Method for estimating potential recognition capacity of texture-based biometrics

Yiding Wang [1], Chenyan Yang [1], Lik-Kwan Shark [2*]

[1] College of Information Engineering, North China University of Technology, Beijing, China
[2] School of Engineering, University of Central Lancashire, Preston, United Kingdom
[*] LShark@uclan.ac.uk

Abstract: When adopting an image-based biometric system, an important factor for consideration is its potential recognition capacity, since it not only defines the potential number of individuals likely to be identifiable, but also serves as a useful figure of merit for performance. Based on block transform coding commonly used for image compression, this paper presents a method to enable coarse estimation of potential recognition capacity for texture-based biometrics. Essentially, each image block is treated as a constituent biometric component, and image texture contained in each block is binary coded to represent the corresponding texture class. The statistical variability among the binary values assigned to corresponding blocks is then exploited for estimation of potential recognition capacity. In particular, methodologies are proposed to determine appropriate image partition based on separation between texture classes and informativeness of an image block based on statistical randomness. By applying the proposed method to a commercial fingerprint system and a bespoke hand vein system, the potential recognition capacity is estimated to around 10^{36} for a fingerprint area of 25 mm^2 which is in good agreement with the estimates reported, and around 10^{15} for a hand vein area of 2268 mm^2 which has not been reported before.

1. Introduction

With an increasing demand of biometric systems for automatic and secure personal identification as well as an increasing number of biometric modalities available [1], an important factor for consideration when adopting a particular biometric system is its potential recognition capacity, defined as the potential number of individuals that is likely to be identifiable. For personal identification based on practical biometric images, potential recognition capacity depends not only on distinctiveness related to modality, but also on other factors, such as image coverage and quality related to imaging system and acquisition control [2], as well as feature extraction, strength and matching related to image processing methodologies. Hence, if potential recognition capacity can be estimated, it can also serve as a useful figure of merit for performance comparison.

In stark contrast to the vast literature available on a plethora of approaches to personal identification using different behavioural and biological characteristics [3], only a limited number of papers exists on estimation of biometric recognition capacity, with most of the work conducted from an information theoretic view point. The methodologies reported in the early papers tend to be specific to a particular type of biometric modality from the perspective of distinctiveness, namely, the probability of random correspondence (PRC) between two sets of biometric features extracted from two arbitrary individuals in a population to yield a false match. The earliest paper can be traced back to 1892 on fingerprint distinctiveness, with the first statistical model produced based on subjective statistical analysis of pattern frequencies and minutiae occurrences [4]. Since then, various statistical models with increasing complexity have been reported [5-7], and these models can be categorised into (a) block based approaches by dividing the fingerprint into square regions and capturing the statistical variability of local patterns within each region [4, 8, 9]; (b) event based approaches by deriving a fixed probability for the occurrence of each minutiae type [10, 11]; (c) relative measurement based on the positions and orientations of minutiae with respect to landmark based spatial references [12-14]; and (d) generative approaches such as using point processes, mixture models, and Bayesian networks to model minutiae spatial distribution [15-18], minutiae position-orientation dependencies [19-21] and inter-minutiae relationship [22]. There is a significant weakness in the early models as they were established using an ideal condition that minutiae are independent and identically distributed random events. Although the latter models have considered the minutiae variability in a more realistic manner, more complex statistical relationships were ignored in order to simplify model formulation. Furthermore, only minutiae were considered in these models without including other discriminatory features in fingerprints, such as pattern types, apart from [23]. There are two other important aspects which have not been explicitly taken into account, namely, image quality and intraclass variations which include non-linear deformation caused by factors such as skin elasticity and non-uniform fingertip pressure.

For other biometric modalities, face distinctiveness and iris distinctiveness have also been reported. While the former was studied based on unusual facial features with its development still in infancy [24], the latter was approached through large scale empirical evaluation at the matching score level instead of feature level due to high complexity of iris patterns for direct statistical modelling [25-27].

By treating image quality and intraclass variation as possible sources of noise, biometric recognition has been modelled as information transmission through a noisy channel [28], and this enables the application of the information theory to yield the constrained capacity, defined as the number of individuals that can be reliably identified with a low error rate. Such model has been reported for face and iris images based on global appearance features obtained

from principal component analysis and independent component analysis [29], and for palmprint based on local frequency features obtained from the discrete cosine transform [30]. For the former, it was formulated based on the mutual information of biometric features with the assumption of their distributions being independent and identical. For the latter, it was formulated based on the information capacity derived from the genuine and imposter matching score distributions which were assumed to be Gaussian. Comparing the two approaches, the feature-level based approach is seen to be more relevant for characterisation of biometric distinctiveness with quantifiable confidence level, whereas the score-level based approach is seen to be more appropriate for evaluation of system performance with a computation advantage. Nevertheless, the assumptions made by both are generally not applicable in practice. An alternative but related approach is based on biometric information loss via the relative entropy derived from the interclass and intraclass feature distributions [31]. Although it has been applied to face images, the assumption of Gaussian distributions was again required in order to tackle the inherent computation issues related to high-dimensional feature space and limited availability of data.

