Multi-Perspective Finger-Vein Biometrics

Bernhard Prommegger and Christof Kauba and Andreas Uhl
University of Salzburg
Jakob-Haringer-Str. 2
5020 Salzburg, AUSTRIA
{bprommeg, ckauba, uhl}@cosy.sbg.ac.at

Abstract

Most finger vein recognition systems use palmar finger images. There is some work on the dorsal view, but the remaining views have not been sufficiently investigated yet. All major public available finger-vein databases contain only images from the palmar view and only one smaller database has images from the dorsal view. We aim to fill this gap and evaluate the performance using other perspectives than dorsal and palmar. Therefore, we established a new finger vein data set that consists of videos showing the vein structure all around the finger. We carried out several experiments utilizing common finger-vein recognition algorithms to quantify the recognition performance of each single projection. We further analyzed if a fusion of different views can improve the recognition performance of the system.

1. Introduction

Biometric authentication systems are well established today as they exhibit many advantages over traditional password and token based ones. The most prominent examples are fingerprint and face recognition systems. In recent times authentication based on finger- and hand-veins has gained more attention as they provide several advantages over the well established fingerprint ones. Finger-vein recognition utilizes the pattern of the blood vessels inside the hand of a human which is captured using near infrared (NIR) illumination. Finger-vein recognition is more resistant against forgery because the veins are underneath the skin and only visible in infrared light. In addition the vein patterns are neither susceptible to abrasion nor skin surface conditions. The drawbacks of finger-vein based recognition systems include relatively big capturing devices compared to fingerprint sensors, images having low contrast and quality in general and that the vein structure may be influenced by temperature, physical activity and certain injuries and diseases.

Currently there has only been little research on finger-veins biometric recognition systems using images from different viewpoints. Most works focus on the analysis of the palmar perspective [1, 7, 15, 16, 18, 19]. Raghavendra and Busch [12] proposed a vein recognition system for the dorsal perspective, Lu et al. [8] fused images from two different views which are positioned quite close to each other. Zhang et al. [22] applied point cloud matching on hand-veins using two cameras. All major publicly available finger-vein databases contain only images from the palmar or dorsal perspective. Table 1 lists these available data sets. The only dorsal database has just been released and was not used in any publication so far. To the best of our knowledge, there is no work analyzing finger-veins using perspectives all around the finger. Hence, it is not clear if there are other perspectives that provides better or enough additional information to improve the performance of the recognition system. Another advantage of using several perspectives is an increased robustness against spoofing attacks. It has been shown that finger- as well as hand-vein recognition systems are susceptible to spoofing [15, 14]. The proposed spoofing technique is based on a simple paper printout of the vein pattern. Capturing the vein images from different perspectives will prevent such simple kinds of spoofing attacks.

Name	Subjects	Finger	Images	View
UTFVP [16]	60	6	1440	palmar
FV-USM [1]	123	4	5940	palmar
MMCBNU 6000 [7]	100	6	6000	palmar
SDUMLA HMT [19]	106	6	3816	palmar
VERA FV DB [15]	110	2	440	palmar
THU-FVFDT [18]	610	2	6540	palmar
PROTECT MM DB [17]	20	4	240	dorsal
HKPU-FID [5]	156	2	3132	palmar

Table 1. Publicly available finger-vein data sets

The main goal of this work is to evaluate the recognition performance of finger-vein images taken from different perspectives. We evaluate the performance of these additional perspectives in order to find out if they exhibit a better or similar performance compared to the palmar one or

978-1-5386-7180-1/18/$31.00

at least provide enough information to improve the recognition performance when fusing them. The provided data set can also be used to evaluate the robustness of finger recognition systems against longitudinal finger rotation. Based on this data set, the different projections' individual recognition performances are evaluated utilizing some well established vein recognition schemes, compared and ranked. In addition, we conducted experiments using score-level fusion of selected projections in order to find out if the recognition performance can be increased further.

The rest of the paper is organized as follows: Section 2 describes the used finger vein recognition system. At first the image acquisition is explained, then the algorithms for ROI extraction, preprocessing, feature extraction and comparison are briefly outlined. Section 3 contains informations about the custom build multi-perspective finger-vein scanner hardware used to acquire the new data set as well as a description of the data set itself. Section 4 presents the experimental protocol and discusses the results. Section 5 concludes this paper.

2. Finger-Vein Recognition System

2.1. Image Acquisition

A hand-vein scanner consists of 2 basic components: a near-infrared (NIR) sensitive camera and a NIR light source. Usually there is some automatic illumination intensity control to achieve an optimal contrast of the vein images. The wavelength of the NIR light source is typically between 730 and 950 nm. The near-infrared light is absorbed by the haemoglobin in the blood flowing through the veins and arteries. Thus, they appear as dark lines in the captured images. The camera should be equipped with an NIR pass-through filter to block the ambient light and further enhance the image contrast.

2.2. Preprocessing, Feature Extraction and Comparison

ROI Extraction Prior to the extraction of the region of interest (ROI), the finger is aligned and normalized. The alignment should place the finger always in the same position - independent of the relative position of the finger during the acquisition. To achieve this, we detected the finger lines (edge between finger and the background of the image) and calculate the center line (in the middle of the two finger lines). Next we rotate and translate the center line of the finger in a way that it is placed in the middle of the image and mask the image outside of the finger region. The final step is to extract a rectangular ROI. In order to keep the whole information of the vein structure, we first adjust the finger region to fit into a rectangle of defined height and just cut off some pixels on the border. The three steps are visualized in Figure 1.

