5 Comparison of time of feature matching by SIFT, SURF, PILP, and proposed R-PILP on BATH, CASIAv3, UBIRISv2, and FERETv4 databases

Approaches →		SIFT	SURF	PILP	Proposed R-PILP
Databases ↓		Average time for feature matching (s)	Average time for feature matching (s)	Average time for feature matching (s)	Average time for feature matching (s)
NIR Databases	BATH	0.61	0.42	3.81	2.38
	CASIAv3	1.54	0.73	3.59	2.28
VS Databases	UBIRISv2	1.2	0.68	3.75	2.41
	FERETv4	0.99	0.62	2.69	1.71

images yield $> Q3$ keypoints when a feature extraction method is applied on a database of size N. However, from Table 4, it is clear that R-PILP takes more time for feature extraction than SIFT, SURF, and PILP.

As feature matching time is theoretically proportional to the number of keypoints, in Table 5 it can be observed that matching time for R-PILP is less than that of PILP by a factor of 63%, which supports our theoretical analysis.

5 Conclusion

R-PILP, having 20% reduction from PILP, delivers approximately the same performance as PILP, as confirmed by testing on four datasets. However, reduction beyond this 20% margin degrades the performance as found experimentally. We have hence empirically limited our reduction to 20% of existing keypoints only. The reduction, while maintaining approximately the same accuracy as PILP, offers two additional benefits: *a*. a low template size resulting in reduction of the whole database and query communication cost for a networked biometric system and *b*. reduced matching time making the response to user query faster. The matching time of R-PILP is theoretically analysed to speed up 1.56 times than matching time of PILP. The analysis has been supported by experimentation where 63% reduction in matching time of R-PILP with respect to PILP is obtained. This reduction in matching time, with help of removal of outlier features, surely places R-PILP as a candidate feature for mobile biometric authentication systems.

Abbreviations used

ACC :	Accuracy
CMC :	Cumulative Match Characteristics
FAR :	False Acceptance Rate
FRR :	False Rejection Rate
PIGP :	Phase Intensive Global Pattern
PILP :	Phase Intensive Local Pattern
R-PILP :	Reduced Phase Intensive Local Pattern
ROC :	Receiver Operating Characteristic
RR :	Recognition Rate
SIFT :	Scale Invariant Feature Transform
SURF :	Speeded Up Robust Features

Acknowledgments The research is funded by Grant no. 12(5)/2012-ESD by Department of Electronics and Information Technology, Government of India. This research is an extension to:

S. Bakshi, P.K. Sa, and B. Majhi. A Novel Phase-intensive Local Pattern for Periocular Recognition under Visible Spectrum. *Biocybernetics and Biomedical Engineering* 35(1):30–44, 2015. doi:10.1016/j.bbe.2014.05.003.

References

1. Adams J, Woodard DL, Dozier G, Miller PE, Bryant K, Glenn G (2010) Genetic - based type II feature extraction for periocular biometric recognition: less is more. In: 20th international conference on pattern recognition (ICPR), pp 205–208. doi:10.1109/ICPR.2010.59
2. Ahuja K, Islam R, Barbhuiya FA, Dey K (2017) Convolutional neural networks for ocular smartphone-based biometrics. Pattern Recogn Lett 91:17–26. doi:10.1016/j.patrec.2017.04.002

