Efficient hand vein recognition using local keypoint descriptors and directional gradients

Mohammad H. Alshayeji [1] · Suood Abdulaziz Al-Roomi [1] · Sa'ed Abed [1]

Abstract

This paper proposes a computationally efficient palm and wrist vein biometric system through finely tuning computer-vision algorithms. In particular, a comprehensive analysis of the scale-invariant feature transform (SIFT) and speeded-up robust features (SURF) keypoint descriptors was conducted along with a novel idea of a score-based fusion of directional image derivatives to achieve outstanding recognition results. The work demonstrates that appropriate vein image processing, keypoint extraction, optimal matching metrics, and combination of classification scores from a group of directional gradients lead to robust and stable vein recognition. It was shown through experimental analysis that the developed biometric system outperforms all state-of-the-art results other than deep learning methods on the two public hand vein databases (VERA and PUT). Moreover, an absolute 100% recognition for the PUT palm dataset was achieved without using deep learning. The proposed method is more suitable for embedded implementation compared to deep learning algorithms, with only a slight penalty in performance compared to deep learning architectures.

1 Introduction

Personal identification and verification play an important role in our society. Valuable information, access points, and permissions are granted using passwords, magnetic cards, keys, and personal identification numbers (PINs). However, such traditional methods are limited in their convenience, reliability, and security. Passwords can be forgotten or compromised, keys and magnetic cards can be stolen or lost, and PINs could be known or guessed by

✉ Mohammad H. Alshayeji
  m.alshayeji@ku.edu.kw

[1] Computer Engineering Department, Kuwait University, Kuwait City, Kuwait

unauthorized individuals. These issues inspired the search for alternatives to traditional methods, which sparked the interest in research and development for computer-based biometric authentication systems in the 1970s [21].

Biometric authentication refers to the use of unique and measurable human physiological and behavioral characteristics to authenticate an individual. Such traits include fingerprint, face, iris, palm print, gait, voice, and hand geometry. The reason for such diverse interests is that each biometric trait has its own merits and shortcomings when it comes to satisfying different evaluation factors [38]. The work in [15] defined seven factors that can be used to evaluate a biometric trait: universality, uniqueness, permanence, measurability, performance, acceptability, and circumvention. Each biometric trait varies in its score on these factors; thus, the best one is situational.

The fingerprint as a biometric trait captured considerable attention in research and development. This can be explained because it satisfies most of Jain's biometric factors with a high degree. Still, the fingerprint as a biometric trait has its flaws. It is an external trait, which makes it susceptible to forgery. In addition, it is also vulnerable to factors that hinder its performance in the case of dirt or grease on the finger. Furthermore, some individuals do not possess a clear fingerprint due to their physical work (e.g. manual laborers), skin issues, and the clarity of the fingerprint is an issue among elderly people [10, 14, 19].

Other popular biometric traits include face and iris recognition. The former suffers greatly from bad lighting conditions and any variation in viewpoint, while the latter requires a costly acquisition device [18]. In both, the acceptability level is quite low to be used in frequent authentication situations. It is evident that there is no single biometric trait that fits all situations. In fact, the applicability of the biometric trait not only differs based on the security level, speed, and convenience required in the implemented environment and on the protected data, but it is also heavily subjected to cultural and societal principles and acceptance. For example, using face recognition as a primary biometric trait in many countries may not be acceptable. Thus, there is a clear need to research and develop techniques for obscure unique human characteristics apart from the common ones.

In this work, we focus our attention on palm and wrist vein biometrics. A review [37] gives a good overview of palm biometrics. The motivation behind such a focus is based on two main reasons. First, veins, as a biometric trait, have several advantages that make them ideal in many situations. This trait has a high collectability, acceptability, and convenience due to being a biometric trait attached to the hand and its extension (in the case of the wrist). Second, veins are internal organs, which means they are harder to notice, spoof, or damage. This makes this biometric trait a highly secure trait by design when compared with external traits, such as fingerprints and face recognition. For example, in the fingerprint, an individual may unintentionally leave a latent print of the palm that opens the possibility for malicious activities.

Deep Learning has been extensively applied to this problem with excellent results. Babalola et al. [4] use Deep Learning to achieve nearly 100% recognition accuracy in several public palm vein datasets. Jhong et al. [16] is another recent work that uses Deep Learning for palm vein recognition. They achieve good performance, but unfortunately use a private dataset. In [20], VGG Net-16 is adopted for the problem of dorsal vein pattern matching. Another important work is [24] that uses the Deep Hashing network to achieve high performance in palm vein recognition. In another work by Chen et al. [8], the authors attempt to reduce the computational complexity of deep learning based method by using Gabor filters. This method achieves 0% EER.

Considering recent works that do not use Deep Learning, Ananthi et al. [2] describe a work that uses curvelet multiresolution features to achieve high recognition performance in public

palm vein datasets. Arakala et al. [3] use a graph matching approach to biometric recognition and achieve good performance. Das et al. [11] use a bag of features approach and use SVM as a classifier. In another important work, Tome et al. [33] used a Local Binary Patterns approach to pattern matching in this problem. Cho et al. [9] attempts to use only visible spectrum images for palm vein recognition to achieve reasonably good performance.

We look at the potential of computer-vision methods for hand vein recognition. In particular, we comprehensively analyze scale-invariant feature transform (SIFT) [22] and speeded-up robust features (SURF) [23]. Here, all the important parameters for keypoint detection, description, and matching are studied. In addition, we describe some optimization steps for vein images. We also propose a novel fusion scheme based on directional image derivatives. It is shown that combining classification results at the score level achieves superior output. All the analysis and results produced in this work were applied on the palm and wrist vein images from two public vein databases, PUT and VERA. We aim to achieve good recognition performance without the help of Deep Learning and with low computational complexity.

In this work, we comprehensively analyze the potential of SIFT and SURF as feature extractors for hand vein recognition. The work achieves very good performance, one that is very close to Deep Learning-based work, but at a much lower computational cost. Due to the lower computational cost, the proposed algorithm is much more suitable for embedded standalone devices. In addition, we propose an effective score fusion technique that is based on directional image derivatives of the vein images.

The paper is organized as follows. Section 2 provides general information about biometric system components and an overview of some successful vein recognition methods proposed in the literature. Section 3 presents theoretical information about the analysed local keypoint descriptors and the developed fusion scheme. Public vein datasets, experimental analysis, and recognition results are described in Section 4. Section 5 concludes the research and provides some future work.

2 Hand vein biometrics: key steps and methods

The general structure of a biometric system (Fig. 1) contains some common components such as an acquisition device, a pre-processing module, feature extraction, and matcher parts. Database storage is needed to save samples of the enrolled individual. The saved samples are also retrieved to verify or authenticate the unknown sample.

