SP-FVR: SuperPoint-Based Finger Vein Recognition

Xianjing Meng[✉] and Shuai Yuan

School of Computer Science and Technology, Shandong University of Finance
and Economics, Jinan 250014, People's Republic of China
xianjing.meng@sdufe.edu.cn

Abstract. Among existing finger vein recognition methods, the keypoint-based method has gained much attention owning to its deformation-tolerance. However, the adopted keypoint descriptors may be insufficient in keypoint quantity and descriptor distinctiveness, since they are generally artificially-designed and sensitive to the poor quality of the finger vein images. To address the above-mentioned problem, an automatically-learned keypoint-based method utilizing fully convolutional neural networks (FCNNs), called SuperPoint-based finger vein recognition (SP-FVR), is proposed in this paper. We first adjust the finger vein images by pertinent intensity correction, noise removal and resizing. Then, we locate and learn the keypoint descriptors utilizing the SuperPoint model. Finally, the keypoints are matched by a bilateral strategy. Experiments on the HKPU and SDU-MLA databases demonstrate the effectiveness of the proposed method in real applications, the EERs are 0.0025 and 0.0138, with the FRRs-at-0-FAR of 0.0218 and 0.0890, respectively.

Keywords: Finger vein recognition · Keypoint · SuperPoint · Biometrics

1 Introduction

Finger vein recognition is competitive among the prevalent biometrics owning to its security and convenience, therefore has been widely accepted and increasingly adopted in commercial applications [1]. However, its performance is still limited by the quality and deformation problems [2] brought by the non-contact capturing under Near Infrared Rays (NIRs) and non-rigidity of the fingers. Extensive research works [3] have been proposed to deal with the problems, which can be roughly categorized as local pattern-based method, Region of Interest (ROI) -based method, keypoint-based method and vessel-based method.

Among the existing methods, this paper pays attention to the keypoint-based method, since it is comparatively advantageous in many aspects [4]. The local pattern-based method is mainly based on pixel-level features, such as Local Binary Pattern (LBP), Local Line Binary Pattern (LLBP) and many high-order

or learning-based variants [5]. The local patterns are fine in granularity and dense, however, are relatively low in determinability and sensitive to deformations. The ROI-based method mainly refers to recognition of whole finger ROIs based on machine learning algorithms. Typical methods are recognition based on Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA) and Convolutional Neural Networks (CNNs) [6]. This kind of method can extract features automatically, but is sensitive to image quality and deformations. Moreover, additional images are required to train the transformation matrix, which are not practical is real applications. The vessel-based method is based on the vasculature structures spread in the images, thus the vessel segmentation is prerequisite. Due to the low contrast and obscure image quality, the vessels segmented may be problematic and affect the following recognition [7]. Hence, the vessel segmentation methods are stressed and researched attentively by the researchers. However, we think the vessel-based method still needs further improvements [8]. The keypoint-based method is mainly based on descriptors extracted from structurally meaningful or geometrically explicable positions, such as cross or end points of vessels and extreme points. Comparing to other methods, the advantages of the keypoint-based method can be summarized as follows: (a) High distinctiveness, the descriptors extracted in keypoint-based method are large in scope and relatively high in discriminability. (b) Deformation-tolerance, the descriptors are usually tolerant to rotation and translations, etc., which are harmful to other methods. (c) Coarse granularity and small templates, the keypoints are localized in typical positions, thus are coarse and small in templates, which can save storage space and matching time. (d) Avoidance of vessel segmentation, keypoints can be automatically extracted, thus avoid the error propagation of the vessel segmentation. (e) Relevance to vessel structures, the distribution of the keypoints is in accordance with the vessel topologies, hence their matching results can be further refined in the guidance of the position relationships.

However, the keypoint-based method is not thoroughly explored, which may be caused by the following reasons: (a) affected by the image quality, the quantity and stability of the keypoint descriptors are always questionable. (b) the descriptors are generally artificially-designed, which are not pertinent in representativeness. (c) the similarity of main vessel can bring inter-class resemblances, which may introduce false pairings. To deal with the above-mentioned problems, we learn the keypoints using Fully Convolutional Neural Networks (FCNNs). A conventional framework called SuperPoint is introduced to finger vein recognition, which automatically learn the keypoint descriptors to improve intra-class similarities and suppress inter-class resemblances.

