Why Vein Recognition Needs Privacy Protection

Daniel Hartung, Christoph Busch
Gjøvik University College, Norwegian Information Security Laboratory (NISlab)
{daniel.hartung, christoph.busch}@hig.no

Abstract

This paper describes the emerging biometric modality of vein recognition and privacy concerns that arise with its widespread use. Current sensors are able to capture vein patterns inside the human body, this is considered as a private biometric characteristic. In fact two medical disease patterns are presented that can be extracted from the vein patterns. In order to be compliant with data privacy protection laws privacy enhancing mechanisms have to be applied in vein recognition systems. Experiments of applying the helper data scheme to a back-hand vein database were conducted with remarkable results. A privacy-enhanced verification system can be realized, which shows good biometric performance under laboratory conditions.

1. Vein Recognition

The randomness of vein patterns is epigenetic, even identical twins can be distinguished. During the embryonic period the blood vessels are formed, this process of growth is not determined by the DNA sequence.

The pattern is available at every healthy human, making it an interesting research objective. Commercial applications evolved out of this research, nowadays many ATMs in Japan and Brazil are secured using this biometric modality. With the upcoming changes of the liability situation in the Single Euro Payments Area (SEPA) it is likely that this biometric technology will also be widespread in Europe.

The patterns are commonly extracted from images of the palm, the back of the hand or fingers as seen in figure 1. Recently Yanagawa et al. showed that the diversity of finger vein patterns among different persons is competitive to iris-based systems [13]. The International Biometrics Group (IBG) 6th report 2006 confirms recognition rates fairly at the same level for two different vein and one iris-based authentication system [2]. An interesting aspect of vein recognition is the fact that the information is not visible, it is hidden inside the body. Unlike fingerprints it is not possible to leave a vein pattern representation unintentionally in public places and thus it is not possible for an attacker to acquire the pattern in daily life or to replicate it. Furthermore there is no relation to criminal prosecution. How vein images are captured is described in the next section.

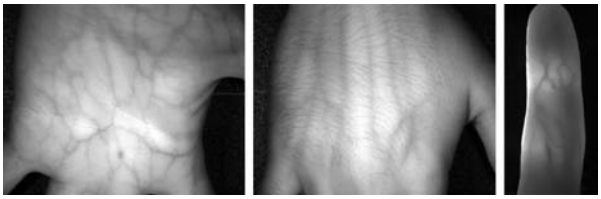

Figure 1. Palm, back-hand taken from [11] and finger vein images [1]

1.1. Imaging Techniques

The imaging approach makes use of the absorption capacity of particular substances in the blood running through the veins. To capture the image, the region of interest is illuminated with a near-infrared (NIR) light source with wavelengths around 700 to 1,000 nm. A reflection or transmission technique can be used. Deoxygenized hemoglobin highly absorbs rays within this wavelength band while the surrounding tissue does not. NIR-sensitive optical sensors are used to capture the image of the vein pattern. Examples are shown in figure 1.

The diameter of the blood vessels and their depth inside the body are limiting factors for the feature extraction process, which is described in the next section.

1.2. Feature Extraction

The feature extraction process is starting from a captured vein image sample. The pattern, the abstract structure of the veins, has to be extracted from the noisy vein image, a sample is shown in figure 2 taken from [7]. Features used for comparison are localized in the extracted vein skeletal pattern. Various algorithms are published based on line track-

ing [6], local thresholding [11], curvelets and neuronal networks [14], as well as maximum curvature points [7]. The algorithms of those publications all have the same outcome: a vein pattern for authentication purposes, which needs to be stored and processed. Example vein patterns are shown in figure 2 and 3. Why this may conflict with privacy protection directives is explained in the following section.