For vein biometrics, the data are typically acquired in the near infrared (NIR) frequency band [6]. Vein images can be acquired in contact-based or contactless modes. Figure 2 illustrates an example of wrist and palm vein images. The wrist vein image from the PUT
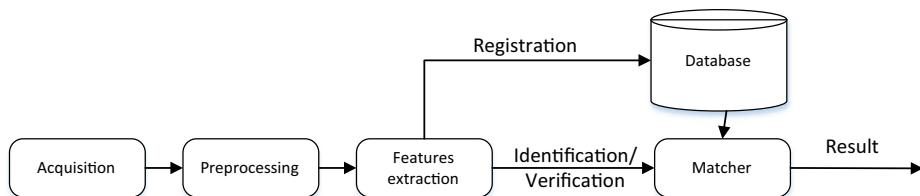

Fig. 1 General biometric system structure

database [17] is shown in Fig. 2a, and a palm vein sample from the VERA dataset [33] is given in Fig. 2b. The goal is to find the most suitable image pre-processing, vein image features, and matching scheme demonstrating the highest discriminative power and potential for accurate biometric identification.

Low contrast and registration noise usually make acquired images unsuitable for direct feature extraction. Specific processing is required. Image pre-processing includes region of interest (ROI) segmentation [18, 23], denoising, low-pass filtration, and contrast enhancement (CE).

For low-pass filtration, Gaussian smoothing is applied to the vein image. This is to enhance the low frequency components of the image. For vein biometrics, CE is the most important part of pre-processing since initial vein images often have poor contrast (see Fig. 2). Different techniques can be applied [27].

After pre-processing, feature extraction is applied. This stage is the most widely studied part of biometric systems. Here, 'features' indicate some distinctive image characteristics that are efficiently described numerically. Local keypoint descriptors were used as proposed in [34] to address the problem of vein recognition. The keypoint descriptors used in this work, SIFT and SURF are described in the next section. A feature extraction technique was applied to both known samples and the unknown samples. Following that, Euclidean distance, [35] metric was used to match the keypoint descriptors.

3 Analysis of vein image structure

This section contains theoretical information about SIFT and SURF algorithms. Peculiarities of interest such as pixel detection, descriptor construction, and vein image matching are described. Moreover, the proposed idea for score-based fusion for a sequence of image directional derivatives is revealed.

3.1 Local keypoint descriptors

Interest point or keypoint descriptors are widely used in various computer-vision problems, such as image matching, image retrieval, panorama stitching, 3D reconstruction, object detection, etc. [7]. In our case, we can consider biometric identification based on vein images as an image retrieval task. The task is to find the most similar sample in the biometric database
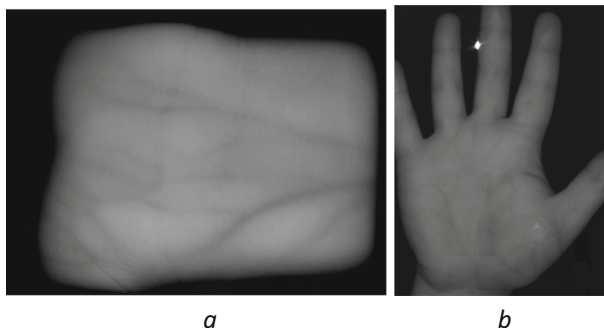

a   *b*

Fig. 2 Examples of hand vein images **a** PUT DB, left hand, wrist, **b** VERA DB, right hand, palm

for a given test vein image. There are several successful examples of SIFT and SURF applications for vein recognition. In [1, 39] SIFT was successfully applied to palm vein images from CASIA and PolyU databases, demonstrating accuracy up to 96.13% [1]. As for SURF, its advantage over general SIFT in palm vein recognition was demonstrated in [26]. In addition, in [28], SURF features were applied to finger vein recognition.

It was noticed that in all existing works there is a lack of information about algorithm parameters and optimization steps for both SIFT and SURF. In our study, the gap is filled by providing theoretical details about both computer-vision techniques and experimental analysis.

Generally, image matching based on keypoint descriptors is accomplished via several steps (Fig. 3a). First, keypoints are localized in the input image. They are pixels that demonstrate a high intensity gradient with respect to the local neighborhood. To capture the image structure in multiple scales, the scale-space concept is used [22]. A typical step is to form a set of vein image versions convolved with Gaussian functions $G(x, y, \sigma)$ (Fig. 3b). Changing the σ parameter allows the retrieval of veins of different sizes.

Keypoint detection is accomplished via analysis of the second-order image gradients (Hessian) [22]. The problem is that direct calculation of Hessian matrix elements is computationally inefficient. Thus, alternatives were proposed in both SIFT and SURF algorithms. The former method uses an approximation of the Laplacian of Gaussian (LoG) using the difference of Gaussians (DoG) [23]. The latter technique utilizes specifically designed box filters as an approximation of Gaussian derivatives. In addition, the concept of integral images is used for performance reasons [5, 12] making SURF detection faster than SIFT. The output of the detection step is both the location of and scale of the keypoints (Fig. 4). Figure 4a, b contain an example of localized keypoints for SIFT and SURF, respectively. One can see that the found keypoint scales are larger for SURF. This effect will be discussed in the experimental analysis section.

After finding the keypoint, the corresponding descriptor is constructed (Fig. 5). For SIFT, a square area around the keypoint is extracted [5, 23]. Its slope is aligned with an estimated keypoint orientation. The region is divided into 4 × 4 subregions of 3σ size, and in each, 8 bin orientation histograms are calculated. All the orientation histograms are combined into the SIFT descriptor (Fig. 5a). Its size is equal to 128 (8 bin orientation histogram for 4*4 subregions).
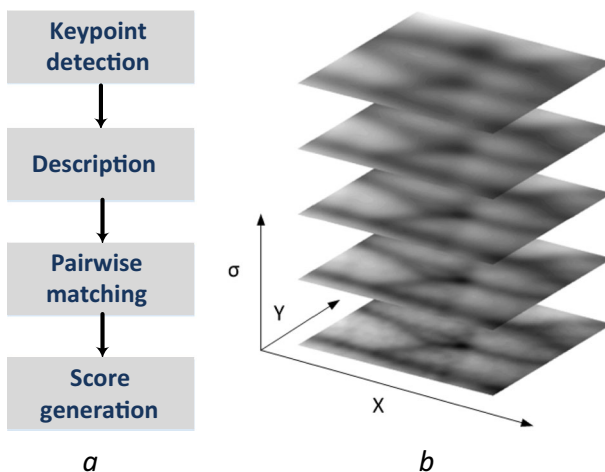