The proposed method is composed of three main stages. First, the finger vein images is pre-processed by illumination inhomogeneity correction, bias removal and scale adjusting. Then, the keypoints are localized and described by the FCNNs-based SuperPoint model. Finally, the descriptors are matched with a bilateral matching strategy. Extensive experiments have been conducted on the HKPU and SDU-MLA databases, the EERs are 0.0025 and 0.0138, respectively, which demonstrate the effectiveness of the proposed method. We also provide

the FRRs-at-0-FAR, which are usually adopted in commercial scenarios. The values are 0.0218 and 0.0890, respectively, which also present the applicability of the proposed method in real applications.

Our method have the following contributions. First, a finger vein recognition framework based on automatically learned keypoints is designed, which achieves high performance and is applicable in high security applications. Second, the learning strategy is introduced to the keypoint-based finger vein recognition, which is novel and provides new directions in improving finger recognition performance. Finally, effective pre-processing strategies are introduced to deal with the influences brought by the image quality and deformations, etc.

The remaining of this paper is organized as follows. Section 2 presents the proposed method. Section 3 demonstrates the experimental results and its effectiveness. This paper concludes with a discussion of future research consideration in Sect. 4.

2   Proposed Method

The proposed finger vein recognition is a systematic framework, which consists of three main stages including pre-processing, feature extraction, matching and false pairing removal. In this part, each stage is elaborated to make our method reproducible.

2.1   Pre-processing

The finger vein images suffer from severe image quality problems, such as low intensity contrast and blurry vessel margin, etc, which can be seen in Fig. 1a. These quality problems make the finger vein images different from the natural ones and limit the recognition performance. To improve the image quality, illumination inhomogeneity correction, noise removal and scale adjusting are applied.

In this paper, an image I is defined as Eq. (1):

$$I = b \cdot Img + n \tag{1}$$

here, Img represents the original 'pure' image. b denotes the additive or multiplicative noise, and n is defined as the random noise. According to this model, we take b as additive model and estimate it using a Gaussian-like convolution with the template t of $ones(11, 11)$ [9] (Fig. 1b), thus the image after bias removal can be defined as Eq. (2) (Fig. 1c):

$$rimg = I - conv2(I, t) \tag{2}$$

then, the noise is restrained by the average filtering of template size 3×3. Finally, the image is enhanced by histogram equalization resized to 192×128 by nearest interpolation, which can be seen in Fig. 1(d). Though the pre-processing strategy seems artificial, it is highly effective in promoting the performance in the proposed keypoint-based recognition. The reason might be the acquiring of the perfect images with distinct, smooth and precise location edges. The effectiveness is graphically illustrated in Fig. 2.
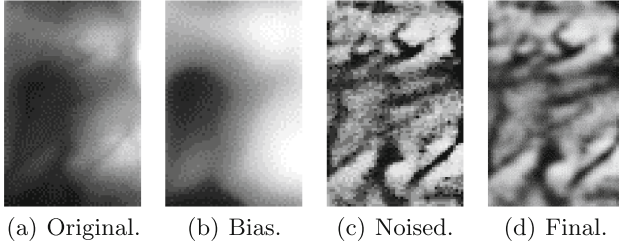
(a) Original.   (b) Bias.   (c) Noised.   (d) Final.

Fig. 1. Demo images in the pre-processing procedure. (a) The ROI extracted from original finger vein iamges; (b) The estimated bias; (c) Image after bias removal; (d) The final image after pre-processing.
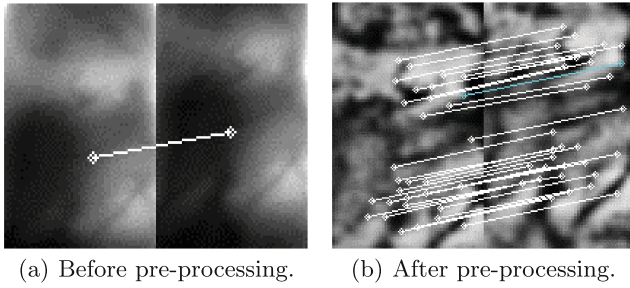

(a) Before pre-processing.   (b) After pre-processing.