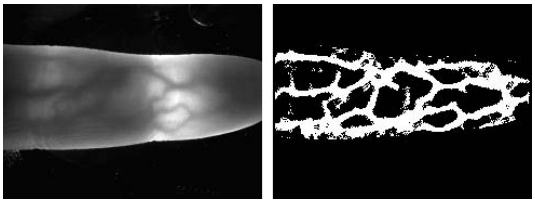

Figure 2. Finger vein image and corresponding vein pattern based on maximum curvature points [7]

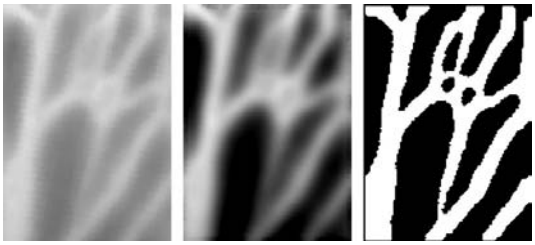

Figure 3. Original vein image, after noise reduction and after local thresholding [11]

2. Privacy Concerns

Biometric systems are exposed to privacy concerns since there is an intrinsic link between the stored biometric template and the person it originates from. The advantages, offered by biometric authentication, are inverted when the stored data is stolen – the data theft becomes an identity theft. The problem is that you can not simply change a biometric characteristic like a key or password. Revocation and reissuing of a specific biometric characteristic is not feasible in common biometric systems.

Another privacy related aspect is cross-matching: if the same modality is used in different application contexts (access control, financial services, eCommerce services etc.), a profile can be constructed linking the stored data in different databases.

Furthermore biometric data can contain information about physical traits of humans - the risk of storing medical or health related information is therefore always existent.

2.1. Medical Issues with Vein Patterns

Diseases related to the cardiovascular system are among the main causes of death in the world [12]. Vein patterns used in biometric systems could also reveal information about the medical state of a human. The authors found several examples of vascular diseases related to the finger or the hand.

In general the diameter and the position of the veins are of medical interest. An example is thrombosis, where a blood clot (thrombus) blocks the blood flow in the cardiovascular system. Diseases changing the position and the structure of the vein network affect all feature extraction methods resulting in a vein pattern.

After a literature survey on radiological publication two examples of disease patterns are found that change the appearance of the hand vein pattern: arteriovenous malformation (AVM) is a congenital disorder where veins and arteries are connected in an abnormal way. A contrast-enhanced radiographic example is given in figure 5 taken from [4]. Another abnormality is the hypothenar hammer syndrome (HHS) which is also identifiable throughout the vein pattern of the hand (Fig.4 taken from [4]). HHS is a thrombosis of the superficial palmar arch of the ulnar artery and is caused by repeated mechanical force, as seen in fighting sports or the work with vibrating tools (e.g. a hammer).

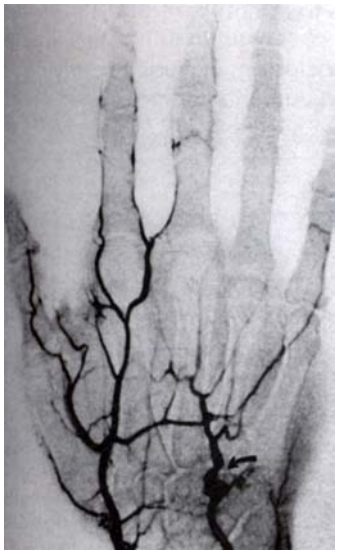

Figure 4. Hypothenar hammer syndrome [4]

Those examples illustrate that vascular image data may contain health related information. Since the ISO standard for the vascular interchange format [3] uses those image based vein patterns, the medical information is still available in the stored references. The next section describes the legislative view on this special kind of personal data.

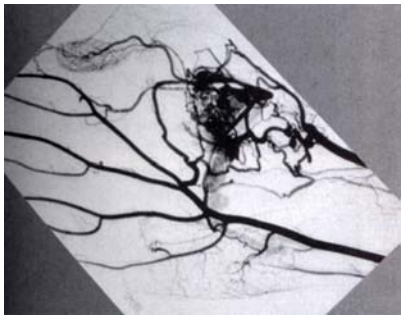

Figure 5. Arteriovenous malformation [4]

2.2. Legislative Regulations

Special acts are implemented to secure the privacy rights of individuals. An introduction to data protection and biometric systems is given by Meints in [5]. In the following the relevant European and the Norwegian regulations are introduced.