Fig. 3 Image matching steps and scale-space concept **a** image matching steps, **b** Gaussian pyramid
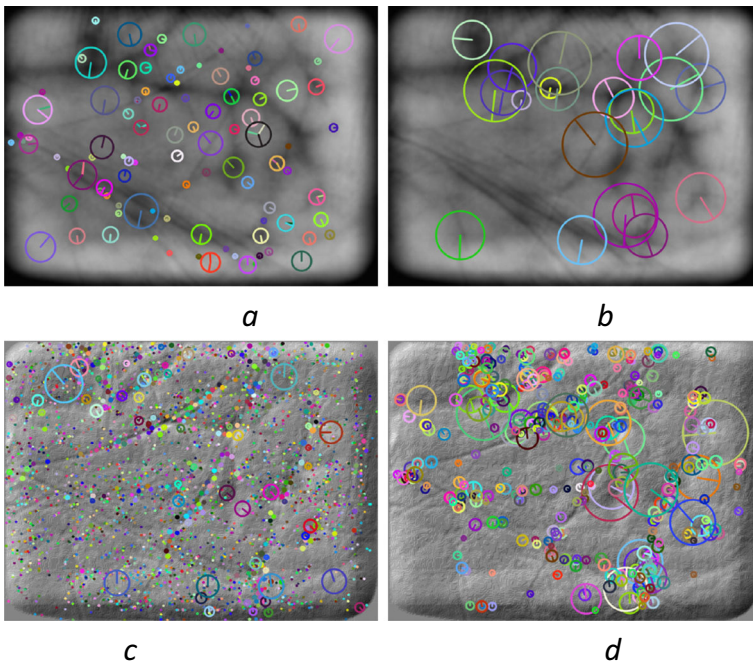
Fig. 4 Detected SURF and SIFT keypoints **a** original image (SIFT), **b** original image (SURF), **c** 135° directional derivative (SIFT), **d** 135° directional derivative (SURF)

In SURF, the extracted squared area of $6\,s$ size (where s is the keypoint scale) is divided into 16 quadrants (Fig. 5b). For a particular quadrant, 25 cells are considered as an input for the calculation of Haar filter responses. Four values are evaluated for each quadrant as a sum of Haar responses $\sum dx$, $\sum dy$, $\sum |dx|$, and $\sum |dy|$ over the quadrant (blue dots). Thus, a 64-dimensional SURF descriptor is constructed (4×4 quadrants $\times 4$ vector dimensions) [5, 29].

In order to lower the processing time, so-called octaves are used. In SIFT, the input image is resized multiple times (different octaves) instead of building multiple Gaussians with large σ values. Thus, this allows the detection to be more efficient. In SURF, the octaves concept is
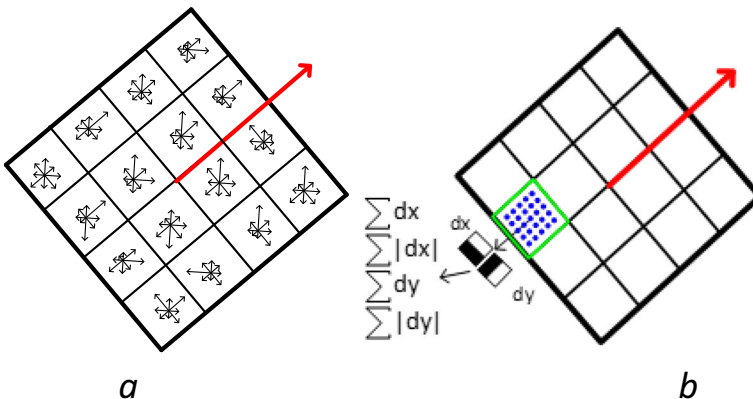

Fig. 5 Descriptor construction **a** SIFT, **b** SURF

implemented via building separate groups of box filters of different sizes. Balancing the number of scales and octaves is one of the tuning instruments for the considered keypoint descriptors. Local descriptors are formed along the estimated keypoint orientation. Details about this step can be found in [5, 23].

After constructing the keypoint in a given vein image, we match it with one or more samples (verification), or all samples (identification) in the database. Image matching scores for SIFT and SURF can be produced using several matching techniques. We evaluate the metric (score) as a sum of the best matches:

$$d = \frac{N_{best}}{N_{\min}} \sum_{i=1}^{N_{\min}} D_i,$$ (1)

where N_{best} represents the number of the most similar keypoint pairs. If the number of detected keypoints is less than N_{best}, $N_{\min}$ is introduced as a safety measure to avoid incorrectly producing a lower distance metric for the cases in which the vein image contains a small number of detected keypoints. In (1), D_i corresponds to the distance between a particular pair of descriptors. The Euclidean distance metric is used to find D_i as in [33]:

$$D_E = \sqrt{\sum_{i=1}^{n} (u_i - v_i)^2},$$ (2)

where u_i and v_i are the two compared descriptors summed over the dimensions (128 for SIFT and 64 for SURF).

Thus, while the mechanism of keypoint detection is the same for both SIFT and SURF, they use different approximations of second-order image gradients and the orientation estimation process. This provides a significant difference in keypoint location, orientation, and scale.

3.2 Image directional gradients and fusion

Textural features like LBP and LDP are often used for vein recognition [25]. In this paper, we propose to utilize textural information in a new way. The general idea is to calculate the image directional derivatives, apply SIFT and SURF descriptors for those images separately and, finally, perform a score-based fusion.

Directional derivatives can be easily evaluated as follows:

$$I_0(x,y) = I(x+1,y) - I(x,y),$$ (3a)

$$I_{45}(x,y) = I(x+1,y+1) - I(x,y),$$ (3b)

$$I_{90}(x,y) = I(x,y+1) - (x,y)$$ (3c)

$$I_{135}(x,y) = I(x-1,y+1) - I(x,y),$$ (3d)

$$I_{180}(x,y) = I(x-1,y) - I(x,y), \tag{3e}$$

$$I_{225}(x,y) = I(x-1,y-1) - I(x,y), \tag{3f}$$

$$I_{270}(x,y) = I(x,y-1) - I(x,y), \tag{3g}$$

$$I_{315}(x,y) = I(x+1,y-1) - I(x,y), \tag{3h}$$

where $I(x, y)$ is a given image pixel. Figure 6 illustrates an example of a vein image and its constructed directional derivatives. Notice that each gradient retrieves some specific features of the vein image, which are invisible in the original image. Comparing detected keypoints (SIFT and SURF) in the original vein image (Fig. 4a, b) and in the calculated directional derivative (3d; Fig. 4c, d) shows that the number of retrieved small-scale details is higher. In addition, the scales of the keypoints for the directional gradient are much smaller. The diversity in the size and location of the keypoints in each direction allows for additional recognition power when combining them with the original vein image keypoints.

