A Study of Similarity between Genetically Identical Body Vein Patterns

Hengyi Zhang[#1], Chaoying Tang[#*2], Xiaojie Li[#3] and Adams Wai Kin Kong[#4]

[#]School of Computer Engineering
Nanyang Technological University
Singapore
{[1]hengyi, [3]xli16, [4]adamskong}@ntu.edu.sg

[*]College of Automation
Nanjing University of Aeronautics and Astronautics
Nanjing, P.R. China
[2]cytang@nuaa.edu.cn

***Abstract*— Vein patterns have been used in commercial biometric systems for many years and are recently considered for criminal authentication. Understanding the similarity between genetically identical vein patterns is important, especially when using them in legal cases involving identical twins. Vein patterns sharing the same Deoxyribonucleic acid (DNA) sequence are generally regarded as vein patterns with maximum similarity. If they are completely distinguishable, it implies that the uniqueness of vein patterns is high. Though the genetic dependence of other biometric traits, including fingerprints, faces, palmprints, and irises, have been studied, genetically identical vein patterns have not been studied systematically. With the help of an automatic vein pattern matching algorithm, this paper analyzes and measures the similarity between genetically identical vein patterns. 234 genetically identical forearm pairs and 204 genetically identical thigh pairs were collected for this study. Experimental results indicate that genetically identical vein patterns have extra similarity, but they are distinguishable.**

Keywords*— Identical twins, vein patterns, biometrics, criminal and victim identification

I. Introduction

Medical reports point out that the current rate of twin births in many industrialized countries is around 3% [1-2]. Twins can be classified into two types, dizygotic and monozygotic. Dizygotic twins are evolved from two different fertilized eggs and have different DNA sequences. Monozygotic twins, also called identical twins, are from a single fertilized egg splitting into two individual cells, which finally develop into two individuals. Thus, monozygotic twins have the same DNA sequence. DNA contains all genetic information required to create an organ of a species. The mapping from DNA to an organ is very complex. The final products are influenced by not only genetic information, but also other factors, such as living styles, diets and climate. In spite of this, some biometric traits of monozygotic twins are still very similar. Studying biometric traits of monozygotic twins is an important topic because they are expected to have maximum similarity. If there is no or only limited extra similarity between them, the corresponding biometric trait is believed to be highly unique. In addition, if the biometric traits are used for criminal and victim verification in legal cases involving monozygotic twins, their genetic dependence has to be understood completely. Genetic dependence of many biometric traits, including fingerprints, faces, irises, and palmprints, have been studied [2-5].

Vein recognition systems have been available for commercial applications. Most of them utilize near infrared (NIR) to capture vein patterns from palms, fingers or back of the hands in controlled environments. Recently, vein patterns are proposed for identifying criminals and victims in evidence images, where their faces are not observable, but partial skin is available for identification, because methods have been developed to visualize vein patterns hidden in color images [6-7]. These evidence images can be collected from legal cases of child pornography, masked gunmen, and violent protests, where criminals always cover or hide their faces. Fig. 1 shows four color skin images, the corresponding vein patterns in the NIR images, and the vein patterns given by the visualization method [7]. It should be emphasized that the visualization method only took the color images as inputs, without using any information from the NIR images. These methods demonstrate that vein patterns have potential for criminal and victim identification in some legal cases.

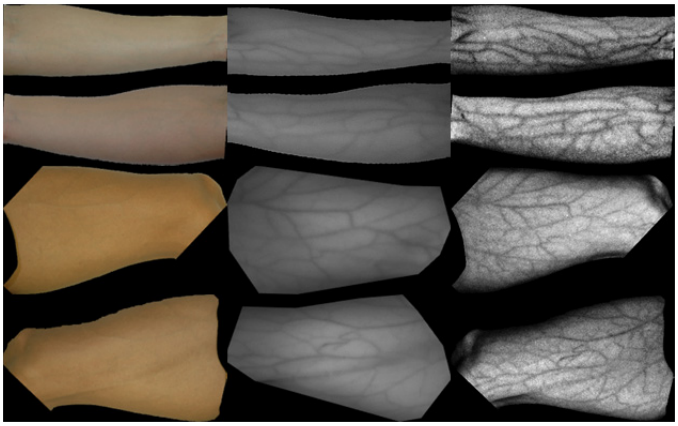


Fig. 1 Results from the vein visualization method [7]. Each row shows vein patterns from the same subject. The first column is color images taken by a consumer camera. The second column is the corresponding NIR images. The third column is the results given by the vein visualization method [7], which only took images in the first column as inputs.

To point out the importance of development of new biometric traits, such as vein patterns, for criminal and victim identification, some statistics of child pornography are given. In the U.S., the National Center for Missing and Exploited Children (NCMEC) has reviewed and analyzed more than 90 million child pornographic images over the past ten years [8]. The CyberTipline, an NCMEC's congressionally mandated

service for victims of child sexual exploitation operated in partnership with the Federal Bureau of Investigation (FBI), the U.S. Immigration and Customs Enforcement (ICE) and the U.S. Department of Justice, has received more than 1.9 million reports of suspected child sexual exploitation since it was launched in 1998 [8]. With the improvement of digital imaging technology, more and more child pornographic images have been posted on the internet and it is the driving force for the increase of child sex offenses. Anyone who possesses, makes, prints, publishes, distributes, sells, or imports child pornography commits a criminal offense. However, due to a lack of effective identification technology, the number of prosecuted suspects was much less than the number of suspects [47]. Pedophiles are usually careful not to show their faces in the images for fear of identification. Development of effective identification technology based on partial skin in the images can greatly enhance the ability of prosecutors to achieve convictions. In addition to child pornography, many criminals in violent protests, including the recent Thailand, Rome, and English riots, wear masks to cover their faces for fear of possible identification and prosecution, but a good portion of their skin is observable in evidence images. Therefore, development of new biometric traits based on non-facial skin is highly demanded.

Tattoos and skin marks have been used to identify criminals and victims in images where only partial skin is observable. However, they are ineffective in some of the cases mentioned before because skin exposed in evidence images may neither have unique tattoos nor enough skin marks [9-10]. Some people, including gang members, can have the same tattoo pattern, while many do not have tattoos. Skin marks suffer from the similar problems. Although vein patterns have been used in commercial systems and are now being considered as a tool for forensic investigation, according to our best knowledge, the similarity between genetically identical vein patterns, e.g., from monozygotic twins, has not been studied systemically. The aim of this paper is to analyze and measure the similarity between genetically identical forearm and thigh vein patterns for mainly criminal and victim authentication. 234 genetically identical forearm pairs and 204 genetically identical thigh pairs were collected for this study.

The rest of this paper is organized as follows. Section II summaries previous research results. Section III presents a vein pattern matching algorithm for this study. Section IV reports and analyzes the experimental results on the forearms and thighs. Section V offers some conclusive remarks.