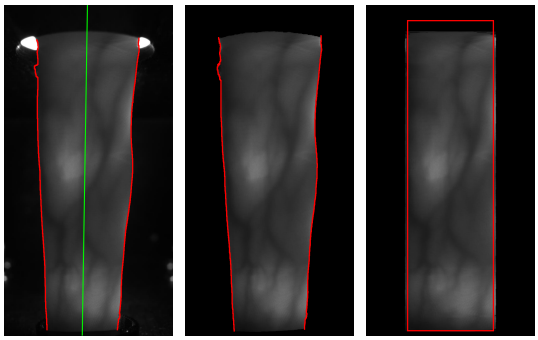

Figure 1. ROI Extraction

Preprocessing Preprocessing tries to enhance the low contrast and improve the image quality. Simple **CLAHE** [24] or other local histogram equalization techniques are most prevalent according to the literature for this purpose. We used **High Frequency Emphasis Filtering (HFE)** which was originally proposed for hand vein image enhancement [23]. In addition, filtering using a **Circular Gabor Filter** (**CGF**) as proposed by Zhang and Yang [20] was applied. Furthermore, the images were resized to half of their original size, which not only speed up the comparison process but also improved the results. For more details on the preprocessing methods the interested reader is referred to the authors' original publication [3].

Feature Extraction and Comparison We used two different types of feature extraction and comparison methods. The first three techniques discussed aim to extract the vein pattern from the background resulting in a binary image followed by a comparison of these binary images using a correlation measure. All algorithms are well established and therefore deliver reproducible results. We used the publicly available implementations published in [4].

Maximum Curvature (MC [11]) aims to emphasize only the center lines of the veins and is therefore insensitive to varying vein widths. The first step is the extraction of the center positions of the veins. Afterwards, a score according to the width and curvature of the vein region is assigned to each center position which is recorded in a matrix called locus space. Due to noise or other distortions some pixels may not have been classified correctly at the first step, thus the center positions of the veins are connected using a filtering operation. Finally binarization is done by thresholding using the median of the locus space.

Principal Curvature (PC [2]): At first the gradient field of the image is calculated. Hard thresholding is done to filter out small noise components and then the gradient at each pixel is normalized to 1 to get a normalized gradient field.

This is smoothed by applying a Gaussian filter. The next step is the actual principal curvature calculation. It is obtained from the Eigenvalues of the Hessian matrix at each pixel. The two Eigenvectors of the Hessian matrix represent the directions of the maximum and minimum curvature and the corresponding Eigenvalues are the principal curvatures. Only the bigger one which corresponds to the maximum curvature is used. The last step is a binarization of the principal curvature values to get the binary vein output image.

Gabor Filter (GF [5]): A filter bank consisting of several 2D even symmetric Gabor filters with different orientations (in $\frac{\pi}{k}$ steps where k is the number of orientations) is created. Several features images are extracted by filtering the vein image using the different filter kernels of the Gabor filter bank. The final feature image is obtained by fusing all the single images from the previous step. This final vein output image is then post-processing using morphological operations to remove noise.

For comparison the binary feature images we extended the approach in [10] and [11]. As the input images are neither registered to each other nor aligned, the correlation between the input image in x- and y-direction shifted versions of the reference image is calculated. The maximum of these correlation values is normalized and then used as final comparison score.

In contrast to the techniques described above, key-point based techniques try to use information from the most discriminative points as well as considering the neighborhood and context information of these points by extracting key-points and assigning a descriptor to each key-point. We used a **SIFT** [6] based technique with additional key-point filtering. Details are described in [3].

3. Multi-Perspective Finger-Vein Data Set

Due to the lack of an existing data set consisting of finger-vein images from different perspectives, we established a new data set which will be made publicly available. The images have been acquired using a custom build scanner. The different projection angles are achieved by rotating a NIR camera and the illumination unit around the finger. The principle is shown in Figure 2: The finger is positioned at the axis of rotation, whereas the camera and the illumination module are placed on the opposite sides, rotating around the finger, i.e. the scanner is based on the transillumination principle. The rotation of camera and light source enables the scanner to acquire images from different views.

3.1. Multi-Perspective Finger-Vein Scanner

Our custom build sensor is based on the above mentioned principle. All non-commercially available parts were engineered and manufactured by ourselves using a 3D printer and a laser cutter for the wooden parts. Figure 3 shows the unwrapped scanner with all its components. In the middle of