Figure 7 shows the structure of our proposed multimodal biometric system. First, we evaluate the direction derivatives for the test sample using (3a)-(3 h). After that, we extract the feature (keypoints) using either SIFT or SURF. For each vein image in the database, we obtain matching scores with the test sample using the scheme described in (1). Scores are calculated for the original image and its directional derivatives. As a result, nine predicted class labels are obtained: one for the original vein image and eight derivatives. Note, the calculations related to different directional derivatives are independent, so these calculations may be performed in parallel. The last step in the biometric system is combining all the produced
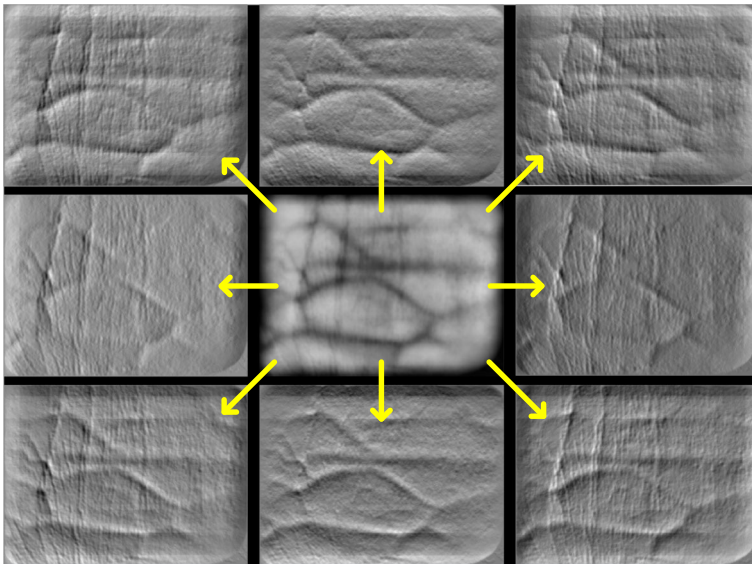

Fig. 6 The original image and its 8 directional gradients

scores using a majority voting scheme. The leading class (individual) will be selected as the final prediction. Majority voting was found effective as even though certain derivatives are more effective than others, since they are identifying the same person, the majority agrees on the class. This fusion method was used as it was the simplest approach to fusion. If there is no major class label, the class with the lowest classification distance is selected from predicted classes.

4 Experimental analysis

Significant research in this field has been gated behind private databases where the results of different biometric systems are incomparable [5, 12, 26, 30, 35, 36]. Each private database captures the vein at a specific angle, with different orientations, lighting conditions, environments, and hardware. Moreover, the produced raw images are processed in variant sizes and resolutions. What is even more challenging is the sample size and diversity of the samples being tested are vastly different from one to another. All these configurations make comparing biometric systems of different private databases almost impossible.

Fortunately, the recent emergence of public vein databases has allowed researchers of this field to evaluate and compare their developed systems on common ground. In addition, evaluating the performance of the developed work in multiple public databases can highlight the stability and resilience of the biometric work across different sets of environments.
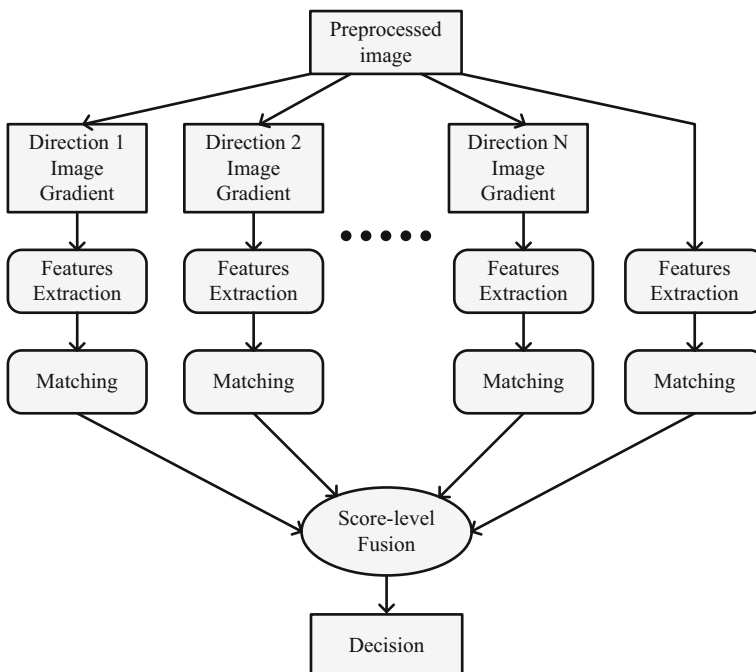

Fig. 7 Score-fusion scheme

4.1 Public hand vein databases

An example of an accessible public vein database is the PUT database [23]. It is a contact-based database that has 2400 images, of which 1200 are of palms and 1200 are of wrists. All the images were captured using NIR imaging with a wavelength of 880 nm from 50 individuals. Images were obtained in three sessions with four images in each session, and the timeline between each session is at least one week. One of the advantages of this database is the availability of both the palm and wrist vein images of each participant. This allows the researchers to evaluate the developed biometric systems on both parts in isolation and to have the ability to fuse them together.

Another contactless database is called VERA palm vein [33]. The dataset was created in the Idiap research institute with a wavelength of 940 nm. The acquisition for this database was performed in two different indoor locations, both having normal lighting conditions. The database captured 1000 images from 50 individuals. The actual number of participants was 110, but only 50 are currently available, with 64% of the participants being male and 36% female. The age of the participants ranged from 18 to 60 with a mean age of 33. Two sessions were conducted to acquire five palm images per side for a total of 10 images per participant.

4.2 Vein recognition experiments

In this research, three sets of images are analyzed: PUT wrist, PUT palm, and VERA palm images. In each classification experiment, 75% of samples for each person were considered as the biometric system database, and the remaining images were used as the test set. For the PUT database this means 3 base samples and 1 test sample for each person: for the Vera base they are 4 and 1 sample. To provide reliable results and a better simulation of real scenarios, 10 independent experiments were conducted. In each experiment, a random division to the database and the test samples was applied, and the accuracy was determined. The final accuracy of the biometric system is estimated by averaging the accuracy results of the 10 random selections (experiments). Such a procedure eliminates the effect of the random sample and makes the final results more deterministic.