Different from the information theoretic approaches, a data-driven methodology is proposed in this paper to provide a practically applicable approach for estimation of potential biometric recognition capacity, whereby realistic biometric feature characteristics are learnt from the actual biometric images collected from the real biometric system under consideration, without the need to make unsubstantiated assumptions of biometric feature distributions. In particular, by drawing some parallels between classification and compression of images, the proposed method borrows the concept from block transform coding that is commonly used for image compression. Essentially, an image is partitioned into image blocks with each block considered as a constituent biometric component, and the image texture contained in each block is binary coded to represent the corresponding texture class. The statistical variability among the binary codes of different images is then used to identify the informativeness of each image block, thereby enabling coarse estimation of potential recognition capacity. Although the proposed methodology shares some similarities with previous works on block-based approaches for statistical modelling of fingerprints [4, 8, 9] as well as block-based coding of iris images [25-27] and palmprint images [30], a significant contribution of the paper lies in providing a solution to the tricky problem of appropriate block size.

To illustrate the proposed methodology and its applicability, near-infrared (NIR) dorsal hand vein images and fingerprint images acquired from two practical biometric systems in an unsupervised environment were used. With no prior work on the recognition capacity of hand vein images which are considered to have poor image quality, a practical contribution is seen to be there for this work to be the first to provide a reference point. Unlike various prior works on fingerprint distinctiveness based on minutiae only, the block-based texture coding approach proposed in this paper directly takes in raw fingerprint images for estimation of recognition capacity, whereby the whole fingerprint image is encoded as a binary string with its local and global patterns represented by binary value and bit position, respectively. The texture-based approach with the advantage of including all possible fingerprint features for recognition capacity estimation is seen to form another contribution of the paper, and a comparison of the results with other approaches is presented to confirm the validity of the proposed methodology.

In the rest of the paper, presented first in Section 2 is the proposed methodology, which is then applied to fingerprint and NIR dorsal hand vein images in Section 3 to estimate their potential recognition capacities. Finally, conclusions are given in Section 4.

2. Methodology

Although compression and classification of images are fundamentally different in their goals, several parallels can be drawn from their implementation frameworks, since both of them map an input image to a binary string, which is seen as representing a compact codeword for transmission/storage in the case of image compression, or a class label in the case of image classification. These similarities have led to the use of block-based texture coding as the first stage in estimation of potential recognition capacity by adapting the three standard steps in block transform coding, which are image transform, coefficient quantisation, and binary encoding.

In image transform as the first step of coding, the input image is partitioned into small square blocks of equal size so as to enable the selected transform operation to be performed on a block basis, and this step can be expressed mathematically as $f(k)'=\boldsymbol{T}b(k)'$, where $k\in[1, \ldots, K]$ with K denoting the number of blocks; $b(k)$ denotes the row vector formed by concatenating $N\times N$ pixels of the kth block with $b(k) \in \mathbb{R}^{N\times N}$; and $\boldsymbol{T}$ denotes the transform operator to give a new representation $f(k) \in \mathbb{R}^{D}$ with D denoting the number of dimensions in the transform domain. Unlike image compression that normally has $D=N\times N$ because it seeks a reversible transform to map most of the energy in the input image to a few large transform coefficients, biometric identification allows $D<<N\times N$ because it seeks a mapping function (that does not have to be reversible) to enable classification to be performed in a low-dimensional feature space.