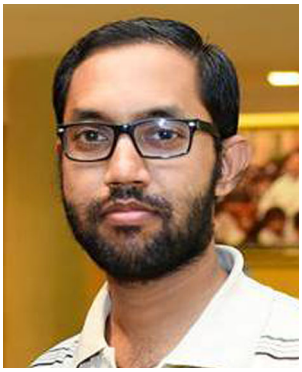
3. Bakshi S, Sa PK, Majhi B (2014) Phase intensive global pattern for periocular recognition. In: Proceedings of the the 11th IEEE India conference (INDICON). doi:10.1109/INDICON.2014.7030362
4. Bakshi S, Sa PK, Majhi B (2015) A novel phase-intensive local pattern for periocular recognition under visible spectrum. Biocybernetics Biomed Eng 35(1):30–44. doi:10.1016/j.bbe.2014.05.003
5. BATH University database. http://www.bath.ac.uk/elec-eng/research/sipg/irisweb, last accessed on 4.4.2016
6. Bay H, Ess A, Tuytelaars T, Gool LV (2008) Speeded-up robust features (SURF). Comput Vis Image Underst 110(3):346–359. doi:10.1016/j.cviu.2007.09.014
7. CASIA Database. http://www.cbsr.ia.ac.cn/english/irisdatabase.asp, last accessed on 4.4.2016
8. Hollingsworth KP, Bowyer KW, Flynn PJ (2010) Identifying useful features for recognition in near-infrared periocular images. In: 4th IEEE international conference on biometrics: Theory, Applications and Systems (BTAS). doi:10.1109/BTAS.2010.5634529
9. Hollingsworth KP, Darnell SS, Miller PE, Woodard DL, Bowyer KW, Flynn PJ (2010) Human and machine performance on periocular biometrics under near-infrared light and visible light. IEEE Trans Inf Forensics Secur 7:588–601. doi:10.1109/TIFS.2011.2173932
10. Jain AK, Flynn P, Ross A (2008) Handbook of biometrics. Springer. ISBN : 978-0-387-71040-2
11. Juefei-Xu F, Cha M, Heyman J, Venugopalan S, Abiantun R, Savvides M (2010) Robust local binary pattern feature sets for periocular biometric identification. In: 4th IEEE international conference on biometrics: theory, applications, and systems (BTAS). doi:10.1109/BTAS.2010.5634504
12. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110. doi:10.1023/B:VISI.0000029664.99615.94
13. Lu C, Lu Z (2008) Local feature extraction for iris recognition with automatic scale selection. Image Vis Comput 26(7):935–940. doi:10.1016/j.imavis.2007.10.011
14. Mahalingam G, Ricanek Jr K (2013) LBP-Based periocular recognition on challenging face datasets. EURASIP Journal on Image and Video Processing, 36. doi:10.1186/1687-5281-2013-36
15. Miller PE, Lyle JR, Pundlik SJ, Woodard DL (2010) Performance evaluation of local appearance based periocular recognition. In: 4th IEEE international conference on biometrics: theory, applications and systems (BTAS). doi:10.1109/BTAS.2010.5634536
16. Miller PE, Rawls AW, Pundlik SJ (2010) Personal identification using periocular skin texture. In: Proceedings of the 2010 ACM symposium on applied computing (SAC '10), pp 1496–1500. doi:10.1145/1774088.1774408
17. Oh K, Oh B-S, Toh K-A, Yau W-Y, Eng H-L (2014) Combining sclera and periocular features for multi-modal identity verification. Neurocomputing 128:185–198. doi:10.1016/j.neucom.2013.01.066
18. Ojala T, Pietikäinen M, Harwood D (1996) A comparative study of texture measures with classification based on featured distributions. IEEE Trans Inf Forensics Secur 29(1):51–59. doi:10.1016/0031-3203(95)00067-4
19. Padole CN, Proenca H (2012) Periocular recognition: analysis of performance degradation factors. In: 5Th IAPR international conference on biometrics (ICB), pp 439–445. doi:10.1109/ICB.2012.6199790
20. Park U, Ross A, Jain AK (2009) Periocular biometrics in the visible spectrum: a feasibility study. In: IEEE 3rd international conference on biometrics: theory, applications and systems (BTAS). doi:10.1109/BTAS.2009.5339068
21. Patel B, Maheshwari RP, Raman B (2017) Evaluation of periocular features for kinship verification in the wild. Comput Vis Image Underst. doi:10.1016/j.cviu.2017.04.009
22. Phillips JP, Moon H, Rizvi SA, Rauss PJ (2000) The FERET evaluation methodology for face-recognition algorithms. IEEE Trans Pattern Anal Mach Intell 22(10):1090–1104. doi:10.1109/34.879790
23. Proenca H, Alexandre LA (2005) UBIRIS: a noisy iris image database. In: 13Th international conference on image analysis and processing (ICIAP 2005), pp 970–977. doi:10.1007/11553595_119
24. Proenca H, Alexandre LA (2007) Toward non-cooperative iris recognition: a classification approach using multiple signatures. IEEE Trans Pattern Anal Mach Intell 29(4):607–612. doi:10.1109/TPAMI.2007.1016
25. Proenca H, Filipe S, Santos R, Oliveira J, Alexandre LA (2010) The UBIRIS.v2: a database of visible wavelength iris images captured on-the-move and at-a-distance. IEEE Trans Pattern Anal Mach Intell 32(8):1529–1535. doi:10.1109/TPAMI.2009.66
26. Santos G, Grancho E, Bernardo MV, Fiadeiro PT (2014) Fusing iris and periocular information for cross-sensor recognition. Pattern Recogn Lett. doi:10.1016/j.patrec.2014.09.012
27. Uzair M, Mahmood A, Mian A, McDonald C (2013) A compact discriminative representation for efficient image-set classification with application to biometric recognition. In: 6th IAPR international conference on biometrics (ICB). doi:10.1109/ICB.2013.6612959
28. Vatsa M, Singh R, Noore A (2008) Improving iris recognition performance using segmentation, quality enhancement, match score fusion, and indexing. IEEE Trans Syst Man Cybern B Cybern 38(4):1021–1035. doi:10.1109/TSMCB.2008.922059