Initially, different options for vein image pre-processing were analyzed. Noise reduction and CE are essential steps for vein image analysis. The image contrast and noise level are especially crucial for local keypoint descriptors (e.g. SIFT and SURF). Finding interest points even on small vessels may significantly increase the classification performance. To enhance the contrast, we use CLAHE. It outperforms other common techniques [34] and was applied to obtain state-of-the-art results on PUT images [22, 35]. Its application increased the recognition accuracy from about 50% (with ROI segmentation as the only pre-processing step) to 90%.

Noise suppression is also an important pre-processing step since high-frequency intensity fluctuations may cause the detection of spurious keypoints and lead to false matching of descriptors. In particular, in the case of SURF, the large utilized box filters are very sensitive to high-frequency noise [29]. Several options were examined for noise reduction. In particular, Gaussian smoothing, one of the simplest examples of low-pass filtration, was used. In addition, we tested Wiener, median, and non-local mean filtering. A comparative analysis indicated that the Gaussian filter is the best pre-processing option. Removing high-frequency noise with any of the methods additionally increases the recognition accuracy to up to 96% for both SIFT and SURF. Optimal pre-processing parameters were determined for PUT and VERA separately.

To compare two sets of keypoints, the scheme (2) was applied with different distance metrics: Euclidean distance, its normalised form [1, 32], Minkowski distance [31], city block (or Manhattan) [1, 32] and Hausdorff distance [13]. The common Euclidean metric was found to be the best for SIFT descriptor matching, while normalized Euclidean distance demonstrated the highest accuracy for the SURF algorithm.

For feature extraction, an important initial parameter is the keypoint detection threshold. This parameter regulates the detectability mechanism in the input image. It determines the minimum brightness change required for treating a particular pixel as an interest point. Different decision mechanisms for detection are used in SIFT and SURF. The former method utilizes the contrast threshold, while the latter is calculated based on Hessian matrix elements directly [5]. Figure 8 illustrates the detected keypoints in a vein image as a function of SIFT (Fig. 8a) and SURF (Fig. 8b) thresholds. The higher the value of the threshold (both SIFT and SURF), the fewer keypoints are extracted. In the case of a high value of the threshold, we can expect more stable and distinctive keypoints to be found; however, the number of keypoints should be properly controlled. Thus, there is a trade-off between the amount of captured information and the robustness of keypoint matching. The optimal threshold level depends on the vein image characteristics. Different optimal threshold values for SIFT and SURF thresholds were found. In the case of SIFT, it was 0.01 for PUT wrist, 0.007 for PUT palm, and 0.003 for VERA palm images. For SURF, the corresponding Hessian threshold was equal to 150 for PUT (palm and wrist) and 50 for VERA images.

In any scale-space analysis, we have two important variables that we can vary: the number of octaves and its scale, which represents the number of layers within a single octave. Figure 9 contains the dependence of the recognition accuracy for different numbers of scales.

For SIFT, a smaller number of scales gives higher accuracy, while for SURF, it is the opposite. This is due to the difference in key point selection method of SIFT and SURF. With an increase in scale, the number of key points in SURF increases. This increases the number of features extracted from the image. This improves the accuracy of our method since the accuracy depends on the number of features extracted. However, for SIFT, the number of key points selected decreases with an increase in scale. Hence, the number of features extracted from the image decreases with an increase in scale which in turn decreases the accuracy.
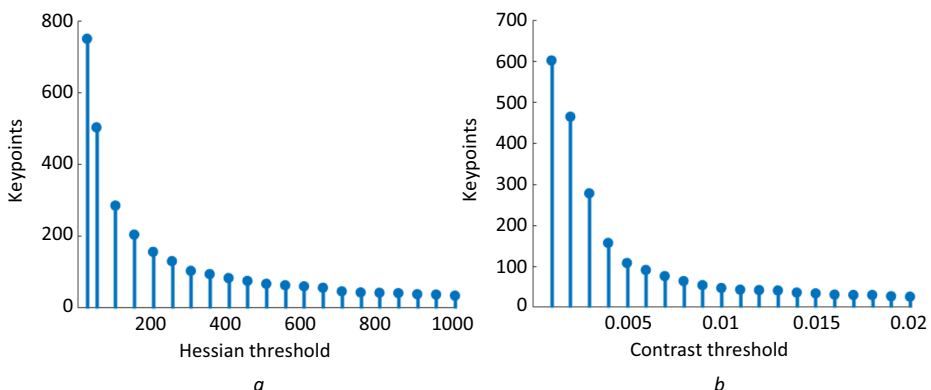

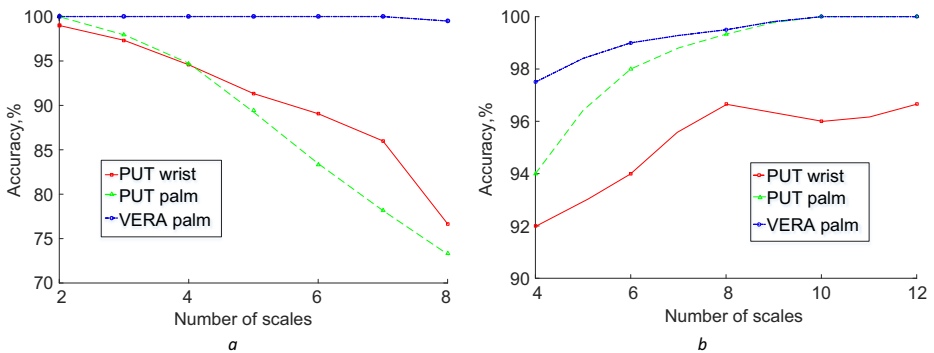Fig. 8 Number of keypoints vs threshold value **a** SURF **b** SIFT

Fig. 9 Recognition accuracy using SIFT and SURF **a** SIFT, **b** SURF

Diminishing returns are reached early for octave values. We found that, for all three datasets and for both SIFT and SURF, using four octaves provides better results than two octaves, but further increases in the number of octaves does not change the accuracy. This means that keypoints with a scale higher than a certain threshold (which is determined by vein pattern characteristics) do not bring additional information on vein structure to improve accuracy. Thus, the optimal values in our case are three octaves and four scales for SIFT and four octaves and 12 scales for SURF.

Putting it all together, we depict the accuracy of our proposed biometric system with and without the fusion of the directional derivatives. First, vein recognition experiments have been conducted with different directional gradient images as input. Since the structure of every directional gradient image contains more textures and small-scale details, the optimal threshold values have been found separately. As for the scales and octaves, the observed effects were similar to the original vein images for both SIFT and SURF. Looking at Fig. 4c, d, large keypoints still can be noticed, but small keypoints dominate.