The adaption of this image transform step for estimation of potential recognition capacity requires consideration of two intertwined parameters, namely, image block size and the number of texture classes to be used. With each image block treated as a constituent biometric component in this work, how to determine an appropriate block size in order to capture discriminative biometric content presents a tricky problem, because it depends not only on image resolution but also on the effective geometric coverage of the texture descriptor adopted. From an information perspective, an image block with $N\times N$ pixels and G possible grey-level values per pixel can appear in any one of $G^{N\times N}$ possible states, and the maximum amount of information conveyed by each image block is $N\times N\times\log_2 G$ bits, if all possible states are equally likely to occur. For a given G, as N increases, more image information are contained within each image block, and more texture classes are likely to be needed in order to provide a good representation without loss of salient image information. However, in the case of many biometrics based on imaging of geometric appearance, which include fingerprint and hand vein, the actual amount of information conveyed by each

image block is significantly less than $N\times N\times\log_2 G$ bits, due to high occurrence probabilities of certain grey-level ranges and spatial dependency among the neighbouring pixels. This gives the possibility of representing the salient biometric content of an image block by using a small number of texture classes. For the sake of implementation simplicity, the number of possible texture classes is set to the smallest value of two, thereby enabling the use of a single between-class distance value instead of multiple between-class distances to determine an appropriate image block size, as presented in the next section.

There are various texture descriptors available to provide a compact representation of the image content in each block. A well-known statistical mapping function to characterise image texture is the grey-level co-occurrence matrix (GLCM) [32] that describes the spatial relationship of grey-level values based on their co-occurrence frequencies, and it is expressed for the kth image block as:

$$P_k(i,j)=\frac{1}{N^2}\sum_{x=1}^{N}\sum_{y=1}^{N}\begin{cases}1 & \text{if } I_k(x,y)=i \text{ and} \\ & \quad I_k(x+d\cos\theta, y+d\sin\theta)=j \\ 0 & \text{otherwise}\end{cases} \tag{1}$$

where $I_k(x, y)$ and $I_k(x + d\cos\theta,\ y + d\sin\theta)$ are pixel pairs separated by distance d at angle θ with grey-level values of i and j. For an image represented by L grey-level ranges, the size of GLCM produced by (1) is $L\times L$. By applying different statistical measures to GLCM, a set of texture descriptors is available to represent the grey-level variation in each image block [32]. By employing one of the statistical texture descriptors, the whole image is mapped to a K-element feature vector denoted by $f(k)$.

In the second step of coefficient quantisation, the feature values contained in $f(k)$ is restricted from a continuous set of possible values to a discrete set of allowed levels. Unlike image compression that seeks a quantisation function to spread transform coefficients uniformly among as many allowed quantisation levels as possible to minimise approximation errors, the objective of this step is to divide the feature values into as few classes as possible with minimum classification errors. This leads to 2-level quantisation to divide texture feature values into two classes, which may be considered as indicating absence or presence of a particular biometric texture pattern in an image block. With T_q denoting the threshold to separate two different texture classes, 2-level quantisation is expressed as:

$$q(k)=\begin{cases}0 & \text{if } f(k)<T_q \\ 1 & \text{otherwise}\end{cases} \tag{2}$$

Since a sufficiently discriminative texture descriptor will result in the feature values of two texture classes following a bi-modal statistical distribution, the optimum value for T_q can be found by using the Otsu method [33] to maximise the separation distance between the two texture classes:

$$T_q=\underset{T}{\operatorname{argmax}}\{w_{<T}(\mu_{<T}-\mu_f)^2+w_{\geq T}(\mu_{\geq T}-\mu_f)^2 \tag{3}$$

where $w_{<T}$ and $w_{\geq T}$ denote the probabilities of the feature values contained in $f(k)$ below and above the threshold value of T with the respective mean values of $\mu_{<T}$ and $\mu_{\geq T}$, and μ_f denotes the mean value of $f(k)$.

In the third step of binary encoding, the multiple texture classes in the quantised feature vector are coded using binary bits and concatenated into a binary string to represent a biometric image. For 2-level quantisation with only 1 bit needed for binary representation of each image block, the quantised feature vector is already in the required binary form.

Given a set of biometric images acquired from different individuals and mapped to their corresponding binary codes based on texture features, the second stage of the proposed method is to estimate the potential recognition capacity based on the informativeness of each image block from the statistical variability of its binary values. For 2-level quantisation and coding of the texture information in each image based on K blocks, it results in a K-bit binary code per image, and the number of possible binary combinations is 2^K. However, not all the image blocks are likely to provide useful discriminatory information. With binary 0 and 1 representing the absence or present of a biometric texture pattern in an image block, the informativeness of an image block is reflected by the statistical variability of its binary values across images of different individuals. The most informative image blocks correspond to those having the equal occurrence probability of 0.5 for both binary values, as a result of statistical independence among different individuals; and the non-informative image blocks correspond to those having a high or low occurrence probability for one of the binary values, as a result of similar image appearance among different individuals. Let p_k denote the occurrence probability of binary 1 for the kth image block, and let R denote an informativeness threshold to exclude those image blocks with the occurrence probability of binary 1 significantly above and below the equal probability of 0.5, the number of informative image blocks based on statistical entropy, denoted by K', is given by

$$K'=-\sum_{k=1}^{K} p_k\, log_2\, p_k+(1-p_k)log_2(1-p_k)|_{p_k\in[0.5-R,0.5+R]} \tag{4}$$

With $K'<K$, the number of possible binary combinations is reduced to $2^{k'}$.