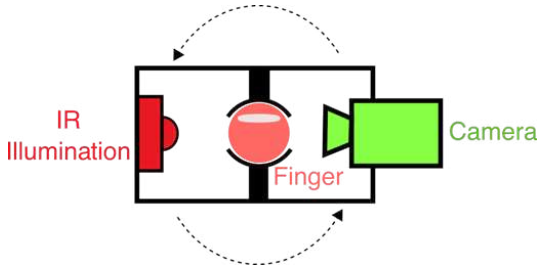

Figure 2. Basic principle of our finger vein scanner

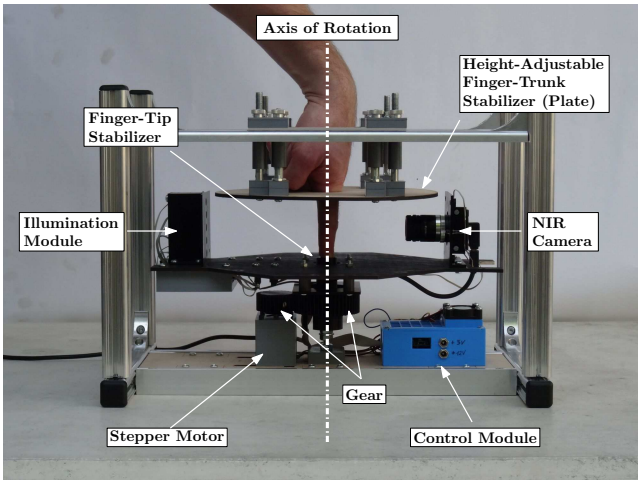

Figure 3. Custom build multi-perspective finger-vein scanner

the image you can see the finger. In order to keep the movement of the finger during the acquisition to a minimum, we added two parts to help to stabilize the finger. For the finger-tip, we constructed a part that has a finger-tip shaped hole. Putting the finger into this hole keeps it in its position. For the hand we added a height-adjustable wooden plate. Placing the hand on this plate stabilizes the trunk of the finger. The height of the plate is adjusted according to the length of the captured finger. The illumination module on the left side consists of 5 NIR laser diodes (808 nm) placed on a strip. The illumination intensity of each laser diode is controlled separately. The plane of focus is set at the axis of rotation where the center of the finger is located. This results in a uniform illumination of the finger. During data acquisition, the intensity of the different laser diodes is set automatically. This is achieved by individually setting the intensity value (controlled by the operating current) for each laser. The best value is selected by evaluating the image contrast in the corresponding section of the image. The images are captured by a NIR enhanced industrial camera (IDS Imaging UI-1240ML-NIR, max. resolution 1280x1024 pixels) with a 9 mm wide-angle-lens (Fujifilm, HF9HA-1b, 9mm, 2/3"). An additional NIR longpass filter (Midopt LP780, useful range: 800-1100mm) mounted on the lens blocks visible light up to a wavelength 780 nm. The rotation is ac-

complished using a stepping motor (SY42STH47-1684A). The stepper and the rotor are connected via self printed cogwheels having a gear ratio of 1:5/3 (motor to rotor). One step corresponds to $0.0675°$, thus it is possible to capture a maximum of 5.333 different projections. The sensor has a size of 25.8 x 32.5 x 45.5 cm (width x height x depth). The rotor has a length of 38 cm.

The acquisition process is semi-automated. After the finger is put into the device and the capture process is initiated, the illumination for the finger is set automatically in order to achieve an optimal image contrast with the help of a contrast measure. After this, the video acquisition is started. To achieve a defined resolution (in degrees) of images (video frames), the speed of the rotation and the video frame rate are coordinated with each other. All perspectives are captured in one run using the same illumination conditions to ensure the comparability of the different projections.

The automated illumination algorithm evaluates the average gray level of the image around in the center of each laser diode ($GL_{i,current}$)) and tries to achieve a preconfigured target gray value ($GL_{i,target}$). The centers of the diodes are arranged along the longitudinal axis of the finger. The individual intensity values of all diodes are set at once. Initially all diodes are set to half of its max intensity (I_{max}). The intensity is corrected by:

$$correction_i = \frac{GL_{i,target} - GL_{i,current}}{GL_{max}} * \frac{I_{max}}{2 * n} \quad (1)$$

where GL_{max} is the maximum gray value and n is the number of the current iteration. The maximum number of iterations is $\log_2(I_{max})$.

3.2. Data Set

The data set currently contains of a total of 252 unique fingers from 63 different subjects, 4 fingers per subject. We acquired videos from the index and middle finger of both hands where the target resolution is $1°$. As acquiring of the ring finger would be ergonomically uncomfortable for our volunteers, we skipped capturing of this finger.

We extracted the video frames as images which leads to 361 different perspectives (361 as we captured one frame for $0°$ and $360°$). Due to some variations in the video frame rate and the speed of the rotation during the capturing, we got between 357 and 362 frames for a full rotation (ideal would be 361 frames). To get 361 perspectives for every finger, we mapped the frame with the minimum deviation to the desired position for every perspective. This results in a maximum deviation of $0.5042°$ to its desired position.

Every finger was acquired 5 times - each time removing the finger from the scanner and putting it in again. This results in $252 * 361 * 5 = 454.860$ images in total. One projections consists of $252 * 5 = 1260$ images. The extracted frames are 8-bit gray scale images with a resolution of 1024*1280 pixels. Due to the fact that the finger is always positioned in the middle of the scanner, we cut of the borders. This results in to a final image resolution of 650x1280. Figure 4 shows our sensor during data acquisition. Recording a single person with 4 fingers and 5 iterations takes approximately 15 minutes.