The European data privacy principles are formulated with the "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data". In article 2 important terms are defined, interesting in our context are the first two definitions (a) and (b) of 'personal data' and its processing, which covers the usage in biometric systems. Article 8 specifies the handling of health related personal data. Generally member states should prohibit processing of this "special category of data". Article 6 (c) defines the need for the adequateness and non-extensiveness of the data in relation to its purpose.

The Norwegian "Act of 14 April 2000 No. 31 relating to the processing of personal data (Personal Data Act)" follows the European directive with some differences, nevertheless the definitions for 'personal data' and 'processing of personal data' also apply to our context. Section 2 defines health related information as 'sensitive personal data'.

These two regulations both challenge the common praxis in biometric systems based on vein recognition. As shown, vein data may contain medical, health-related information and therefore the principle of adequateness in a biometric authentication system is violated. Since the usage is prohibited by law, the stored data has to be transformed into a non-revealing form.

Current research on template protection [9, 10, 15] show one possible way to construct biometric authentication systems that satisfy the regulations. After the enrolment of a data subject, no information about the original biometric sample is revealed. Even for the comparison of templates the biometric information does not need to be revealed as needed in the classical case of encrypted databases.

3. Experiments

The following experiments show a solution to the data privacy challenges associated with vein images. A feature extraction algorithm based on local thresholding [11] is used to extract vein patterns from backhand vein images. These feature vectors are then further processed to be compliant with the helper data scheme for privacy enhancement [10]. Finally the performance based on the raw feature vectors and the processed versions is evaluated.

3.1. Database

The database was gathered by the Nanyang Technological University in Singapore [8] and consists of a near infrared and a far infrared part. The near infrared part, that is used in the experiments, contains 122 data subjects with 3 samples for each hand taken in one session with a reflection technique. The resolution of the gray-scale images is 644 x 492 pixels.

3.2. Feature Extraction

Features are extracted directly from the histogram-optimized images, the local thresholding algorithm [11] was used. The value of each pixel in the feature space f is affected by the surrounding pixels of the original image I: if the pixel intensity exceeds the mean value of a defined area around the actual one it is set to 255, otherwise 0.

$$f(x,y) = \begin{cases} 255, & \text{if } I(x,y) > \mu_{I(x,y)} \\ 0, & \text{otherwise} \end{cases}$$

The parameter μ was set to the mean of 30 x 30 pixel blocks, after this step 50 pixels were cropped from each edge of the feature image due to irregularities at the boarders. The resulting feature image has a resolution of 544 x 392 pixels. To further decrease the variance caused by noise, the images were resized with factor 20 to a size of 28 x 20 pixels. Resulting feature vectors have 560 elements. figure 6 shows the three different steps towards the feature vector.

3.3. Privacy Enhancing Mechanism

Among the several privacy enhancing mechanisms is the helper data scheme [10] which is sketched in figure 7. The scheme can process any biometric data, as long as the created feature vectors have the same dimension. Basically the biometric data is merged with a random secret into a secure form during the enrolment. Beforehand feature vectors are binarized, reliable bits are extracted. These bits are combined using boolean exclusive or (XOR) with an error encoded random and secret bit vector. The hash of the secret,

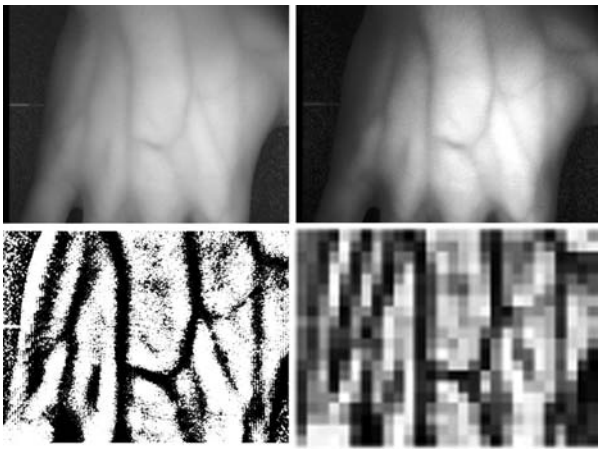

Figure 6. Processing steps of the feature extraction algorithm: normal image, histogram optimized, local thresholding applied, resized feature image.