Fig. 2. Matching comparison before and after pre-processing. There is only 1 keypoint matching pair in the genuine matching (a), while the number is boosted to 31 in the same matching after pre-processing.

2.2   Keypoint Extraction and Description

In order to provide robust features for the finger vein image description, we involve the automatically learned keypoint extracted by Superpoint network [10], which demonstrates good performance for integrating robust structural information during keypoint detection. The Superpoint model is a FCNN framework that consists of keypoint localization and description, which can be seen in Fig. 3.

In the whole framework, the single encoder layer is shared to reduce the dimensionality of the image, which include a series of convolutional layers, max-pooling layers, rectified linear units (ReLUs) for activation and batch normalization (BN) to get a informative feature map. Through three max-pooling layers, the image with size $W \times H$ is transformed to image sized $W_c \times H_c$ with $W_c = W/8, H_c = H/8$. Hence, the encoder maps the input image I to an intermediate tensor with smaller spatial dimension and greater channel depth.

The decoders are keypoint decoder for keypoint localization and descriptor decoder for keypoint depiction. The keypoint decoder transform the input of size $W_c \times H_c \times 65$, which corresponds to local, non-overlapping 8×8 grid regions of pixels and an extra "pointness" channel. After a channel-wise softmax, the grid regions are reshaped to size of $W \times H$. The descriptor decoder computes
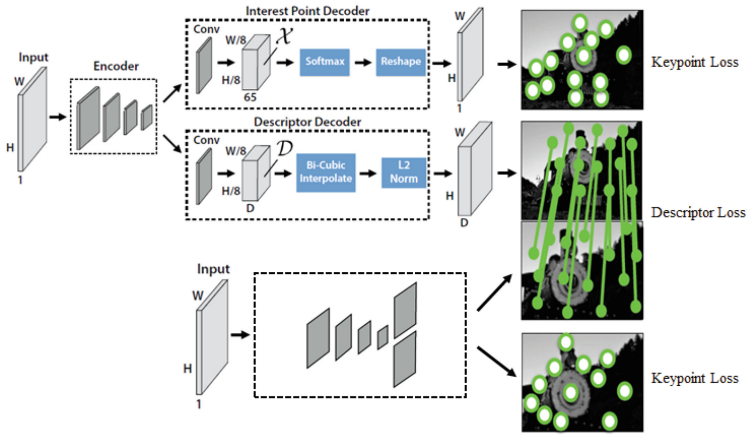
Fig. 3. The flowchart of the SuperoPoint model.

the input of size $W_c \times H_c \times 256$ and outputs a tensor sized $W \times H \times 256$. A model similar to Universal Correspondence Network(UCN) [3] is first utilized to output a semi-dense grid of descriptors, then Bicubic interpolation of the descriptors is performed, and a dense map of L_2-normalized fixed length descriptors is obtained.

The final loss is the weighted sum of keypoint detector and descriptor, the details can be referred in literature [1]. In this paper, the model is not re-trained or fine-tuned in limits of the lack of finger vein samples and ground truths. Hence, we adopt the pre-trained Superpoint model in the proposed method with careful pre- and post-processing strategies. The parameters in this model are all remained unchanged.

2.3   Matching

Two images are matched according to the similarity of descriptors calculated by the Euclidian distances. Descriptors a and b from the two images A and B are matched successfully if and only if the scores of a and b is *threshold* times higher than a with other descriptors in image B and the scores of b and a is *threshold* times higher than b with other descriptors in image A. The final matching score is defined as the number of matching pairs.

3   Experimental Results

In this section, the experimental databases and settings are first introduced. Then, the experimental results in verification and identification modes are presented with parameters and each functional components analyzed. The comparison with state-of-the-art methods is also provided.

Table 1. Component analysis on the PolyU and SDU-MLA database.