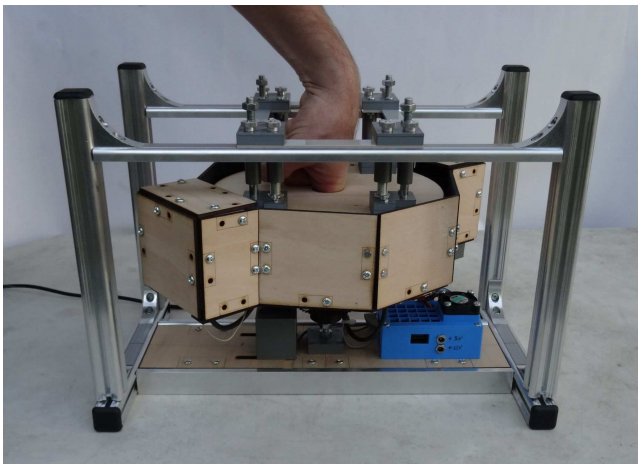

Figure 4. Data capturing

Figure 5 shows 6 example images ($0°$ - $300°$ in $60°$ steps, $0°$ corresponding to palmar view and $180°$ corresponding to dorsal view, respectively). It is apparent, that the number of visible veins in the images differ among the different projections. The black area at the top results from the hand stabilization plate. Depending on the length of the finger, the plate is pushed in further or less far.

The gender distribution of the volunteers is balanced. Among the 63 subjects, 27 are female (43%) and 36 men (57%). The dataset represents a good cross section among all age groups with a slight overhang among the 20-40 year olds. The youngest participant was 18, the oldest 79 years. Due to national law, we were not allowed to acquire data from people younger than 18. The actual distribution is shown in Figure 6. Our subjects are from 11 different countries[1] where the majority is white Europeans (73%).

4. Experiments

The experiments are split into two main parts: in the first part we analyze the recognition performance of the different projections. Every perspective is considered as a separate data set. We did not perform cross-projection comparison. We processed the images as described in section 2. As we aim for analyzing the recognition performance from views all around the finger, we used 73 perspectives extracted in $5°$ steps. To quantify the performance we used the EER as well as the FMR100 (the lowest $FNMR$ for $FMR \leq 1\%$), the

[1] Austria, Brazil, China, Ethiopia, Germany, Hungary, Iran, Italy, Russia, Slovenia, USA

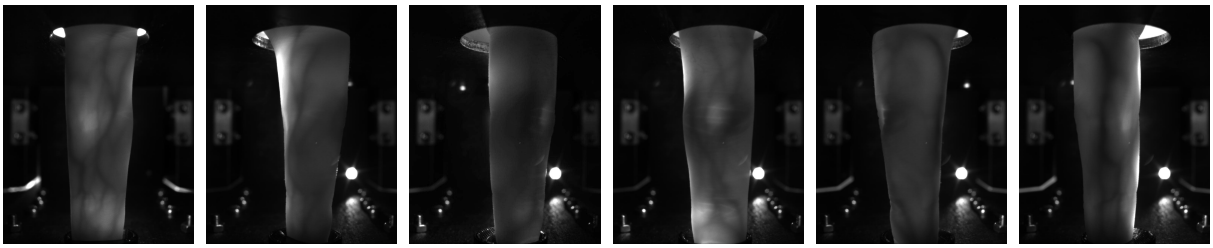

Figure 5. Example images of the data set acquired from 0° to 360° in 5° steps

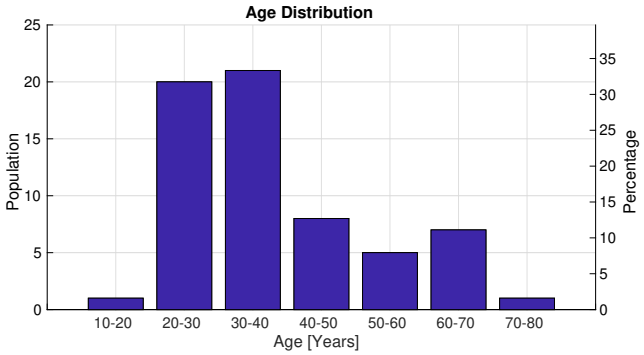

Figure 6. Age distribution among all subjects

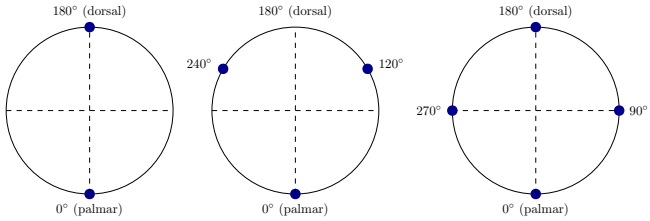

Figure 7. Selection of view angles for fusion with 2, 3 and 4 perspectives

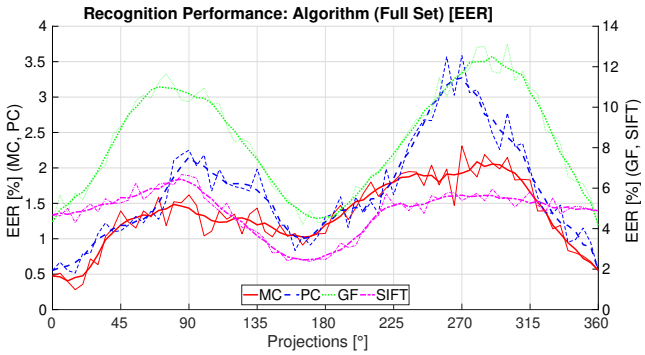

Figure 8. Recognition performance for different projections (EER)

FMR1000 (the lowest $FNMR$ for $FMR \leq 0.1\%$) and the ZeroFMR (the lowest $FNMR$ for $FMR = 0\%$). For their calculation we followed the test protocol of the FVC2004 [9]: for calculating the genuine scores for each projection, all possible genuine matches are done, which are $63 * 4 * \frac{5*4}{2} = 2520$ matches. For calculating the impostor scores, only the first image of a finger is matched against the first image of all other fingers, resulting in $4 * \frac{63*62}{2} = 7812$ matches, so together 10332 matches in total. All values are given in percentage terms, e.g. 2.35 means 2.35%.