II. Related work

A. Related work on vein biometrics

In this subsection, the previous research on vein recognition is summarized and in the next subsection, the previous research on different biometric traits of monozygotic twins is reported. A vein pattern is formed by blood vessels that lie between the skin and muscle covering most parts of human bodies. It is regarded as a stable biometric trait with strong immunity to forgery [11]. The vast network of blood vessels is believed to be "hardwired" into the body at birth, and remains relatively unaffected by aging, except for venous diseases, which can cause spider and varicose veins [12]. In addition, as the blood vessels are hidden underneath the skin and are almost invisible, they are much harder to be duplicated comparing with other biometric traits.

(a) (b) (c)

Fig. 2 Vein patterns commonly used in commercial systems: vein patterns in (a) a back of the hand, (b) a palm and (c) a finger.

Vein patterns in the back of the hands, palms and fingers are commonly used in commercial systems. Figs. 2(a)-(c) show respectively vein patterns in a back of the hand, a palm and a finger collected in near infrared environments. Most of these vein verification systems rely on high quality images captured in controlled infrared imaging environments with great cooperation from users. Hand veins were proposed as a biometric trait in 1990s because of the high user acceptance and the simple acquisition process. Since then, hand veins have been extensively studied and many recognition methods have been developed for them.

In 1991, MacGregor and Welford [13] presented an authentication system called Veincheck using vein patterns acquired from the back of a clenched fist under near infrared light. Its commercial version was developed by Cambridge Consultants Ltd and British Technology Group and announced in September 1993 [14]. Wang et al. [15] proposed an automatic recognition method using vein patterns in the back of the hand captured in a far infrared environment. Their method utilizes minutiae features and modified Hausdorff distance to measure the similarity between two vein patterns. Subspace methods based on independent component analysis (ICA) [16], and locality preserving projection (LPP) [17] which uses a multimodal method to combine a palmprint image and a palm vein image into a single image were developed. Neighborhood matching radon transform (NMRT), Hessian phase [18], and scale invariant feature transform (SIFT) were also studied for hand [19] and palm vein [20] recognition. Line based methods utilizing blood vessel patterns extracted by Gabor filters [21], cutoff Gaussian filters [22], or matched filters [23] were also proposed.

The finger vein matching method proposed by Miura et al. [24] uses transmitting infrared light to capture finger vein images. Their method utilizes a line tracking scheme starting from various positions to extract finger vein patterns from the images and template matching is adopted to compare finger vein patterns. Miura et al. also used local maximum curvature to improve the robustness of their vein pattern extraction method [25]. Wu et al. [26] employed radon transform, neural network and support vector machines (SVM) for finger vein recognition. A method using curvelet to extract finger vein information and a back-propagation neural network as a classifier was also developed [27]. Lee and Park [28] used a point spread function to restore finger vein images and reported

significant improvement in terms of identification performance. Kumar and Zhou [29] proposed a method utilizing a score-level fusion strategy to combine information from finger vein images and fingerprints together.

Currently, vein recognition systems rely on infrared and laser imaging technologies to capture high quality blood vessel patterns from palms, wrists, and fingers, where the skin is thin, for commercial applications, e.g., access control [15, 30]. In addition to commercial applications, vein patterns have recently been proposed for criminal and victim identification because algorithms have been developed to visualize body vein patterns (e.g., forearm and thigh vein patterns) hidden in color images, which can be evidence images exposing partial skin of criminals or victims [6, 7]. There are two types of vein visualization algorithms. One is based on skin optics and the other is based on mapping RGB values to NIR intensity [6, 7]. A matching algorithm was also developed to automatically extract and match forearm and thigh vein patterns [31]. Because NIR has a higher penetration capability and is expected to offer the best-quality vein patterns, NIR images were selected to compare the results given by the algorithms [6, 7]. The experimental results indicated that the algorithms can visualize vein patterns hidden in color images for personal identification and the matching performance from visualized vein patterns is comparable with that from NIR images.

B. Related work on monozygotic twins' biometrics

In the early days, biometric researchers concentrated mainly on the development of robust and accurate recognition systems for normal people. They developed methods for handling lighting, pose and viewpoint variations, and extracting and matching features effectively. The issue of distinguishing monozygotic twins was a bit neglected. Some early studies on monozygotic twins' voices, ears, and faces were reported in [2, 32-34]. Most of them were based on manual comparison. With the advance of computing techniques and automatic matching methods, researchers are empowered to conduct more throughout studies on genetically identical biometric traits.

Doddington [33] observed that monozygotic twins' sound is quite similar to each other, but their spectrograms look very different. He showed that machines can outperform human beings when classifying speakers of monozygotic twins. Ariyaeeinia et al. [35] reported some encouraging results on a dataset with 49 pairs of monozygotic twins. Their experiments had two configurations: general cases where imposters were anyone in the database, and twin cases where imposters were his/her twin sibling. The reported equal error rates were 0.5% for the general experiment and 1% for the twin experiment. Later, Loakes studied voice patterns of monozygotic twins for forensic analysis [36]. He demonstrated that variations of voice patterns are governed by speakers' physical dimensions and their learned behavior. Despite similarity and correlation in phonetic outputs between speakers, specific speaker behavior was not the same even between monozygotic twins.

In addition to voice patterns, genetically identical faces, irises, ears, and palmprints have also been studied. Faces of monozygotic twins are easily misrecognized. Bronstein et al. [37] and Vijayan et al. [38] proposed 3D face recognition methods to address the challenge of monozygotic twins. Zhang et al. utilized facial motions to separate twin pairs [39]. Research on face recognition between identical twins was conducted on different aspects. Klare et al. compared different features, including MLBP, SIFT, and facial marks, from monozygotic twins [40]. Philips et al. performed cross-year matching with different lighting conditions to analyze similarity between faces of monozygotic twins [41]. Facial expressions and facial aging features were employed by Le et al. [42]. Facial marks were also utilized to distinguish monozygotic twins by Srinivas et al. [43]. With manual annotation and auto detection, their experiments demonstrated that facial marks are a useful biometric trait and their geometric distribution is genetic dependent.

Zhang et al. [44] reported that voice recognition could perform better than face recognition when classifying monozygotic twins. They also fused voice and face features together. Kong et al. [5] performed an automatic palmprint matching experiment on 53 pairs of monozygotic twins and reported that principal lines in twins' palmprints are highly correlated, but non-genetic dependent features are rich enough to differentiate them. Similar observations were also reported by Jain et al. [3] for fingerprints. Though certain appearance similarity can be found in fingerprints of monozygotic twins, current fingerprint matching algorithms are still able to separate them. Not surprisingly, genetically identical irises were also tested. Daugman and Downing found no correlation [4]. The first report on monozygotic twins' ear comparison was conducted manually by Iannarelli in 1989 [34]. An accuracy of 92% was obtained. In addition to physical biometric traits, behavioral biometric traits of monozygotic twins were also examined. Mohd-Isa et al. studied gaits of 12 pairs of monozygotic twins [45]. Sun et al. [2] used multiple biometric traits, including fingerprints, faces, and irises to identify monozygotic twins. They found that genetically identical irises and fingerprints cause slight performance degradation, but the tested face recognition method has difficulty separating face images from monozygotic twins. Further experiments on their multimodal biometric system combining fingerprints and irises demonstrated its great separation capability on monozygotic twins. Zhang et al. [44] also proposed a multimodal biometric system, but they used voices and faces to separate monozygotic twins. According to our best knowledge, no one studied genetically identical vein patterns for forensic verification before.