29. Woodard DL, Pundlik SJ, Miller PE (2010) On the fusion of periocular and iris biometrics in non-ideal imagery. In: 20Th international conference on pattern recognition (ICPR), pp 201–204. doi:10.1109/ICPR.2010.58
30. Woodard DL, Pundlik SJ, Miller PE, Lyle JR (2011) Appearance - based periocular features in the context of face and non-ideal iris recognition. J Signal Image Video Process 5(4):443–455. doi:10.1007/s11760-011-0248-2
31. Zhao Z, Kumar A (2017) Accurate periocular recognition under less constrained environment using semantics-assisted convolutional neural network. IEEE Trans Inf Forensics Secur 12(5):1017–1030. doi:10.1109/TIFS.2016.2636093

Sambit Bakshi is currently with Centre for Computer Vision and Pattern Recognition of National Institute of Technology Rourkela, India. He also serves as Assistant Professor in Department of Computer Science & Engineering of the institute. He earned his PhD degree in Computer Science & Engineering in 2015. He serves as associate editor of International Journal of Biometrics, IEEE Access, and Plos One. He is technical committee member of IEEE Computer Society Technical Committee on Pattern Analysis and Machine Intelligence. He received the prestigious Innovative Student Projects Award - 2011 from Indian National Academy of Engineering (INAE) for his master's thesis. He has more than 30 publications in journals, reports, conferences.

Pankaj K. Sa earned his PhD in Computer Science from NIT Rourkela in the year 2010. He is currently an Assistant Professor in the CSE Department of NIT Rourkela. His research interest includes Computer Vision, Biometrics, Visual Surveillance. He is a member of CSI and IEEE. He has coauthored a number of research articles in various journals, conferences, and book chapters. He has co-investigated some R&D projects funded by SERB, PXE, DeitY, ISRO. He has been conferred with various prestigious awards and honors. Apart from research and teaching, he is also actively involved with the automation of NIT Rourkela, where he conceptualizes and engineers the automation process.

Haoxiang Wang is currently the director and lead executive faculty member of GoPerception Laboratory, NY, USA and has been the Cornell University Alumni since 2014. His research interests include multimedia information processing, pattern recognition and machine learning, remote sensing image processing and data-driven business intelligence. He has co-authored over 30 journal and conference papers in these areas, including IEEE SMC, ICPR, ICTAI, etc. He severed as the Guest Editor for Multimedia Tools and Applications, International Journal of Information Technology and Web Engineering and International Journal of Computer Science & Applications. He has been serving as the technical program committee member or session chair for dozens of international conferences and the reviewer for over 20 prestigious international journals.



Soubhagya Sankar Barpanda received his doctoral degree from Department of Computer Science & Engineering, National Institute of Technology Rourkela, India. He has completed his M.Tech. degree from the same institute. His research interests include biometric security and classical image processing.

Banshidhar Majhi is presently serving as Academic Dean of National Institute of Technology Rourkela, India. He is also a professor in the Department of Computer Science & Engineering there. He heads the Centre for Computer Vision and Pattern Recognition. He has more than 100 publications in reputed journals and conferences. His area of research spans from video technology to statistical image processing.