In the second part of our experiments, we apply score-level fusion [13] to selected projections in order to improve the recognition performance. We start with the fusion of two views and increase the number to the maximum of 72 views. The perspectives used are evenly distributed over the whole circle. The step width of 5° of the selected images allows us to fuse 2, 3, 4, 6, 8, 9, 12, 18, 24, 36 and 72 different projections. Figure 7 shows the principle for the first 3 options. In addition, we fuse certain angles in a 2-view-fusion against all other perspectives. Since all scores are from the same modality using the same feature extraction method (we do not fuse results from different feature extraction algorithms), a score normalization was not necessary. For the fusion we used the simple sum rule.

4.1. Recognition Performance Results

Figure 8 visualizes the results for the evaluation of the recognition performance using MC, PC, GF and SIFT based on the EER. For every method there are two lines: the thin line shows the actual EER values of the relevant view, the thicker line is calculated from the EER values using a moving average filter of size 5 and should highlight the trend of the recognition performance. As the step between two projections is only 5°, the acquired images of neighboring perspectives show similar vein structures. This results in similar recognition performances. The best results are obtained around the palmar (0°) and dorsal (180°) region. The inferior results of the perspectives between those two view

can be explained by the fact that they contain fewer vein information, as it can be seen in Figure 9. It shows the original ROI, the ROI after preprocessing, and the extracted veins (using MC) for the views at 0°, 90°, 180° and 270°. It reveals, that the palmar and dorsal perspectives contain more vein information than the other two views. Moreover, it turns out that vein extraction - especially at 180° - compromises some features related with the knuckles of the finger. This features can be recognized as horizontal lines in the feature image.

For SIFT, the best performance is achieved around the dorsal region. The palmar region exhibits better performance than the remaining perspectives as well, but it is inferior compared to the dorsal one. This is due to the applied preprocessing: for SIFT we only apply algorithms that enhance the vein structure, but not a vein extraction algorithm (binarization) ahead of the SIFT point calculation. This prevails the texture of the finger. Especially the structure of finger knuckles seem to contain a lot of information in it. Finger knuckles have been introduced by Zhang et al. [21] as its own biometric modality. This could explain the better performance at the dorsal view. Yang [18] et al. experienced similar behavior. They fused the finger vein structure of the palmar view and the finger texture of the dorsal view which improved the recognition performance.

The EERs for the best projections are in accordance with the rates achieved in well-established implementations. For projections other than palmar/dorsal no comparisons are available. The best/worst rates are shown in Table 2.

	Worst Result		Best Result	
	View	EER	View	EER
MC	270°	2.31	15°	0.28
PC	270°	3.59	15°	0.52
GF	300°	13.12	360°	4.16
SIFT	85°	6.67	170°	2.38

Table 2. Recognition performance (EER) for single views

The results for FMR100 (Figure 10), FMR1000 (Figure 11) and ZeroFMR (Figure 12) show the same trend as the EER. Again, the best performance is achieved around 0° and 180°.

To ensure that two opposing views do not contain the same (just mirrored) information, we further investigated the palmar and dorsal perspective. We mirrored the images of the dorsal view along the longitudinal axis of the finger and matched them against the palmar ones. If both perspectives show the same blood vessels, they should - due to the mirroring - be registered to each other, and a comparison against each other would show similar performance as comparing the single views itself. Our results show exactly the opposite behavior: The EER of all four used algorithms is close to 50% which means that the vein structure of the two perspectives is not related to each other. Figure 9 shows the original ROI, the ROI after preprocessing, and the extracted veins (using MC) for both projections. It is obvious that the vessel structure differs between the palmar and dorsal view. Table 3 shows the results in detail.

	EER for Perspective		
	0°	180°	0° vs 180°
MC	0.47	1.08	47.28
PC	0.55	1.31	49.41
GF	4.33	4.38	50.04
SIFT	4.68	2.48	46.74

Table 3. Recognition performance (EER) for palmar vs dorsal view

4.2. Score-Level Fusion Results

In the second part of our experiments we analyze the impact of fusing selected perspectives. In the first experiment we fuse an increasing number of views which are, as described in section 4, evenly distributed over the whole circle. As starting angles we used 0° (palmar view), 45°, 90°, 180° (dorsal view) and 270°.

Figure 13 visualizes the results for start angle 0°. The fusion results for MC, PC, GF and SIFT are similar: Fusing the palmar with dorsal view improves the result. With the fusion of 3 views (60°, 180° and 300°), the result is slightly inferior to the one with two views. This can be explained by the fact that we removed the well performing palmar view and replaced it with two less performing ones. When adding additional views, the recognition performance further increases and stabilizes at a high level.