III. An Algorithm for Measuring Similarity between Body Vein Patterns

For the completeness of this paper, the algorithm [31] utilized to measure similarity between body vein patterns is presented in this section. This algorithm uses a series of preprocessing steps to extract vein patterns from NIR images and a three-step matching scheme to measure similarity between body vein patterns. In the preprocessing steps, a Gabor filter bank is used as a major tool to extract potential vein segments and an enhancement scheme is used to increase the contrast between authentic veins and generic skin. Then, thresholding and thinning methods are applied to obtain skeletonized vein patterns. They are sampled into points and their coordinates are used to describe the vein patterns. The

three-step matching scheme employs rigid and non-rigid point registration methods to align two input point sets, which are called data set and model set. Their similarity score is calculated based on the average distance of corresponding points. A more detailed description of this algorithm is given below. Sections III.A and III.B respectively present the preprocessing steps and the three-step matching scheme.

A. Preprocessing

(1) A region of interest (ROI) in an input NIR image is first selected manually for the experiment. The input image is cropped to enclose the selected area and rotated so that the principal axis of the ROI is horizontal. The ROI of each input image is stored in a mask image for latter usage. Note that the preprocessing steps (2) to (4) stated below are all applied to the whole cropped image to avoid border effect. The column (a) in Fig. 3 shows the ROIs of two NIR forearm images and two NIR thigh images.

(2) Contrast limited adaptive histogram equalization (CLAHE) is performed on the image to enhance its local contrast. The image is divided into 8×8 blocks and the histograms of the blocks are stretched to a uniform distribution with limited contrast. The column (b) in Fig. 3 shows the enhanced results of the column (a).

(3) A filter bank composed of the real parts of 16 zero DC Gabor filters with 2 scales and 8 orientations is utilized to extract vein patterns. The input image is convolved with the filters. Let the output from the filter with the orientation θ_k and the scale λ_m be $F_{\lambda_m,\theta_k}(x,y)$. The filter outputs are combined to form a response map $R(x,y)$, which is defined as

$$R(x,y) = \max_{m,k} F_{\lambda_{mk},\theta_k}(x,y)/P_m(x,y), \qquad (1)$$

where P_m is the local image power around the point (x,y) and is dependent on the Gabor filter scale m. The orientation map is defined as

$$O(x,y) = \arg_{\theta_k} \max_{m,k} F_{\lambda_{mk},\theta_k}(x,y)/P_m(x,y), \qquad (2)$$

The columns (c) and (d) in Fig. 3 show respectively the orientation maps and response maps from the forearms and thighs.

(4) Connected pixels with the same value in the orientation map are grouped into components and those smaller in size than a threshold are regarded as noise and removed. If the orientation and spatial distances between two components are smaller than some thresholds, they are linked together. This operation is performed on all component pairs and an adjacency matrix, whose elements indicate the connection of component pairs, is obtained.

(5) Iteratively tracing the connections in the adjacency matrix, multiple components are linked together. The linked components form a group. The pixels in the response map are enhanced or suppressed based on their group sizes in terms of the number of pixels. The columns (e) and (f) in Fig. 3 illustrate the groups and enhanced response maps.

(6) The enhanced response maps are masked by the ROI and further binarized and skeletonized. Finally, point sets are used to represent the vein patterns for matching. For forearms, points are also sampled from boundaries. Thigh boundaries are not reliable because of shorts. Thus, they are not sampled. The last column in Fig. 3 shows the point representation.

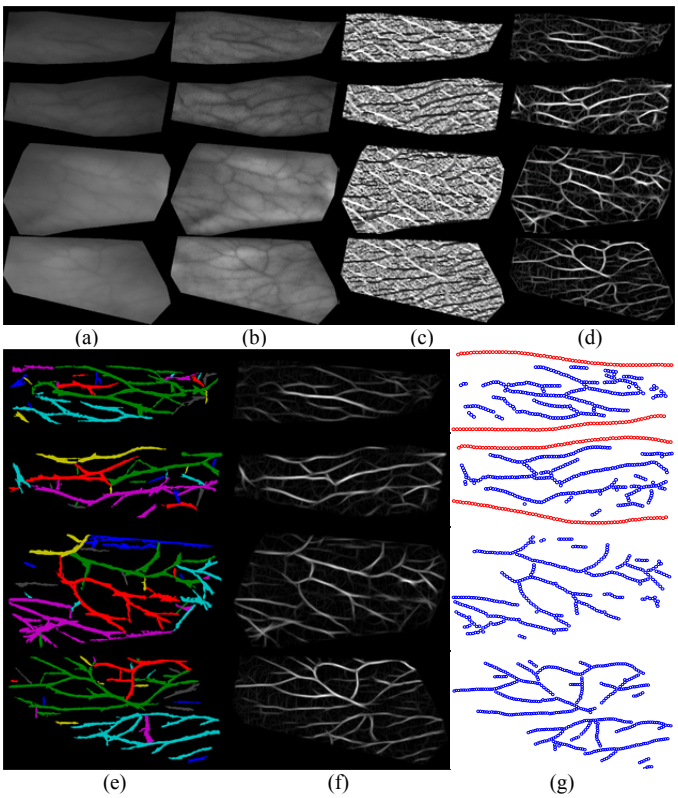


Fig. 3 The blood vessel pattern extraction processes. Four sets of preprocessing results, including two NIR forearm images and two NIR thigh images are presented. Column (a) is the input NIR images; column (b) is the results given by local brightness adjustment; column (c) is the orientation maps generated by the Gabor filters; column (d) is the response maps; column (e) is the maps of grouped vein components (different colors indicating different groups); column (f) is the enhanced response maps; and column (g) is the point sets for matching.

B. Vein Pattern Matching

To measure similarity between vein patterns represented by the point sets, a three-step matching scheme reported in [31] is used. In order to better understand and compare the experimental results, the matching algorithm is described in more details. A set of notations is given first for clear presentation. Let M and D be two point sets from input images I_M, and I_D. M and D are called model and data patterns, respectively. Each point in M, e.g., $(x,y) \in M$, is associated with a response value $R_M(x,y)$ and an orientation value $O_M(x,y)$, where R_M and O_M are the response and orientation maps of I_M, respectively. For forearms, the point sets representing their boundaries are denoted as M_{BD} and D_{BD}. Transformations of the point sets are denoted as $T(\bullet,\tau)$, controlled by a parametric vector τ.