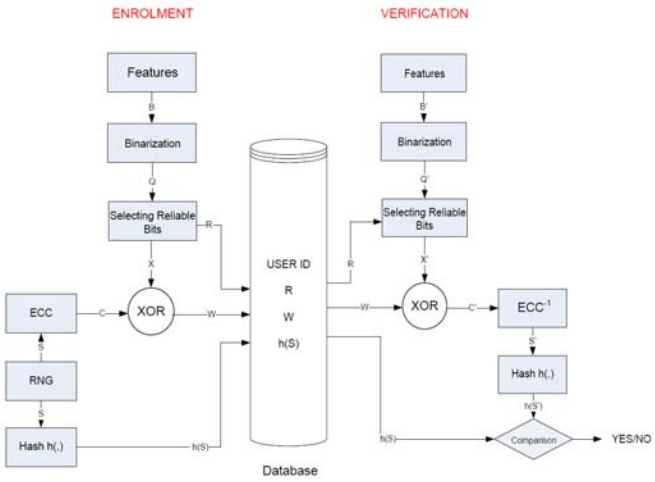

Figure 7. Block diagram of the helper data scheme.

the indexes of the reliable bits (helper data) and the secured template are stored in the database.

During the verification the data belonging to the biometric claim is loaded. The probe feature vector is binarized again, bits are extracted using the stored reliable indexes. The boolean XOR is applied to the stored secured template and the probe reliable bit vector. The hash of this result is compared with the stored hash. If the error decoder is able to correct the bit flips caused by noise, an identical hash value is produced, the biometric claim is verified.

In the experiment a reliable bit vector is generated from an intermediate (unprotected) feature vector in the following manner:

1. Center the feature vectors around its mean (subtract the mean from each feature vector).
2. Map every value to binary 1 exceeding zero (larger than the mean), the rest to binary 0 (smaller than the mean).

In this way, the bits are equally distributed and statistically independent. The binarized feature vectors consist of 560 bit values each.

The reliable bit extraction block estimates the optimized reliability and discrimination power R of every component k for each subject i. Assuming Gaussian distributed components, the following formula can be used:

$$R_{i,k} = \frac{1}{2}(1 + Erf(\frac{\mu_{intra} - \mu_{inter}}{\sqrt{2v_{i,k}}}))$$

Here the variance (v), the intra-class mean (μ_{intra}) and the inter-class mean (μ_{inter}) has to be computed in advance. Erf stands for the Gaussian error function.

The components having the highest reliability value R are selected as candidates for the reliable bit vector. Since those vectors are used in the helper data scheme, performance estimations can be computed for the whole system taking into account the error correction capability of the ECC-block.

4. Results

4.1. Unprotected Features

The histograms of genuine and imposter attempts are shown in figure 8 for the unsecured 560-dimensional feature vectors. A good performance with an equal error rate at 0.55% is measured (figure 11) using the 1-correlation as distance metric.

4.2. Binarized Features

The extracted binary feature vectors perform at about the same level as the unprotected feature vectors. The distribution using the Hamming distance is shown for genuine and imposter attempts in figure 9. The DET is shown in figure 11, the equal error rate is around 0.55%.

4.3. Protected Features

The reliability estimation of the components leads to a mean value of 520 perfectly reliable bits per data subject. When selecting the the 255 most reliable bits from the binarized feature vectors, performance is increasing. The reliable bit estimation was performed on 2 out of 3 samples per data

subject. If training and validation sets are strictly separated, no FNMR can be computed. Considering the 2 samples taken for the reliability estimation as a stored reference for testing, no false non match was measured.

For a threshold in the Hamming distance of 0.18 no false match was measured as well. The distribution of genuine and imposter attempts are shown in figure 10. Those reliable bit vectors could be used in the helper data scheme. In the case that the error correction block is able to correct 18% bit errors, imposters and genunies are perfectly separated and the privacy is protected. The DET-curve is also plotted in figure 11.