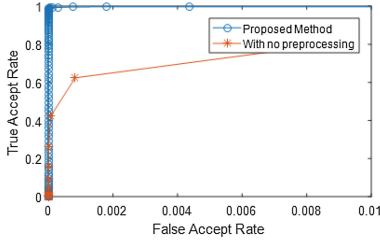| Method	HKPU Database			SDUMLA database		
	EER	FRR-at-0-FAR	RR	EER	FRR-at-0-FAR	RR
No pre-processing	0.0970	0.8397	0.6793	0.0998	0.8842	0.6787
Proposed method	0.0025	0.0218	0.9983	0.0138	0.0890	0.9792

3.1   Experimental Databases

The proposed method is evaluated on the HKPU and SDU-MLA databases, respectively. The first session of the HKPU database [11] is utilized in this paper, which includes 1,872 ($156 \times 2 \times 6$) images, and is captured from the index and middle fingers from the left hand of 156 volunteers, with each finger contributing 6 images. The second session is not conventionally adopted in literature, because not all the volunteers showed up in this session. The images in this database are various in intra-class deformations due to the uncontrolled capturing. Thus, the ROI is correspondingly segmented using methods segmented in literature [12]. The SDU-MLA database [13] consists of 3,816 ($106 \times 6 \times 6$) images, which are captured from 106 volunteers. Each of the volunteers contributes the index, middle and ring fingers of both hands, with 6 images of each finger. Images in this database are more complex in deformations and quality, we extract the ROI using a different method from literature [14].

3.2   Evaluation Protocols

The proposed method is tested in both the verification and identification modes. In the verification mode, the genuine and imposter finger vein images are fully matched. Consequently, there are $312 \times C_6^2$ genuine matchings and $312 \times 6 \times 311 \times 6/2$ imposter matchings on the HKPU database, $636 \times C_6^2$ genuine matchings and $636 \times 6 \times 635 \times 6/2$ imposter matchings on the SDU-MLA database, respectively. According to the score distribution, the Equal Error Rate (EER), False Reject Rate (FRR) at zero False Accept Rate (FAR) (FRR-at-0-FAR) are provided. In the identification mode, the real identity authentication is simulated, i.e., to find the finger of which each image belongs to. In the experiment, each finger vein images is utilized as the probe, then a template is randomly selected from the remaining ones. Hence, there are 312×6 probes and $312 \times 6 \times 312$ matcings on the HKPU database, 636×6 probes and $636 \times 6 \times 636$ matchings on the SDU-MLA database. The rank-1 recognition rate is mainly tested, the average rank-1 recognition rate of ten repeated experiments is provided. The Receiver Operating Characteristic (ROC) is also illustrated to further depict the performance.
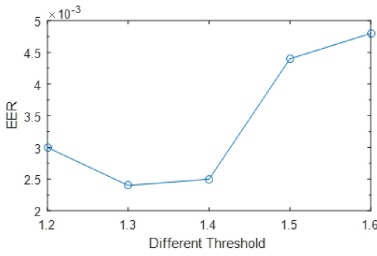
3.3   Performance Analysis

In this part, the performance of the proposed method is first analyzed in comparison to recognition without pre-processings. The threshold in matching is also
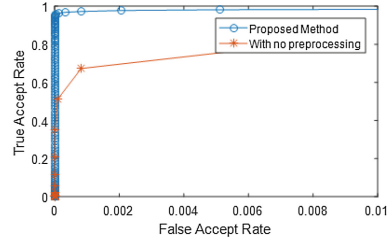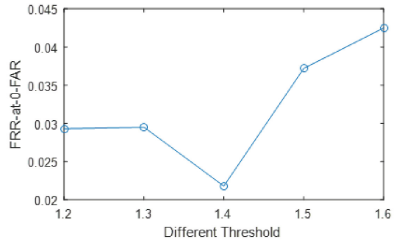
(a) The HKPU database.   (b) The SDU-MLA database.

Fig. 4. ROC curves for the two databases.

(a) EERs.   (b) FRRs-at-0-FAR.

Fig. 5. Analysis of threshold in matching on the HKPU database.

illustrated. The EERs, FRRs-at-0-FAR and Recognition Rates (RRs) on the HKPU and SDU-MLA databases are tabulated in Table 1, with the ROC curves illustrated in Fig. 4. From the results, we can figure out that the EERs on the two databases are 0.0025 and 0.0138, respectively and the FRRs-at-0-FAR of the proposed method are 0.0218 and 0.0890, respectively. The results demonstrate the effectiveness of the proposed method, the FRRs-at-0-FAR also show the applicability of our method in real applications. The average RRs of ten repeated experiments on the two databases are also provided, which are 99.83% and 97.92% with variances of $\pm 5.3584e7$ and $\pm 3.5588e6$, respectively. The method with no pre-processing is compared, the EERs are 0.0970 and 0.0998, the FRRs-at-0-FAR are 0.8397 and 0.08842, the average RRs are $0.6397(\pm 6.4177e5)$ and $0.6787(\pm 2.6026e5)$, respectively. It is obvious that the pre-processing strategies are beneficial and effective.