(1) When matching forearms, the boundaries M_{BD} and D_{BD} are first used to remove points outside the overlapping area of M and D. A rigid transformation T_R is used to describe the deformation between M_{BD} and D_{BD}. The rigid Coherent Point

Drift (CPD) method [46] is used to determine a parametric vector τ_1. Fig. 4(c) shows two boundaries and their alignment result. The point set M is transformed into $M_{t1} = T_R(M, \tau_1)$. Then, the overlapping area B of M_{t1} and D is determined. Points that are not in B are regarded as outliers and removed to increase robustness. The pruned point sets are denoted as M_1 and D_1. If the numbers of points in these two sets are less than a threshold, they are assumed to be from two different blood vessel patterns. Fig. 4(d) shows two point sets after outlier removal. Because thigh boundaries in our database are unreliable, we set $M_1 = M$ and $D_1 = D$ when matching them.

(2) The matching scheme removes outliers and noise inside M_1 and D_1. The rigid transformation $\mathrm{T_R}$ is used again to roughly align M_1 and D_1. Note that M_1 is regarded as a model set in this alignment. Using the rigid CPD method and the resultant parametric vector τ_2, the correspondence c_2 between M_1 and D_1 can be obtained. The point set $M_{t2} = T_R(\mathrm{M}_1, \tau_2)$ is roughly aligned with D_1, and the distances between the corresponding points in M_{t2} and D_1 are calculated. If the distances are very long, the corresponding points in M_{t2} are considered as outliers. All points in M_{t2} are sorted in ascending order according to their distances and the top 80% of the points in M_{t2} are retained. The remaining points are denoted as M_2. The same process is repeated to remove outliers in D_1. The rigid CPD method is applied to align D_1 and M_2. In this alignment, D_1 is considered as a model set and a transformed point set $D_{t2} = T_R(D_1, \tau_2')$ is computed. As with the previous case, only 80% of the points in D_{t2} are retained and distances between corresponding points in D_{t2} and M_2 are used as a selection criterion. The remaining points in D_{t2} are denoted as D_2.

(3) The final step is to match M_2 and D_2 with a non-rigid transformation. The non-rigid CPD method with a Gaussian kernel as a regularizer is used to match M_2 and D_2. With a non-rigid transformation T_N and the corresponding parametric vector τ_3 generated by the CPD method, $M_{t3} = T_N(M_2, \tau_3)$, and D_2 can be aligned optimally. As with the previous steps, a correspondence vector c_3 can be obtained. Note that multiple points in M_2 can correspond to the same point in D_2 and that data points with no correspondence are recognized as outliers.

The dissimilarity function is based on the distance between the two aligned point sets and the corresponding response and orientation differences. Let $m^i = (m_x^i, m_y^i)$ be the i^{th} point in M_{t3} and its corresponding point in D_2 be $d^{c_3(i)} = (d_x^{c_3(i)}, d_y^{c_3(i)})$. In addition, let the response and orientation values of m^i be $M_{t3}^R(m^i)$ and $M_{t3}^O(m^i)$, respectively and the corresponding response and orientation values of $d^{c_3(i)}$ be $M_2^R(d^{c_3(i)})$ and $D_2^O(d^{c_3(i)})$. The dissimilarity function is defined as

$$s(M, D) = |I|^{-1} \times \sum_{i \in I} \left\| m^i - d^{c_3(i)} \right\| \times \left| M_{t3}^R(m^i) - M_2^R(d^{c_3(i)}) \right| \times \Omega\left(M_{t3}^O(m^i), D_2^O(d^{c_3(i)})\right), \quad (3)$$

where $\Omega(M_{t3}^O(m^i), D_2^O(d^{c_3(i)})) = \min\{|M_{t3}^O(m^i) - D_2^O(d^{c_3(i)})|, 2\pi - |M_{t3}^O(m^i) - D_2^O(d^{c_3(i)})|\}$ and I is an index set containing points with dissimilarity values defined as $\left\| m^i - d^{c_3(i)} \right\| \times \left| M_{t3}^R(m^i) - M_2^R(d^{c_3(i)}) \right| \times \Omega\left(M_{t3}^O(m^i), D_2^O(d^{c_3(i)})\right)$ in the 80th percentile.

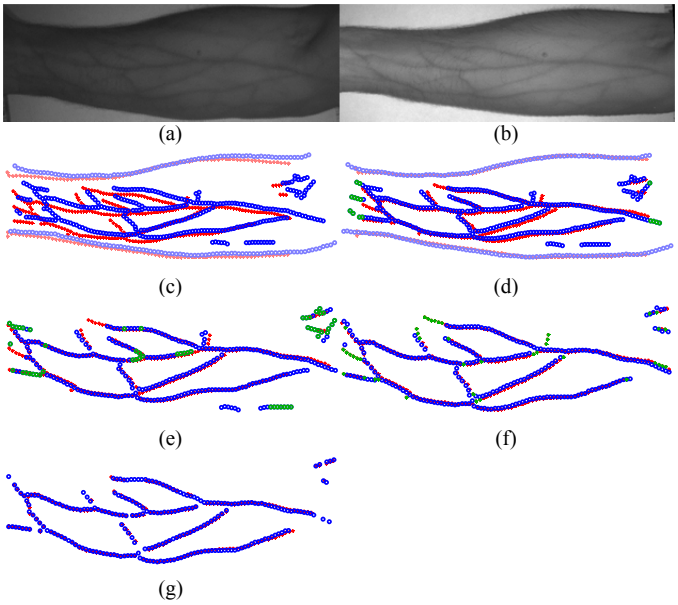


Fig. 4 Illustration of the vein matching scheme. (a) and (b) are two NIR images. (c) shows the points extracted from (a) and (b). Blue and red dots respectively from (a) and (b) represent model and data points. Note that boundary points are shown in lighter color. (d) illustrates M_{t1} and D_1. Outliers outside the overlapping area are marked in green. (e) illustrates M_{t2} and D_1. Green points are regarded as noise. (f) shows the alignment result of M_2 and D_{t2}, and (g) is the final non-rigid alignment result.

IV. Experimental Results

To measure and analyze similarity between genetically identical body vein patterns, NIR images of forearms and thighs were collected in two separate sessions with a gap of 11 days from 328 subjects. Vein patterns in NIR images were used, instead of vein patterns visualized from color images (e.g. the images in the third column of Fig. 1) in order to avoid errors given by the visualization algorithms [6-7]. Each subject in each session contributed four images, including the left and right inner forearms and inner thighs. In total, 8 images were collected from one subject. The captured images are in a grayscale bitmap format with a size of 1,024 × 768 pixels. Not all the collected images were used in this study because of their quality. Images collected from subjects with dense hair, tattoos and improper poses were removed manually. Images with motion blur and under bad illumination were also removed. Fig. 5 shows some removed images. 982 forearm images and 906 thigh images were retained. They were grouped into eight datasets summarized in Table I.