For the above parameters, recognition results for all three datasets have been obtained (Fig. 10). In general, SIFT performs as well as SURF for PUT datasets. As for VERA images, SIFT results regarding different directions appear more uniform than SURF. The results indicate an improved classification accuracy for directional gradient images with respect to the input vein image. In particular, for SIFT a 225° derivative outperforms the original image for the PUT wrist, while, for the PUT palm, both 90° and 270° provide superior results. As for VERA classification using SIFT features, both directional derivatives and original images provide very good results. In the case of SURF, the obtained diagram reveals that the gradient directions $45° + 90°k$ are more efficient than others.

Finally, the scores of the conducted experiments with various directional gradients are fused together to generate the final recognition results. This fusion is done by majority voting. That is, the class with the highest number of matches out of all directional gradients was selected as the best match. The average accuracy for 10 randomly selected samples was found to be very high in the range of 98.07%–100%. In particular, for the PUT palm, all 10 random selections gave an absolute 100% accuracy for both SIFT and SURF (for the fusion scheme). For both SIFT and SURF in the VERA database, the majority of the random selection experiments also gave 100% accuracy.

Apart from accuracy, we measure the performance of our algorithms in terms of the following metrics.
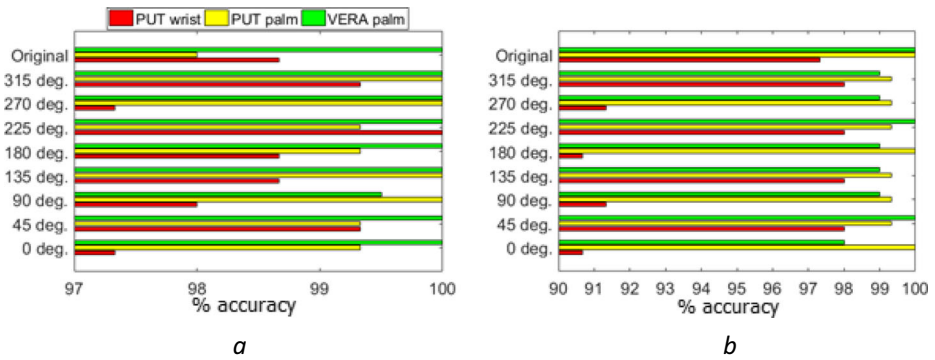
Fig. 10 Accuracy using directional derivatives **a** with SIFT, **b** with SURF

False Rejection Ratio (FRR) is given by

$$FRR = \text{Number of genuine identification claims rejected/Number of genuine identification requests}$$

False Acceptance Ratio FAR is given by

$$FAR = \text{Number of false identification claims accepted/Number of false identification claims}$$

Equal Error Rate (EER) is the operating point at which FRR and FAR are equal

Figure 11 demonstrates the metrics (FAR, FRR and ERR) for these combinations. For the PUT palm dataset all these metrics are equal to zero.

In this section, we found optimal combinations of directional derivatives for each real vein dataset. In each case, it includes an original image and two derivatives. Angles (in degrees) of these derivatives are given in Table 1. The optimal combinations were found by independently fusing all possible pairs of results from directional derivatives and the original image and picking the pair that gives the best results.

As expected, different combinations were found optimal for SIFT and SURF since the mechanisms of keypoint detection and description have some intrinsic peculiarities. In addition, the difference in optimal values of SIFT and SURF parameters also existed between the two databases due to different environments and capturing modes (contact and contactless).
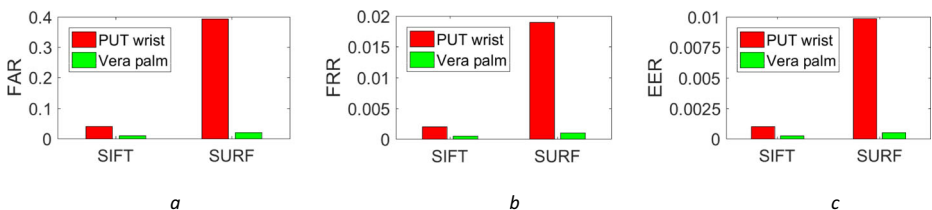

Fig. 11 Performance metrics of the biometric system for the best fusion combinations: **a** FAR, **b** FRR, **c** EER

Table 1 Optimal combinations of directional derivatives

Database	SIFT	SURF
PUT palm	45, 225	0, 45
PUT wrist	90, 270	0, 180
Vera	0, 180	0, 45

Furthermore, these values are not the same between the two experimental biometric traits (palm and wrist veins).

It is also important to note that we generally found that the classification of palm veins is much more stable than that of wrist veins since the palm images contain small contrast details from the palm print, which was evident when comparing the datasets of PUT.

The running time of the program for training and testing while running it on an Intel Xeon four core 3.7 GHz computer with 32GB RAM is given in Table 2. As shown in the table, SURF has a significantly lower running time due to faster feature extraction and the smaller 64-dimensional features.

Since SURF had faster running time and gave good results, we investigated it further to explore the relationship between its running time and accuracy The plot of running time vs accuracy for SURF using a single directional gradient is shown in Fig. 12. This plot was obtained by varying the threshold of SURF from 400 to 100. The number of features extracted from the image by SURF algorithm depends on the given threshold. It should be noted here that most of the testing time was spent in comparing the training features with testing features. Hence a low number of extracted features resulted in low running time, but low accuracy. A threshold of 400 yielded a low number of SURF features and hence the running time was low and accuracy was also low. On the other hand, a threshold of 100 gave 100% accuracy, but with a high running time.

We also plotted the Cumulative Matching Characteristics (CMC) curve for SURF using Vera dataset as shown in Fig. 13. CMC curve is plotted by finding the recognition rate for each rank k by assuming that while matching a sample image with all training classes, if the correct class is above rank k, it is recognized correctly. The figure is plotted by processing the 0° directional gradient image using SURF features. As seen in the figure, 98% of the time, the correct class appears at the first place. Another 1% of the time, the correct class appears in second place. 100% of the time, the correct class appears within rank 20. The percentage recognition rates of the first 5 ranks are given in Table 3.

4.3 Comparison with state-of-the-art results

As mentioned earlier, the recent emergence of public databases for wrist and palm veins has allowed researchers to compare their work on a common ground. Recent studies in hand vein literature have used these databases and showcased the results of their developed system.