The parameter of threshold in matching is tested on the HKPU database, the results of EERs and FRRs-at-0-FAR is provided in Fig. 5. The thresholds are from 1.2 to 1.6 with step of 0.1, from the results, we can figure out that the proposed method is stable, we take the threshold of 1.4 in compromise of the two evaluation criterions.

We also compare the proposed method with typical pre-processings adopted in keypoint-based methods on the HKPU database, the results are tabulated in Table 2. The pre-processing techniques includes histogram equalization [15],

Table 2. The performance comparison with typical pre-processing techniques on the HKPU database.

Method	EER	FRR-at-0-FAR
Histogram equalization [15]	0.0623	0.8579
Intensity correction [9]	0.0583	0.8393
Vessel segmentaion [16]	0.0094	0.1951
Chief curvature extraction [17]	0.0418	0.5021
Intensity correction [4]	0.0429	0.9848
Proposed method	0.0025	0.0218

Table 3. The performance comparison with the keypoint-based methods on the HKPU and SDU-MLA database.

Method	HKPU database		SDU-MLA database	
	EER	FRR-at-0-FAR	EER	FRR-at-0-FAR
Pang et al. [15]	0.1081	0.9274	0.1473	0.9340
Kim et al. [9]	0.0105	0.1079	0.0532	0.2853
Peng et al. [16]	0.0235	0.4359	0.0624	0.5259
Matsuda et al. [17]	0.0379	0.7600	0.0715	0.4785
Liu et al. [2]	0.0501	0.5199	–	–
Meng et al. [4]	0.0114	0.1169	0.0585	0.2875
Proposed Method	0.0025	0.0218	0.0138	0.0890

- the results are not provided in the original paper.

bias removal-based intensity correction [4,9], vessel segmentation [16] and chief curvature extraction [17]. Here, the vessel segmentation method is adopted from our previous work of retinal vasculature segmentation [18]. From Table 2, we can figure out that all the pre-processing strategies are helpful in improving the recogniton performance. The proposed method is superior than other methods, especial for the evaluation criterion of FRR-at-0-FAR. The vessel segmentation based pre-processing is second best, with EER and FRR-at-0-FAR of 0.0094 and 0.1951, it might be brought by the enhancement of the vessel edges.

3.4   Comparison with State-of-the-Art Keypoint-Based Methods

In this part, we compared the proposed method with the existing keypoint-based methods, which consists of the Scale Invariant Feature Transform (SIFT) descriptor-based method and its variants [4,9,15,16], methods based on the minutiae matching [2] and the deformation-tolerant-based feature point matching (DT-FPM) [17], the results are tabulated in Table 3. From the table, one can figure out that the minutiae of cross overs and end points are unsatisfactory with the EER of 0.0501 and FRR-at-0-FAR of 0.5199, respectively. It may caused by the segmentation errors of the vessels. The artificially designed keypoints of

SIFT and DT-FPM also perform unsatisfactory with the EERs of 0.1081 and 0.0379, FRRs-at-0-FAR of 0.9274 and 0.7600 on the HKPU database, EERs of 0.1473 and 0.0715, FRRs-at-0-FAR of 0.9340 and 0.4785 on the SDU-MLA database, respectively, due to the image quality problem. From the work of Kim et al., Peng et al., and Meng et al., we can figure out that the preprocessings are effective to improve the keypoint-based recognition performance. The proposed method is designed utilizing the learning-based keypoints and typical preprocessing strategies are proposed, thus the performance is acceptable with EER and FRR-at-0-FAR of 0.0025 and 0.0218 on the HKPU database, 0.0138 and 0.0890 on the SDU-MLA database, respectively.