There are two approaches to collecting genetically identical body vein patterns for this study. One is to collect from monozygotic twins and the other is to collect from the right and left limbs of the same subject. The latter approach is more cost effective. Daugman et al. and Kong et al. employed it to examine similarity between genetically identical irises and palmprints, respectively [4, 5]. In this study, the latter approach is also employed. The images from the left limbs, including those in LF1, LF2, LT1, and LT2, were horizontally flipped. The flipped datasets are labeled as $\mathrm{LF1^F}$, $\mathrm{LF2^F}$, $\mathrm{LT1^F}$, and $\mathrm{LT2^F}$. Images in these datasets and their corresponding images

from the right limbs form genetically identical pairs. All genetically identical pairs in the datasets were matched. Because of the removal process, some images in the datasets did not form genetically identical pairs. In total, 234 subjects had both left and right forearm images and 204 subjects had both left and right thigh images in the datasets.

To study similarity between genetically identical body vein patterns, the images from the same limb were matched to generate genuine dissimilarity values (Eq. 3); the images from different subjects were also matched to generate imposter dissimilarity values, and the flipped images in $LF1^F$, $LF2^F$, $LT1^F$ and $LT2^F$ were matched with the corresponding images from the right limb of the same subject to generate virtual twin dissimilarity values. Note that the virtual twin dissimilarity values were generated from genetically identical vein pattern pairs. The distributions of the dissimilarity values and cumulative match characteristic (CMC) curves are employed in this analysis. Section IV.A reports the genuine and imposter distributions. Section IV.B presents the virtual twin distributions and the corresponding CMC curves and analyzes the results.

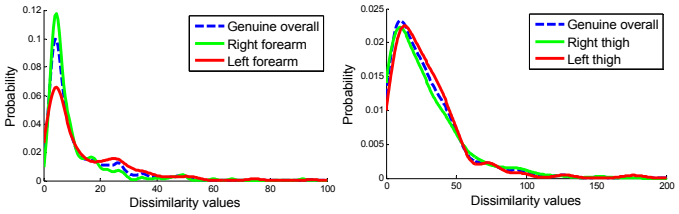
Fig.5 Examples of the removed images.

TABLE I. THE DATASETS USED IN THE EXPERIMENTS

Datasets	Abbreviation	Number of images
Left forearm images collected in the first session	LF1	241
Left forearm images collected in the second session	LF2	241
Right forearm images collected in the first session	RF1	250
Right forearm images collected in the second session	RF2	250
Left thigh images collected in the first session	LT1	228
Left thigh images collected in the second session	LT2	228
Right thigh images collected in the first session	RT1	225
Right thigh images collected in the second session	RT2	225

A. Genuine and Imposter Distributions

Figs. 6(a) and (b) show the distributions of the genuine dissimilarity values of the forearms and the thighs respectively. The vein patterns of the forearms are generally clearer than those of the thighs because the forearm skin is thinner. Thus, their genuine dissimilarity values are smaller than those of the thighs. The mean of the genuine dissimilarity values of the forearms is 13.13, while the mean of the genuine dissimilarity values of the thighs is 28.42. Fig. 6(b) shows that the distributions of the genuine dissimilarity values of the right and left thighs are very similar, but Fig. 6(a) shows a different result for the forearms. The distribution of the genuine dissimilarity values of the left forearms has a larger variance. 178 subjects in our database wore wrist straps on their left forearms. Consequently, their effective ROIs are smaller that those of the right forearms. It is the cause of the larger variance. Fig. 7(a) shows an NIR image of a left forearm with a wrist strap and Fig. 7(b) shows an NIR image of the corresponding right forearm without a wrist strap. The genuine dissimilarity values of the left forearms were separated into two groups: with and without wrist straps and their distributions are given in Fig. 8. Clearly, the variance of the distribution from the forearms without wrist straps is smaller than that from the forearms with wrist straps. For the virtual twin matching analysis, this factor had been taken into our consideration. The distribution of the imposter dissimilarity values are given in Fig. 9. The mean of the dissimilarity values of the imposter forearms is 44.74, while the mean of the dissimilarity values of the imposter thighs is 68.80. The imposter distributions from the right forearms (thighs) and the left forearms (thighs) are similar.



(a) (b)

Fig. 6 Distributions of dissimilarity values from (a) matching genuine forearm pairs and (b) matching genuine thigh pairs.

(a) (b)

Fig. 7 Left and right forearm images from the same subject. (a) the left forearm with a wrist strap and (b) the right forearm without a wrist strap. The wrist strap is covered for privacy protection.

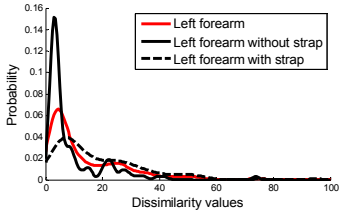
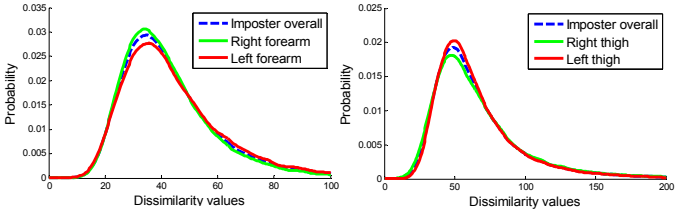


Fig. 8 The distributions of the dissimilarity values from matching left forarms in two sets, with and without wrist staps.



(a) (b)

Fig. 9 The distributions of the dissimilarity values from matching imposter forearm pairs and matching imposter thigh pairs.

B. Virtual Twin Distributions, CMC curves and Analyses

Fig. 10 compares the distributions of the genuine, imposter and virtual twin dissimilarity values. Their means and standard derivations are given in Table II. As mentioned before, some left forearm images have smaller ROIs because of the wrist straps. They were split into two groups: with and without wrist straps. In the 234 virtual twin forearm pairs, 112 pairs belong to the group with wrist straps and the rest belong to the group without wrist straps. Their corresponding distributions are given in Fig. 10(a). Fig. 10 and Table II show that for all the cases, the means of the virtual twin distributions are smaller than the means of the corresponding imposter distributions. They indicate that genetically identical vein patterns have extra

similarity.

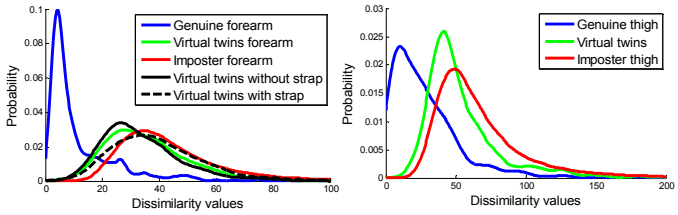


Fig. 10 A comparison of genuine, imposter and virtual twin distributions: (a) distributions from the forearms and (b) distributions from the thighs.