5. Conclusions

In this paper the need for privacy enhancing technologies when processing vein data is shown: disease patterns are presented which can be extracted from vein patterns. No other publication is known to the authors that deals with this delicate topic. One possible solution is the application of a privacy enhancing scheme. The helper data scheme satisfies the regulations, because no information about the biometric characteristic can be extracted from the stored secure template and the helper data.

The experimental section describes a feature extraction algorithm based on local thresholding and the binarization and reliable bit extraction block of the helper data scheme. For the first time a privacy enhancing scheme was applied in the context of back-hand vein data. The results are remarkable, a robust authentication system guaranteeing privacy can be constructed for this specific database of 122 data subjects.

It has to be mentioned that the performance can only be reached in laboratory environments – the database was taken in only one session, the variation in the original data is therefore fairly low. To confirm the results of the experiments a large-scale database is needed and in preparation. Further research is needed in the field of normalization of vein images. To satisfy security constraints in the helper data scheme long reliable bit features are needed, multimodal, multi-spectral solutions could be the way to go.

References

[1] J. Hashimoto. Finger vein authentication technology and its future. *VLSI Circuits, 2006. Digest of Technical Papers. 2006 Symposium on*, pages 5–8, 2006.

[2] International Biometrics Group (IBG). *Comparative Biometric Testing Round 6 Public Report*, Sep 2006.

[3] International Organization for Standardization (ISO), ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 19794-9:2007 Information Technology - Biometric Data Interchange Formats - Part 9: Vascular image data*, 2007.

[4] La Berge et. al. *Interventional Radiology Essentials*. Lippincott Williams and Wilkins, 2000.

[5] M. Meints, H. Biermann, M. Bromba, C. Busch, G. Hornung, and G. Quiring-Kock. Biometric systems and data protection legislation in germany. In *IIH-MSP '08: Proceedings of the 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 1088–1093, Harbin, China, 2008. IEEE Computer Society.

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

[7] N. MIURA, A. NAGASAKA, and T. MIYATAKE. Extraction of Finger-Vein Patterns Using Maximum Curvature Points in Image Profiles. *IEICE Trans Inf Syst*, E90-D(8):1185–1194, 2007.

[8] Nanyang Technological University. *http://www.ntu.edu.sg*, mar 2009.

[9] N. Ratha, J. Connell, and R. Bolle. Enhancing security and privacy of biometric-based authentication systems. *IBM Systems Journal*, 40, 2001.

[10] P. Tuyls, A. M. Akkermans, T. Kevenaar, G.J.Schrijen, A.M.Bazen, and R.N.J.Veldhuis. Practical biometric authentication with template protection. In *Audio- and Video-Based Biometric Person Authentication*, volume 3546 of *Lecture Notes in Computer Science*, pages 436–446. Springer Belin / Heidelberg, June 2005.

[11] L. Wang, G. Leedham, and D. Cho. Infrared imaging of hand vein patterns for biometric purposes. *Computer Vision IET*, 1(3-4):113–122, December 2007.

[12] World Health Organization. *The world health report 2004 - changing history.*, 2004.

[13] T. Yanagawa, S. Aoki, and T. e. a. Ohyama. Human finger vein images are diverse and its patterns are useful for personal identification. 2007.

[14] Z. Zhang, S. Ma, and X. Han. Multiscale feature extraction of finger-vein patterns based on curvelets and local interconnection structure neural network. In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, pages 145–148, Washington, DC, USA, 2006. IEEE Computer Society.

[15] X. Zhou, T. Kevenaar, E. Kelkboom, C. Busch, M. van der Veen, and A. Nouak. Privacy enhancing technology for a 3d-face recognition system. In *BIOSIG 2007: Biometrics and Electronic Signatures*, volume P-108 of *Lecture Notes in Informatics*, pages 3–14. GI-Edition, July 2007.