As the remaining 4 start angles show - in principle - the same general trend, we do not discuss them in detail. Figure 14 shows the results of the other reference angles. Table 4 holds detailed results for the multi-perspective fusion. For every method/reference angle combination it shows the EER for the reference view and the worst/best fusion result. The number of views is always the minimum number of needed perspectives to achieve the stated EER. We achieved a clear performance increase for all combinations.

In our second fusion experiment we applied a 2-view-fusion of a certain perspective against all others. This was done for every feature extraction and comparison algorithm. As reference views we selected a good performing (0°) and an inferior performing view (270°).

Figure 15 shows the results for MC: The solid red line shows the recognition performance without fusion and serves as reference for the other lines. The dotted green line represents the fusion results for 270°. None of the fused values has a worse performance than the single-view performance of 270°. The dashed blue line for 0° shows the same behavior, although not as distinctive.

PC, GF and SIFT (not visualized) show similar behavior.

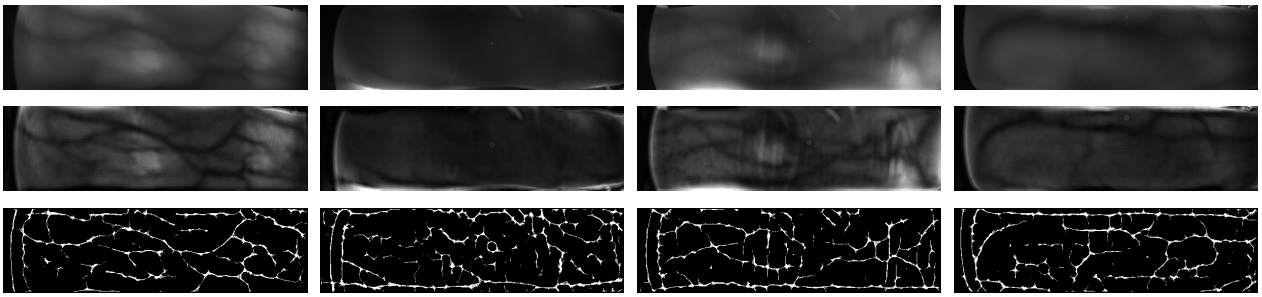

Figure 9. ROI, enhanced image and extracted features (MC) for different projections

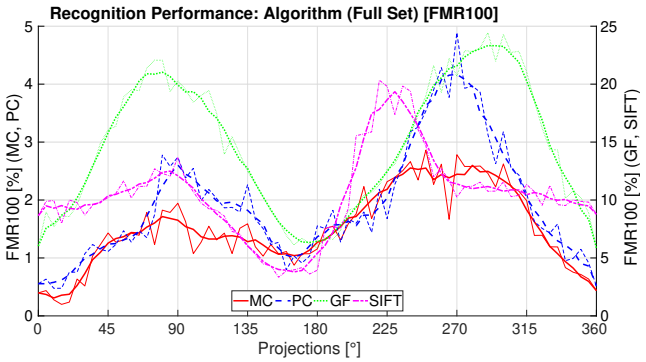

Figure 10. Recognition performance for different projections (FMR100)

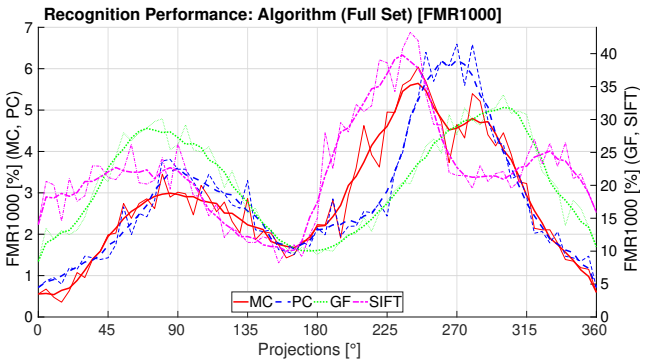

Figure 11. Recognition performance for different projections (FMR1000)

Note that the fusion of two views can result in an inferior recognition performance than the better of the two used views. E.g. this occurs for PC and reference view 0° when fusing it with the 305° perspective. Table 5 holds detailed results for the 2-view-fusion. It shows the EER for the reference view and the worst/best fusion result for every method/reference angle combination.

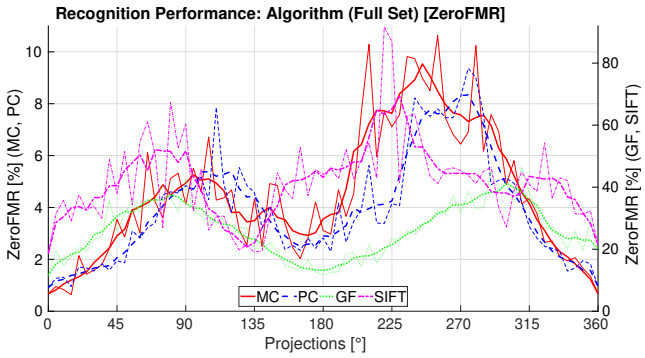

Figure 12. Recognition performance for different projections (ZeroFMR)