TABLE II. MEANS μ AND STANDARD DEVIATIONS σ OF THE GENUINE, IMPOSTER AND VIRTUAL TWIN DISTRIBUTIONS:

(a) forearms and thighs

Similarity statistics	Genuine pair		Imposter pair		Virtual twins pair	
	mean μ	*std. dev.* σ	*mean* μ	*std. dev.* σ	*mean* μ	*std. dev.* σ
Forearm overall	13.13	17.70	44.74	22.60	37.89	27.93
Left forearm	15.50	21.05	46.27	24.24		
Right forearm	10.86	13.58	41.54	21.15		
Thigh overall	28.42	27.66	68.80	41.21	56.55	36.54
Left thigh	28.63	25.48	67.09	35.17		
Right thigh	28.21	29.76	70.56	46.54		

(b) the two forearm groups

Similarity statistics	Genuine pair		Imposter pair		Virtual twins pair	
	mean μ	*std. dev.* σ	*mean* μ	*std. dev.* σ	*mean* μ	*std. dev.* σ
Left forearm without wrist strap	12.12	24.76	43.68	25.72	36.39	33.66
Left forearm with wrist strap	18.82	16.05	49.07	25.08	39.86	17.23

CMC curves are used to analyze the experimental results. LF1, RF1, LT1, and RT1 were used as gallery sets and LF2, RF2, LT2, and RT2 were used respectively as the corresponding probe sets for generating the CMC curves of forearms and thighs. For the CMC curves of the virtual twins' forearms, images in {RF1, $LF1^F$}, {RF1, $LF2^F$}, {RF2, $LF1^F$}, and {RF2, $LF2^F$} were matched. The first datasets in the curly brackets were regarded as a gallery set and the second datasets in the curly brackets were regarded as the corresponding probe set. These four sets of matching were summarized to generate the CMC curve of virtual twins' forearms. The same matching approach was used to generate the CMC curve of virtual twins' thighs. The CMC curves of the forearms with and without wrist straps were also computed. Fig. 11 gives all the CMC curves. The rank-1 identification accuracy of the right and left forearms are 80.4% and 66.8%, respectively and the rank-1 identification accuracy of the right and left thighs are 56.9% and 57.9%, respectively. Because vein patterns in forearms are clearer, the matching algorithm performs better for them. Even though the rank-1 identification accuracy of virtual twins from the forearms and the thighs is below 5%, their rank-20% identification accuracy is more than 40%. The term "rank-20% identification accuracy" refers to the percentage of input vein patterns whose corresponding vein patterns can be found in the database within the top 20% of the vein patterns given by the matching algorithm. The CMC curves of the forearms and thighs are much higher than the corresponding CMC curves of virtual twins. However, the CMC curves of virtual twins are significantly higher than the reference lines. Fig. 11(a) also shows that the virtual twins' forearms without the wrist straps perform better than the virtual twins' forearms with the wrist straps. The size of the virtual twins' forearm groups with and without the wrist straps are similar. One is 112 and the other is 122. This performance difference is caused by the different sizes of ROIs. The ROIs of the forearms without the wrist straps are larger and therefore, their vitual twin pairs can find more genetically dependent veins for matching. Fig. 11 indicates clearly that the genetically identical vein patterns have extra similairy. To further justify this point, Fig. 12 shows 8 genetically identical forearm vein patterns and 4 genetically identical thigh vein patterns. Their enhanced vein patterns are also given. Although genetically identical vein patterns are different, some vein segments are similar.

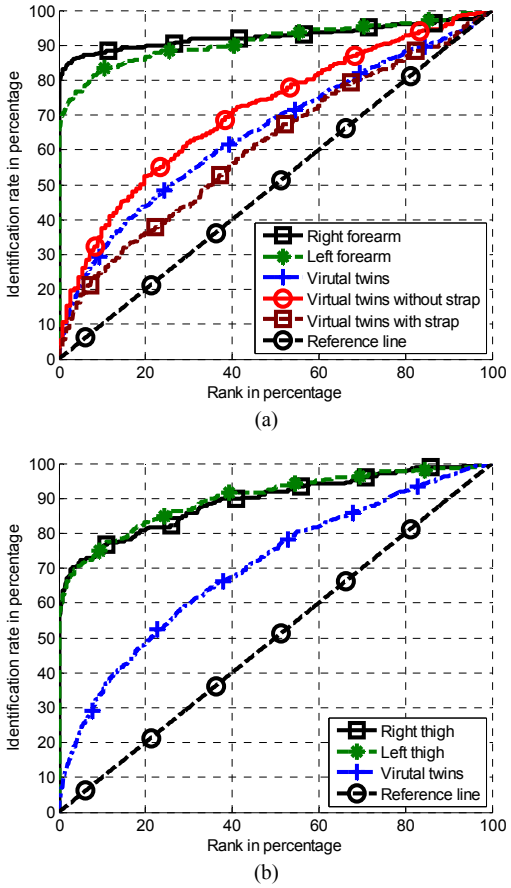


Fig. 11 CMC curves of normal and virtual twins' forearms and thighs. (a) CMC curves of forearms and (b) CMC curves of thighs.

To demonstrate that the similarity between genetically identical vein patterns is not from other factors, including the number of points representing the vein patterns (Fig. 4c), the size of ROIs, the size of vein patterns, and body mass index (BMI), which influences the number of obserable vein segments, further analysis was performed. The size of an ROI is defined by the number of pixels in the ROI and the size of a vein pattern is defined by the number of white pixels in the binarized enhanced vein pattern. Fig. 13 illustrates the size of ROIs and the size of vein patterns. A vein pattern in a gallery set and a vein pattern in a probe set having the closest BMI were considered a target pair to generate the CMC curve of BMI. The same approach was used to provide the CMC curves of the other three factors. Fig. 14 compares the CMC curves of virtual twins and other factors. Clearly, the CMC curves of virtual twins are significantly higher than the CMC curves of other factors, which are close to the reference lines. These

results pinpoint that the extra similarity in genetically identical vein patterns is not from the four factors.

V. Conclusion

In this paper, genetically identical vein patterns are studied. Using an automatic matching algorithm, a series of experiments was performed on 982 NIR forearm images and 906 NIR thigh images, which include 234 genetically identical forearm pairs and 204 genetically identical thigh pairs. The experimental results indicate that genetically identical vein patterns are distinguishable but they, as with fingerprints and palmprints, have some genetic dependence. Experiments were also performed to guarantee that the dependence is not from other factors, including the number of points representing the vein patterns, the size of ROIs, the size of vein patterns, and body mass index.