Table 2 Running time of the program (using Vera dataset and using a fusion of all directional derivatives and original image)

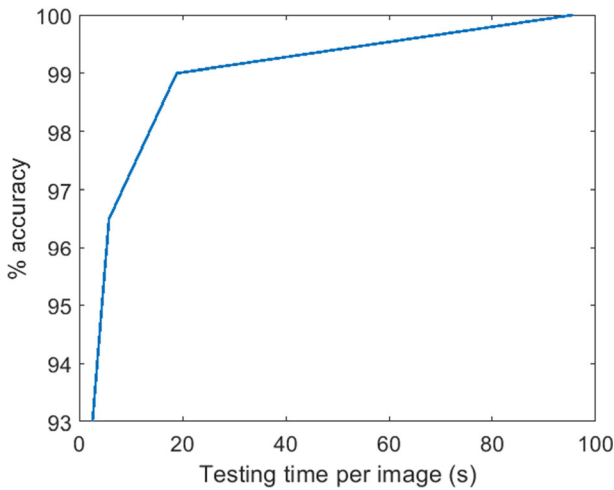| Database	SIFT	SURF
Training (seconds per image)	16.26	0.76
Testing (seconds per image)	50.02	27.23

Fig. 12 Plot of running time vs accuracy for SURF using Vera dataset with 45° directional gradient image

The first reported work for the PUT database came from the original authors of the collected data [23]. The authors segmented the vein vessels with local thresholding and used 2D correlation for matching. They also rotated the test image between −15° and 15° with a 1° step to reduce the influence of rotation between the images. Based on their experiments, the highest accuracy was found to be 96.2%.

A recent study also achieved high results for vein recognition in [33]. Here, they utilized a graph-based representation of palm vein templates of the PUT database. They proposed a novel graph trimming algorithm before matching them. The best accuracy result was found to be 99.01%, which was obtained using the normalized vertex count similarity score for the right palms of the PUT database.

It should be noted, that in [39] 100% accuracy was achieved for PUT wrist images. But the authors used images of both hands, providing score-level fusion. In our case, an image of only one hand is required.
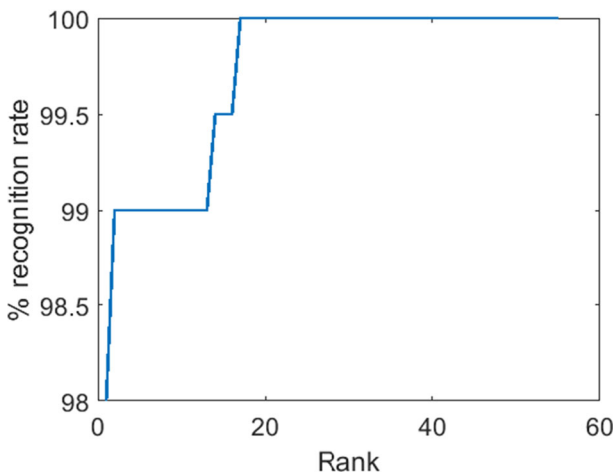

Fig. 13 CMC curve of SURF with 0° directional gradient. 98% of the time, the correct class appears in rank 1

Table 3 % Recognition rate of images with rank 1 to 5

Rank	% Recognition rate
1	98
2	99
3	99
4	99
5	99

The contactless VERA database from Idiap is a very recent database. The original authors of VERA also worked in designing the vein biometric system shown in [33]. In their work, the circular Gabor filter was applied on the vein vessels to enhance the images, and LBP was utilised as a feature of the veins. The extracted features were matched using histogram intersection, and the best reported result was 96.25% using both left and right hands.

Table 4 compares our results with the state-of-the-art work presented in literature for both the PUT palm and wrist VERA databases. We included the best results for SIFT and SURF only and for the developed fusion scheme. The comparison is specifically with works that do not use deep learning. The proposed methodology outperforms all three state-of-the-art results. In particular, both SIFT and SURF with directional derivatives provide absolute recognition accuracy for PUT palm veins. As for the PUT wrist dataset, SIFT with the fusion scheme outperforms the best existing solution, providing 99.8% accuracy. Finally, for the VERA database, both optimized feature descriptors and corresponding directional derivatives provided the above state-of-the-art results. In this case, SIFT provided 99.95% accuracy.

It is important to recall that, in our work, we displayed the stability of our classifier through randomizing the samples with subsequent result averaging. Such a procedure was not conducted in the compared work.

Table 4 Comparison with state-of-the-art results. Comparison is specifically with works that do not use deep learning

Technique	Accuracy
PUT palm	
SIFT with directional gradients fusion	**100**
SURF with directional gradients fusion	**100**
SURF	99.93
SIFT	99.4
A. Arakala et al., 2016	99.01
R. Kabacinski, M. Kowalski, 2011	96.2
PUT wrist	
SIFT with directional gradients fusion	**99.8**
A. Das et al., 2014	99.21
SURF with directional gradients fusion	98.07
SIFT	97.8
SURF	97.8
VERA	
SIFT with directional gradients fusion	**99.95**
SIFT	**99.95**
SURF with directional gradients fusion	99.9
Ananthi et al. 2021	99.86
SURF	99.7
P. Tome, S. Marcel, 2015	96.25

5 Declarations

The authors have no relevant financial or non-financial interests to disclose. Moreover, the authors have no competing interests to declare that are relevant to the content of this article.

6 Conclusion

In this paper, the application of local keypoint descriptors for the hand vein recognition problem is studied. Two of the most common and powerful algorithms, SIFT and SURF, were analyzed. The parameters optimization for the vein images is described. We show that, by carefully adapting optimal values for SURF and SIFT descriptors that differ based on the environment of the enrolled image, modes of capturing, and type of biometric traits, we can gain high classification performance for veins in a biometric system. The performance of SURF was found to be better than or equal to SIFT in most cases with a significantly lower running time.

A novel idea of using directional derivatives in addition to the original image is proposed. We show that different characteristics of an image and its derivatives provide additional discriminative power to the biometric recognition system. The proposed fusion scheme provides further enhancement of the performance, taking it to absolute accuracy in some cases. Such a fusion can be done in parallel, as each direction is independent of the other in its processing.

The experimental tests of the proposed biometric system showed that both fused and non-fused classifications of the palm and wrist vein show superior performance compared to previous work in the literature, and in the case of PUT palm veins, 100% accuracy was achieved.