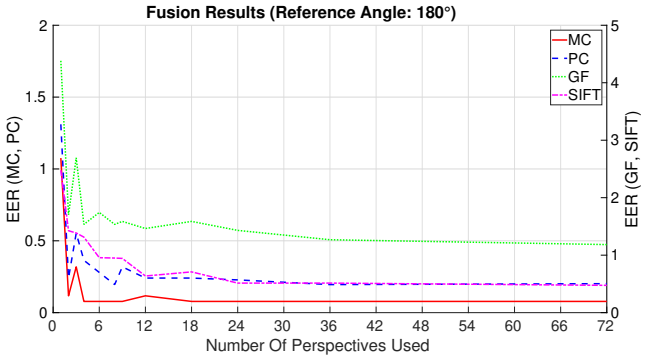

Figure 13. Recognition performance for fusion of different number of perspectives

4.3. Results Discussion

Our analysis of the recognition performance for different projections showed, that the widely used perspectives, palmar and dorsal, perform best. We also showed, that the vein structure of palmar and dorsal view are not connected to each other. The angles inbetween show a slightly worse performance, but it is still acceptable. For further in-depth analysis - e.g. on the individual performance of left/right hand or single fingers - the data set has to be extended by acquiring further subjects.

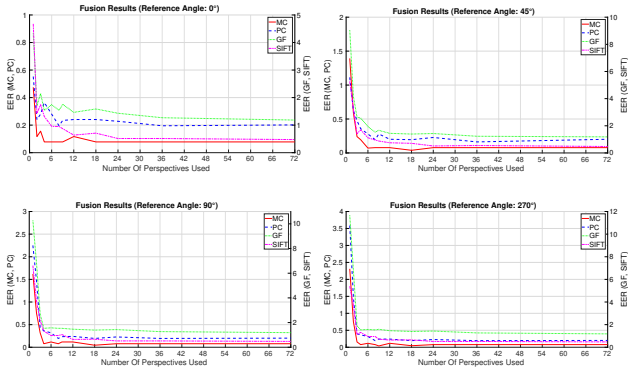

Figure 14. Recognition performance for fusion of different number of perspectives

	Reference		Worst Result		Best Result	
	View	EER	#	EER	#	EER
MC	0°	0.47	3	0.16	4	0.08
	45°	1.39	2	0.59	18	0.04
	90°	1.62	2	0.75	18	0.05
	180°	1.08	3	0.32	4	0.08
	270°	2.31	2	0.75	9	0.04
PC	0°	0.55	4	0.36	8	0.20
	45°	1.11	2	0.62	36	0.16
	90°	2.25	2	1.47	8	0.20
	180°	1.31	3	0.55	8	0.20
	270°	3.59	2	1.47	8	0.20
GF	0°	4.33	3	2.15	72	1.18
	45°	9.04	2	4.17	72	1.18
	90°	10.27	2	6.81	72	1.18
	180°	4.38	3	2.69	72	1.18
	270°	11.65	2	6.81	72	1.18
SIFT	0°	4.68	3	1.74	72	0.47
	45°	5.27	2	3.22	72	0.47
	90°	6.62	2	3.73	72	0.47
	180°	2.48	2	1.42	72	0.47
	270°	5.40	2	3.73	72	0.47

Table 4. Detailed performance results for multi-perspective fusion

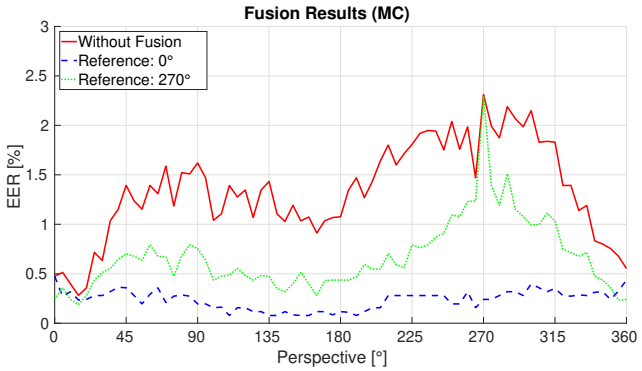

Figure 15. Recognition performance for fusion of two perspectives (MC)

The results indicate that the presence of finger texture has a positive influence on the recognition performance. As

	Reference		Worst Result		Best Result	
	View	EER	View	EER	View	EER
MC	0°	0.47	5°	0.47	110°	0.08
	270°	2.31	275°	2.30	15°	0.19
PC	0°	0.55	305°	0.68	170°	0.15
	270°	3.59	275°	3.58	360°	0.59
GF	0°	4.33	5°	4.33	170°	1.63
	270°	11.65	275°	11.64	165°	3.25
SIFT	0°	4.68	5°	4.68	170°	1.27
	270°	5.40	275°	5.39	175°	1.59

Table 5. Detailed performance results for reference view fusion

it can be seen in Figure 9, regularly used feature extraction algorithms also recognize the texture of the finger and thereby implicitly fuse vein structure and texture. This happens especially in the dorsal region with finger knuckles.

Additionally we showed, that the fusion of multiple perspectives improves the recognition performance. The fusion of two opposite views achieves is already sufficient to achieve superior results compared to a single-view evaluation.