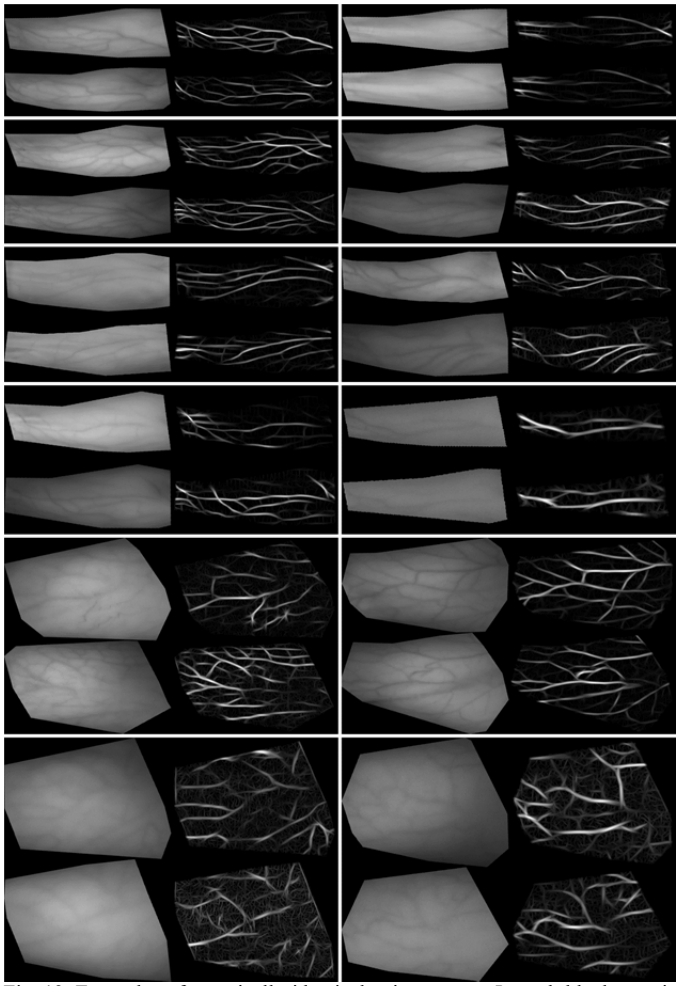
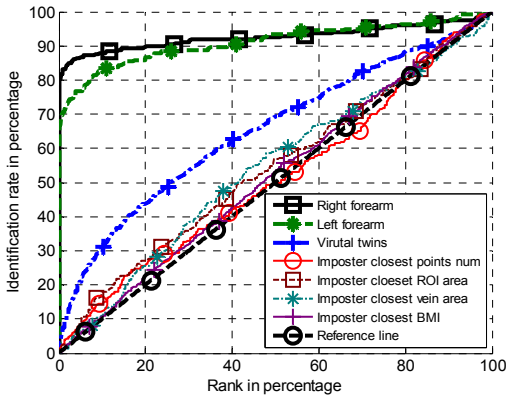


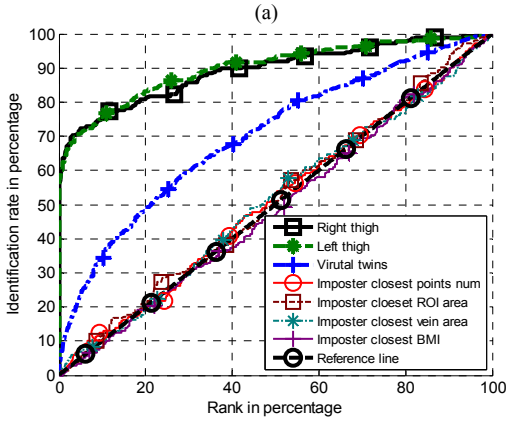
Fig. 12 Examples of genetically identical vein patterns. In each block, a pair of genetically identical vein patterns in NIR images and their enhanced vein patterns are shown.

(a) (b) (c) (d)

Fig. 13 Illistration of the ROI and the vein area size. (a) a segmented NIR image, (b) its ROI, (c) the enhanced vein pattern from (a), and (d) the binarized vein pattern from (c). The number of white pixels in (b) represents the size of ROI and the number of white pixels in (d) represents the size of the vein pattern.



(a)



(b)

Fig. 14 Comparisons of different CMC curves (a) forearms and (b) thighs.

Acknowledgement

The authors would like to thank the Ministry of Education, Singapore for partial support of this research through Academic Research Fund Tier 2, MOE2012-T2-1-024.

References

[1] D. Webbink, J. Roeleveld, and P.M. Visscher, "Identification of twin pairs from large population-based samples," Twin Research and Human Genetics, vol. 9, pp. 496-500, 2006.

[2] Z. Sun, A.A. Paulino, J. Feng, Z. Chai, T. Tan, and A.K. Jain, "A study of multibiometric traits of identical twins," Proceedings of the SPIE, Biometric Technology for Human Identification VII, vol. 7667, pp. 1-12, 2010.

[3] A.K. Jain, S. Prabhakar, and S. Pankanti, "On the similarity of identical twin fingerprints," Pattern Recognition, vol. 35, pp. 2653-2663, 2002.

[4] J. Daugman and C. Downing, "Epigenetic randomness, complexity and singularity of human iris patterns," Proceedings of the Royal Society of London. Series B: Biological Sciences, vol. 268, pp. 1737-1740, 2001.

[5] A.W.K. Kong, D. Zhang, and G. Lu, "A study of identical twins' palmprints for personal verification," Pattern Recognition, vol. 39, pp. 2149-2156, 2006.

[6] C. Tang, A.W.K. Kong, and N. Craft, "Uncovering vein patterns from color skin images for forensic analysis," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 665-672, 2011.

[7] C. Tang, H. Zhang, A.W.K. Kong, and N. Craft, "Visualizing vein patterns from color skin images based on image mapping for forensics analysis," International Conference on Pattern Recognition (ICPR), pp. 2387-2390, 2012.

[8] "Key facts" National Center for Missing & Exploited Children Jun 2013 [Online]. Available: http://www.missingkids.com/KeyFacts

[9] A. Nurhudatiana, A.W.K. Kong, K. Matinpour, S.Y. Cho, and N. Craft, "Fundamental statistics of relatively permanent pigmented or vascular skin marks for criminal and victim identification," International Joint Confernece on Biometrics pp. 1-6, 2011.

[10] A. Nurhudatiana, A.W.K. Kong, K. Matinpour, D. Chon, L. Altieri, S.Y. Cho and N. Craft, "The individuality of relatively permanent pigmented or vascular skin marks (RPPVSM) in independently and uniformly distributed patterns", IEEE Transactions on Information Forensics and Security, vol. 8, no. 6, pp. 998-1012, 2013.

[11] A. K. Jain, R. Bolle, and S. Pankanti, Biometrics: personal identification in networked society: Springer, 1999.

[12] S. Z. Li and A. K. Jain, Encyclopedia of Biometrics: I-Z vol. 2: Springer, 2009.

[13] P. MacGregor and R. Welford, "Veincheck: imaging for security and personnel identification," in Adv. Imaging. vol. 6, ed, pp. 52-56, 1991.