4   Conclusion and Discussion

Our method introduce automatically learned keypoints into finger vein recognition, a systematic recognition framework based on the FCN-based model of SuperPoint is proposed. Our method is demonstrated to be effective with EERs of 0.0025 and 0.0138 on the HKPU and SDU-MLA databases, respectively. In the keypoint-based recognition, unsatisfactory image quality is a big challenge which limits the performance. A targeted pre-processing is designed for vessel enhancement and scale adjusting. The results tabulated in Table demonstrates the effectiveness of the operations. There are also defects in our work, for example, the parameters are not carefully fine-tuned for better performances. What's more, typical keypoint positions should not be confined in corner points, hence our next work will focus on multiple keypoints learning.

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 4–20 (2004)
2. Liu, F., Yang, G., Yin, Y., Wang, S.: Singular value decomposition based minutiae matching method for finger vein recognition. Neurocomputing **145**, 75–89 (2014)
3. Shaheed, K., Mao, A., Qureshi, I., Kumar, M., Hussain, S., Zhang, X.: Recent advancements in finger vein recognition technology: methodology, challenges and opportunities. Inf. Fusion **79**, 84–109 (2022)
4. Meng, X., Xi, X., Yang, L., Yin, Y.: Finger vein recognition based on intensity inhomogeneity correction and scale invarint feature transform. J. Nanjing Univ. Natrual Sci. **54**(1), 1–10 (2018)
5. Lu, Y., Xie, S., Wu, S.: Exploring competitive features using deep convolutional neural network for finger vein recognition. IEEE Access **7**, 35113–35123 (2019)
6. Zeng, J., et al.: Finger vein verification algorithm based on fully convolutional neural network and conditional random field. IEEE Access **8**, 65402–65419 (2020)

7. Yang, L., Yang, G., Yin, Y., Xi, X.: Finger vein recognition with anatomy structure analysis. IEEE Trans. Circuits Syst. Video Technol. **28**(8), 1892–1905 (2017)
8. Yang, W., Hui, C., Chen, Z., Xue, J.-H., Liao, Q.: FV-GAN: finger vein representation using generative adversarial networks. IEEE Trans. Inf. Forensics Secur. **14**(9), 2512–2524 (2019)
9. Kim, H.-G., Lee, E.J., Yoon, G.-J., Yang, S.-D., Lee, E.C., Yoon, S.M.: Illumination normalization for SIFT Based finger vein authentication. In: Bebis, G., et al. (eds.) ISVC 2012. LNCS, vol. 7432, pp. 21–30. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33191-6_3
10. DeTone, D., Malisiewicz, T., Rabinovich, A.: Superpoint: self-supervised interest point detection and description. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 224–236 (2018)
11. Kumar, A., Zhou, Y.: Human identification using finger images. IEEE Trans. Image Process. **21**(4), 2228–2244 (2011)
12. Meng, X., Xi, X., Yang, G., Yin, Y.: Finger vein recognition based on deformation information. Sci. China Inf. Sci. **61**(5), 1–15 (2018)
13. Yin, Y., Liu, L., Sun, X.: SDUMLA-HMT: a multimodal biometric database. In: Sun, Z., Lai, J., Chen, X., Tan, T. (eds.) CCBR 2011. LNCS, vol. 7098, pp. 260–268. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25449-9_33
14. Meng, X., Yang, G., Yin, Y., Xiao, R.: Finger vein recognition based on local directional code. Sensors **12**(11), 14937–14952 (2012)
15. Pang, S., Yin, Y., Yang, G., Li, Y.: Rotation invariant finger vein recognition. In: Zheng, W.-S., Sun, Z., Wang, Y., Chen, X., Yuen, P.C., Lai, J. (eds.) CCBR 2012. LNCS, vol. 7701, pp. 151–156. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35136-5_19
16. Peng, J., Wang, N., Abd El-Latif, A.A., Li, Q., Niu, X., Finger-vein verification using Gabor filter and sift feature matching. In: Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, vol. 2012, pp. 45–48. IEEE (2012)
17. Matsuda, Y., Miura, N., Nagasaka, A., Kiyomizu, H., Miyatake, T.: Finger-vein authentication based on deformation-tolerant feature-point matching. Mach. Vis. Appl. **27**(2), 237–250 (2016)
18. Meng, X., Yin, Y., Yang, G., Han, Z., Yan, X.: A framework for retinal vasculature segmentation based on matched filters. Biomed. Eng. Online **14**(1), 1–20 (2015)