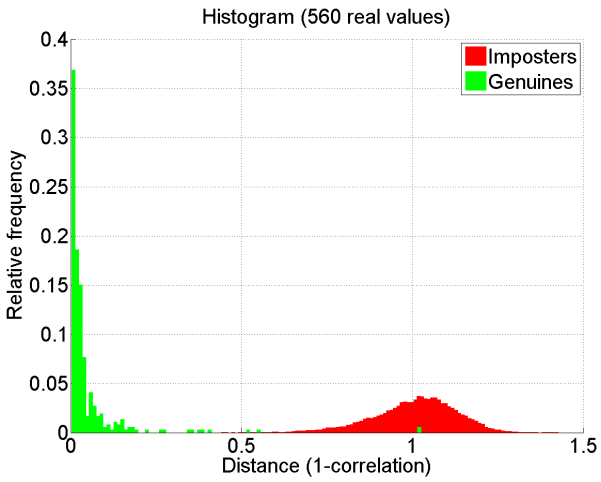

Figure 8. Histogram of genuine and imposter attempts (unprotected 560-dimensional features).

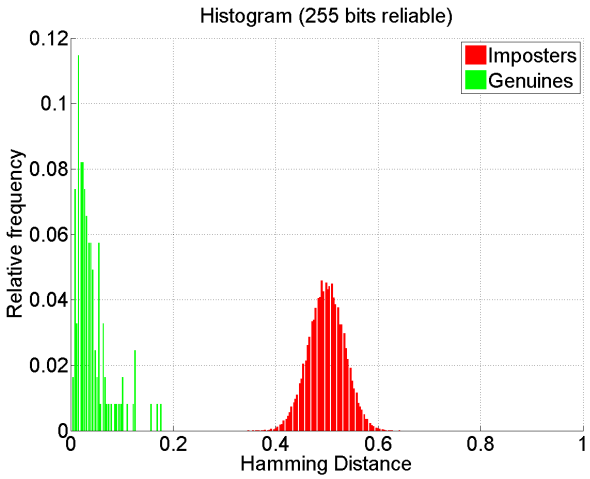

Figure 10. Histogram of genuine and imposter attempts (reliable binarized features).

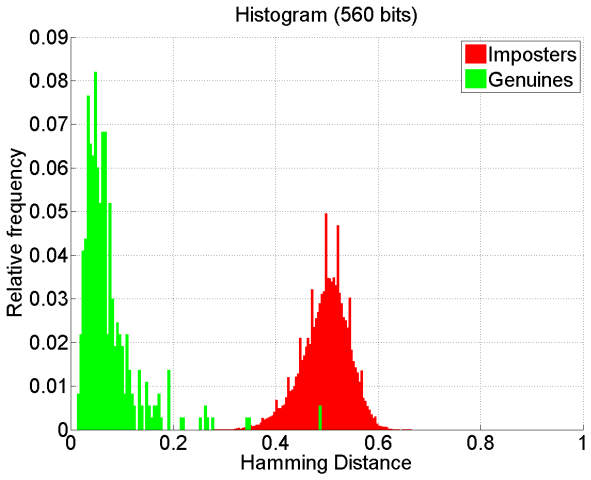

Figure 9. Histogram of genuine and imposter attempts (560-bit features).

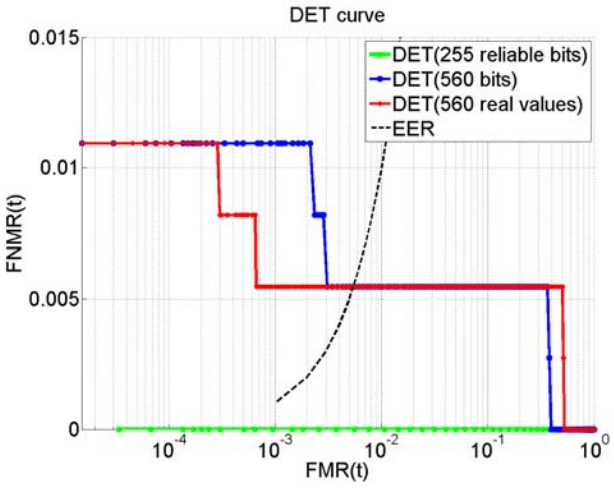

Figure 11. DET curves of the different feature vectors.