[14] P.L. Hawkes, and D.O. Clayden, "Veincheck research for automatic identification of people," presented at the Hand and Fingerprint Seminar at National Physical Laboratory, September 1993.

[15] L. Wang, G. Leedham, and D.S.Y. Cho, "Minutiae feature analysis for infrared hand vein pattern biometrics," Pattern Recognition, vol. 41, pp. 920-929, 2008.

[16] A. Yuksel, L. Akarun, and B. Sankur, "Biometric identification through hand vein patterns," in 2010 International Workshop on Emerging Techniques and Challenges for Hand-Based Biometrics (ETCHB), pp. 1-6, 2010.

[17] J.G. Wang, W.Y. Yau, A. Suwandy, and E. Sung, "Fusion of palmprint and palm vein images for person recognition based on "laplacianpalm" feature," IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2007.

[18] Y. Zhou and A. Kumar, "Human identification using palm-vein images," IEEE Transactions on Information Forensics and Security, vol. 6, pp. 1259-1274, 2011.

[19] Y. Wang, D. Wang, T. Liu, and X. Li, "Local SIFT analysis for hand vein pattern verification," International Conference on Optical Instrumentation and Technology, 2009, pp. 751204-751204-8.

[20] P.-O. Ladoux, C. Rosenberger, and B. Dorizzi, "Palm vein verification system based on SIFT matching," in Advances in Biometrics, Springer, pp. 1290-1298, 2009.

[21] D. Zhang, Z. Guo, G. Lu, and W. Zuo, "An online system of multispectral palmprint verification," IEEE Transactions on Instrumentation and Measurement, vol. 59, pp. 480-490, 2010.

[22] K.-A. Toh, H.-L. Eng, Y.-S. Choo, Y.-L. Cha, W.-Y. Yau, and K.-S. Low, "Identity verification through palm vein and crease texture," in Advances in Biometrics, Springer, pp. 546-553, 2005.

[23] Y.-B. Zhang, Q. Li, J. You, and P. Bhattacharya, "Palm vein extraction and matching for personal authentication," in Advances in Visual Information Systems, Springer, pp. 154-164, 2007.

[24] 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., vol. 15, pp. 194-203, 2004.

[25] 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, vol. 90, pp. 1185-1194, 2007.

[26] J.-D. Wu and C.-T. Liu, "Finger-vein pattern identification using SVM and neural network technique," Expert Systems with Applications, vol. 38, pp. 14284-14289, 2011.

[27] Z. Zhang, S.L. Ma, and X. Han, "Multiscale feature extraction of finger-vein patterns based on curvelets and local interconnection structure neural network," 18th International Conference on Pattern Recognition, pp. 145-148, 2006.

[28] E. C. Lee and K. R. Park, "Restoration method of skin scattering blurred vein image for finger vein recognition," Electronics Letters, vol. 45, pp. 1074-1076, 2009.

[29] A. Kumar and Y. Zhou "Human identification using finger images," IEEE Transactions on Image Processing, vol. 21, pp. 2228-2244, 2012.

[30] Fujitsu-Laboratories-Ltd. (2003, March). Fujitsu Laboratories Develops Technology for World's First Contactless Palm Vein Pattern Biometric Authentication System. Available: http://pr.fujitsu.com/en/news/2003/03/31.html

[31] H. Zhang, C. Tang, A.W.K. Kong, and N. Craft, "Matching vein patterns from color images for forensic investigation," International Conference on Biometrics: Theory, Applications and Systems, pp. 77-84, 2012.

[32] A. Cohen and T. Vaich, "On the identification of twins by their voices," Proceedings of the ESCA Workshop on Automatic Speaker Recognition, Identification and Verification, pp. 213-216, 1994.

[33] G.R. Doddington, "Speaker recognition—Identifying people by their voices," Proceedings of the IEEE, vol. 73, pp. 1651-1664, 1985.

[34] A.V. Iannarelli, Ear identification, Paramont Publishing Company California, 1989.

[35] A. Ariyaeeinia, C. Morrison, A. Malegaonkar, and S. Black, "A test of the effectiveness of speaker verification for differentiating between identical twins," Science & Justice, vol. 48, pp. 182-186, 2008.

[36] D. Loakes, "A forensic phonetic investigation into the speech patterns of identical and non-identical twins," International Journal of Speech Language and the Law, vol. 15, pp. 97-100, 2008.

[37] A.M. Bronstein, M.M. Bronstein, and R. Kimmel, "Three-dimensional face recognition," International Journal of Computer Vision, vol. 64, pp. 5-30, 2005.

[38] V. Vijayan, K.W. Bowyer, P.J. Flynn, D. Huang, L. Chen, M. Hansen, O. Ocegueda, S.K. Shah and I.A. Kakadiaris, "Twins 3D face recognition challenge," International Joint Conference on Biometrics, pp. 1-7, 2011.

[39] L. Zhang, N. Ye, E. M. Marroquin, D. Guo, and T. Sim, "New hope for recognizing twins by using facial motion," IEEE Workshop on Applications of Computer Vision,pp. 209-214, 2012.

[40] B. Klare, A.A. Paulino, and A.K. Jain, "Analysis of facial features in identical twins," International Joint Conference on Biometrics, pp. 1-8, 2011.

[41] P.J. Phillips, P.J. Flynn, K.W. Bowyer, R.W.V. Bruegge, P.J. Grother, G.W. Quinn, and M. Pruitt, "Distinguishing identical twins by face recognition," IEEE International Conference on Automatic Face & Gesture Recognition, pp. 185-192, 2011.

[42] T. Le, K. Luu, K. Seshadri, and M. Savvides, "A facial aging approach to identification of identical twins," IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 91-98, 2012.

[43] N. Srinivas, G. Aggarwal, P.J. Flynn, and R.W. Vorder Bruegge, "Analysis of facial marks to distinguish between identical twins," IEEE Transactions on Information Forensics and Security, vol. 7, pp. 1536-1550, 2012.

[44] L. Zhang, S. Zhu, T. Sim, W. Leow, H. Najati, and D. Guo, "Hearing versus seeing identical twins," in Computer Analysis of Images and Patterns. vol. 8047, R. Wilson, E. Hancock, A. Bors, and W. Smith, Eds., ed: Springer Berlin Heidelberg, pp. 137-144, 2013.

[45] W.N. Mohd-Isa, J. Abdullah, and C. Eswaran, "Classification of gait biometric on identical twins," Journal of Advanced Computer Science and Technology Research, vol. 2, pp. 166-175, 2012.

[46] A. Myronenko and X. Song, "Point set registration: coherent point drift," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, pp. 2262-2275, 2010.

[47] M. Motivans and T. Kyckelhahn, "Federal prosecution of child sex exploitation offenders, 2006," Bureau of Justice Statistics Bulletin, pp. 1-8, 2007.