5. Conclusion

We established a new finger vein data set containing videos that capture the vein structure all around the finger. The videos allow us to extract frames in steps of 1°. Based on this data set, we evaluated the recognition performance using several common finger-vein recognition algorithms on each of the projections which enabled a direct comparison in terms of their accuracy. According to our experimental results, the best performance was achieved around 0° and 180° which corresponds to the palmar and dorsal view. We further showed that a fusion of two or more perspectives can improve the recognition results.

Our future work includes further analysis of the data we acquired using our custom made finger-vein scanner. We will use the data acquired in 1° steps to verify the robustness of existing algorithms with respect to the finger tilt. We aim to further improve recognition performance by fusing the information from different perspectives (further experiments with vein structure only and vein fused with texture are planned). Our final goal is to achieve a complete 3D reconstruction of the finger-vein structure. Our work will also include performance improvements in our scanner hardware.

6. Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 700259.

References

[1] M. S. M. Asaari, S. A. Suandi, and B. A. Rosdi. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. *Expert Systems with Applications*, 41(7):3367–3382, 2014.

[2] J. H. Choi, W. Song, T. Kim, S.-R. Lee, and H. C. Kim. Finger vein extraction using gradient normalization and principal curvature. In *IS&T/SPIE Electronic Imaging*, pages 725111–725111. International Society for Optics and Photonics, 2009.

[3] C. Kauba, J. Reissig, and A. Uhl. Pre-processing cascades and fusion in finger vein recognition. In *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG'14)*, Darmstadt, Germany, Sept. 2014.

[4] C. Kauba and A. Uhl. Shedding light on the veins - reflected light or transillumination in hand-vein recognition. In *Proceedings of the 11th IAPR/IEEE International Conference on Biometrics (ICB'18)*, pages 1–8, 2018.

[5] A. Kumar and Y. Zhou. Human identification using finger images. *Image Processing, IEEE Transactions on*, 21(4):2228–2244, 2012.

[6] D. G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision (CVPR'99)*, volume 2, pages 1150 – 1157. IEEE, 1999.

[7] Y. Lu, S. J. Xie, S. Yoon, Z. Wang, and D. S. Park. An available database for the research of finger vein recognition. In *Image and Signal Processing (CISP), 2013 6th International Congress on*, volume 1, pages 410–415. IEEE, 2013.

[8] Y. Lu, S. Yoon, and D. S. Park. Finger vein identification system using two cameras. *Electronics Letters*, 50(22):1591–1593, 2014.

[9] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain. FVC2004: Third Fingerprint Verification Competition. In *ICBA*, volume 3072 of *LNCS*, pages 1–7. Springer Verlag, 2004.

[10] N. Miura, A. Nagasaka, and T. Miyatake. Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Machine Vision and Applications*, 15(4):194–203, 2004.

[11] N. Miura, A. Nagasaka, and T. Miyatake. Extraction of finger-vein patterns using maximum curvature points in image profiles. *IEICE transactions on information and systems*, 90(8):1185–1194, 2007.

[12] R. Raghavendra and C. Busch. Exploring dorsal finger vein pattern for robust person recognition. In *Biometrics (ICB), 2015 International Conference on*, pages 341–348. IEEE, 2015.

[13] A. Ross and K. Nandakumar. *Encyclopedia of Biometrics: Fusion, Score-Level*, pages 611–616. Springer US, Boston, MA, 2009.

[14] P. Tome and S. Marcel. On the vulnerability of palm vein recognition to spoofing attacks. In *The 8th IAPR International Conference on Biometrics (ICB)*, May 2015.

[15] P. Tome, M. Vanoni, and S. Marcel. On the vulnerability of finger vein recognition to spoofing. In *IEEE International Conference of the Biometrics Special Interest Group (BIOSIG)*, Sept. 2014.

[16] B. T. Ton and R. N. J. Veldhuis. A high quality finger vascular pattern dataset collected using a custom designed capturing device. In *Proceedings of the 2013 International Conference on Biometrics (ICB), Madrid, Spain*, pages 1–5, 2013.

[17] University of Reading. PROTECT Multimodal DB Dataset, June 2017. Available by request at projectprotect.eu/dataset.

[18] W. Yang, X. Yu, and Q. Liao. Personal authentication using finger vein pattern and finger-dorsa texture fusion. In *Proceedings of the 17th ACM international conference on Multimedia*, pages 905–908. ACM, 2009.

[19] Y. Yin, L. Liu, and X. Sun. Sdumla-hmt: a multimodal biometric database. *Biometric Recognition*, pages 260–268, 2011.

[20] J. Zhang and J. Yang. Finger-vein image enhancement based on combination of gray-level grouping and circular gabor filter. In *Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on*, pages 1–4. IEEE, 2009.

[21] L. Zhang, L. Zhang, and D. Zhang. Finger-knuckle-print: a new biometric identifier. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 1981–1984. IEEE, 2009.

[22] Q. Zhang, Y. Zhou, D. Wang, and X. Hu. Personal authentication using hand vein and knuckle shape point cloud matching. In *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*, pages 1–6. IEEE, 2013.

[23] J. Zhao, H. Tian, W. Xu, and X. Li. A new approach to hand vein image enhancement. In *Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on*, volume 1, pages 499–501. IEEE, 2009.

[24] K. Zuiderveld. Contrast limited adaptive histogram equalization. In P. S. Heckbert, editor, *Graphics Gems IV*, pages 474–485. Morgan Kaufmann, 1994.