References

1. Ahmed MA, El-Horbaty E-SM, Salem A-BM (2015) Intelligent techniques for matching palm vein images. Egyptian Computer Science Journal 39:1–14
2. Ananthi G, Sekar JR, Arivazhagan S (2021) Human palm vein authentication using curvelet multiresolution features and score level fusion. Vis Comput:1–14
3. Arakala A, Hao H, Davis S, Horadam K (2015) The palm vein graph - feature extraction and matching. In: International Conference on Information Systems Security and Privacy. SCITEPRESS - Science and and Technology Publications, pp 56–64
4. Babalola FO, Bitirim Y, Toygar Ö (2021) Palm vein recognition through fusion of texture-based and CNN-based methods. SIViP 15:459–466
5. Bay H, Tuytelaars T, Van Gool L (2006) SURF: Speeded Up Robust Features. In: SURF: speeded up robust features. Springer, Berlin, Heidelberg, pp 404–417
6. Bouzida N, Bendada AH, Maldague XP (2010) Near-infrared image formation and processing for the extraction of hand veins. J Mod Opt 57:1731–1737. https://doi.org/10.1080/09500341003725763
7. Bradski G, Kaehler A, Cambridge B·, et al (2008) Learning OpenCV. O'Reilly Media Inc.
8. Chen Y-Y, Hsia C-H, Chen P-H (2021) Contactless multispectral palm-vein recognition with lightweight convolutional neural network. IEEE Access 9:149796–149806
9. Cho S, Oh BS, Kim D, Toh KA (2021) Palm-vein verification using images from the visible Spectrum. IEEE Access
10. Coetzee L, Botha EC (1993) Fingerprint recognition in low quality images. Pattern Recogn 26:1441–1460. https://doi.org/10.1016/0031-3203(93)90151-L

11. Das A, Pal U, Ballester MAF, Blumenstein M (2014) A new wrist vein biometric system. In: 2014 IEEE symposium on computational intelligence in biometrics and identity management (CIBIM). IEEE, pp 68–75
12. Du G, Su F, Cai A (2009) Face recognition using SURF features. In: Ding M, Bhanu B, Wahl FM, Roberts J (eds) SPIE 7496, MIPPR 2009: pattern recognition and computer vision, 749628. International Society for Optics and Photonics, pp 749628–749634
13. Hartung D, Olsen MA, Xu H, Busch C (2011) Spectral minutiae for vein pattern recognition. In: 2011 international joint conference on biometrics (IJCB). IEEE, pp 1–7
14. Jain A, Lin H, Bolle R (1997) On-line fingerprint verification. IEEE Trans Pattern Anal Mach Intell 19:302–314. https://doi.org/10.1109/34.587996
15. Jain AK, Bolle R, Pankanti S (2006) Biometrics : personal identification in networked society. Springer Science & Business Media
16. Jhong S-Y, Tseng P-Y, Siriphockpirom N, Hsia C-H, Huang M-S, Hua K-L, Chen Y-Y (2020) An automated biometric identification system using CNN-based palm vein recognition. In: 2020 international conference on advanced robotics and intelligent systems (ARIS). IEEE, pp 1–6
17. Kabaciński R, Kowalski M (2011) Vein pattern database and benchmark results. Electron Lett 47:1127–1128. https://doi.org/10.1049/el.2011.1441
18. Kang W, Wu Q (2014) Contactless palm vein recognition using a mutual foreground-based local binary pattern. IEEE Transactions on Information Forensics and Security 9:1974–1985. https://doi.org/10.1109/TIFS.2014.2361020
19. Kong WK, Zhang D, Li W (2003) Palmprint feature extraction using 2-D Gabor filters. Pattern Recogn 36:2339–2347. https://doi.org/10.1016/S0031-3203(03)00121-3
20. Kumar R, Singh RC, Kant S (2021) Dorsal hand vein-biometric recognition using convolution neural network. In: International Conference on Innovative Computing and Communications. Springer, pp 1087–1107
21. Li W, Zhang D, Xu Z (2002) Palmprint identification by fourier transform. Int J Pattern Recognit Artif Intell 16:417–432. https://doi.org/10.1142/S0218001402001757
22. Lindeberg T (1994) Scale-space theory: a basic tool for analyzing structures at different scales. J Appl Stat 21:225–270. https://doi.org/10.1080/757582976
23. Lowe DG (2004) Distinctive image features from scale-invariant Keypoints. Int J Comput Vis 60:91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94
24. Ma X, Jing X, Huang H, Cui Y, Mu J (2017) Palm vein recognition scheme based on an adaptive Gabor filter. IET Biometrics 6:325–333
25. Mirmohamadsadeghi L, Drygajlo A (2014) Palm vein recognition with local texture patterns. IET Biometrics 3:198–206. https://doi.org/10.1049/iet-bmt.2013.0041
26. Nivas S, Prakash P (2014) Real-time finger-vein recognition system. International Journal of Engineering Research and General Science 2:580–591
27. Olsen MA, Hartung D, Busch C, Larsen R (2010) Contrast enhancement and metrics for biometric vein pattern recognition. In: International conference on intelligent computing. Springer, Berlin, Heidelberg, pp 425–434
28. Pan M, Kang W (2011) Palm vein recognition based on three local invariant feature extraction algorithms. In: Chinese conference on biometric recognition. Springer, Berlin, Heidelberg, pp 116–124
29. Pedersen JT (2011) Study group SURF: feature detection and description. Department of Computer Science, Aarhus University
30. Raghavendra R, Surbiryala J, Busch C (2015) Hand dorsal vein recognition: sensor, algorithms and evaluation. In: 2015 IEEE international conference on imaging systems and techniques (IST). IEEE, pp 1–6
31. Sangeetha NM, Kumar TA, Natarajan DM (2014) Feature level fusion of WLBP and HOG for hand dorsal vein recognition. International Journal of Emerging Technology and Advanced Engineering 4:9–17
32. Shark L-K, Zhang K, Wang Y (2014) Personal identification based on multiple keypoint sets of dorsal hand vein images. IET Biometrics 3:234–245. https://doi.org/10.1049/iet-bmt.2013.0042
33. Tome P, Marcel S (2015) Palm vein database and experimental framework for reproducible research. In: 2015 international conference of the biometrics special interest group (BIOSIG). IEEE, pp 1–7
34. Wang G, Wang J (2017) SIFT based vein recognition models: analysis and improvement. Computational and Mathematical Methods in Medicine 2017:1–14. https://doi.org/10.1155/2017/2373818
35. Wang K, Zhang Y, Yuan Z, Zhuang D (2006) Hand vein recognition based on multi supplemental features of multi-classifier fusion decision. In: 2006 international conference on mechatronics and automation. IEEE, pp 1790–1795
36. Wang Y, Li K, Shark L, Varley MR (2011) Hand-Dorsa Vein Recognition Based on Coded and Weighted Partition Local Binary Patterns. In: 2011 International Conference on Hand-Based Biometrics. IEEE, pp 1–5

37. Wu W, Elliott SJ, Lin S, Sun S, Tang Y (2019) Review of palm vein recognition. IET Biometrics 9:1–10
38. Xueyan L, Shuxu G (2008) The fourth biometric - vein recognition. In: Pattern recognition techniques, technology and applications. InTech, pp 537–546
39. Zhou Y, Kumar A (2011) Human identification using palm-vein images. IEEE Transactions on Information Forensics and Security 6:1259–1274. https://doi.org/10.1109/TIFS.2011.2158423