Finally, using the Hamming distance measure to give the smallest number of bits in a binary code that must be changed in order to differentiate one individual from the other, the potential recognition capacity is estimated as:

$$C=\frac{2^{K'}}{\bar{H}} \tag{5}$$

where $\bar{H}$ denotes the average interclass Hamming distance.

3. Potential recognition capacities of hand vein and fingerprint

In this section, the proposed methodology is applied to estimate the potential recognition capacities of two different biometric modalities. One is NIR dorsal hand vein images with no previous report found on its recognition capacity, and the other is fingerprint images to confirm the compatibility of the proposed method with respect to other estimates reported.

3.1. Hand vein and fingerprint images

All the results reported in this section are based on images from the NCUT biometric database that is publicly available. While hand vein images were acquired using a self-developed NIR image acquisition device [34], fingerprint images were acquired using a DigitalPersona UareU 4000 fingerprint reader [35]. Image acquisition was conducted in a realistic environment without supervision and with participants recruited from the student population at NCUT. To include intraclass variations, each volunteer was required to alternate their hands and fingers for each image acquisition. 80 hand classes and 80 fingerprint classes were used in this work with 10 images per class.

Since the dataset is relatively small in size and contains significant intraclass variations, geometric pre-processing is required to produce correct alignment of different images acquired from the same individual and consistent extraction of the region-of-interest (ROI) from each image before applying the proposed methodology. While alignment correction is used to minimise intraclass variations caused by posture variations in an unconstrained image acquisition environment, consistent ROI extraction ensures estimation of the recognition capacity based on the texture information contained in the same image blocks. While geometric pre-processing was achieved using a fully automatic approach based on the method developed previously for NIR dorsal hand vein images [36, 37], a semi-automatic approach with manual involvement was used for fingerprint images to achieve the accuracy required.

The geometric pre-processing of hand vein images consisted of (a) background removal to extract the hand area; (b) horizontal shear correction based on relative position of the mid-point along the top and bottom parts of the hand to align hand posture; (c) computation of hand centroid to provide a geometric reference; and (d) image cropping from the hand centroid to provide the ROI with a fixed size. These steps are illustrated in Fig. 1, where an example hand vein image with horizontal tilt is shown on the left, and the hand area extracted by thresholding in step (a) is shown on the right. Superimposed on the extracted hand area are mid-points *E* and *F* found by using horizontal lines *AB* and *CD* along the top and bottom parts of the hand, and this enables the shearing factor to be estimated in step (b) as the slope of line *EF*. The hand centroid computed in step (c) based on the pixel values in the extracted hand area and the ROI produced by image cropping in step (d) are shown as the white dot and white square on Image A of Hand 1 in Fig. 3, where the effect of hand tilting is seen to be reduced with vein lines running more parallel to the vertical axis.

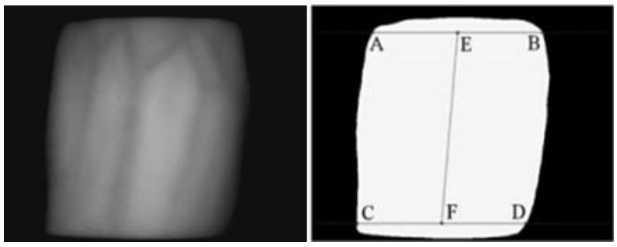


Fig. 1. *Example hand vein image* (*left*), *and extracted hand area with geometry for shearing estimation* (*right*)

The geometric pre-processing of fingerprint images consisted of (a) identification of two corresponding landmark pairs to provide the geometric reference required; (b) position alignment based on the coordinate differences of the landmark pair near to the image centre via image translation; (c) orientation alignment based on the angular differences derived from the two landmark pairs via image rotation; and (d) image cropping from the landmark pair near the image centre to provide the ROI with a fixed size. These steps are illustrated in Fig. 2 using two example fingerprint images from the same individual to be aligned. Superimposed on the two images are two corresponding landmark pairs identified in step (a), and they are denoted by *AA*', and *BB*', respectively. While the coordinate differences between *A* and *A*' are used for position alignment in step (b), the angular difference between lines *AB* and *A*'*B*' with respect to the horizontal is used for orientation alignment in step (c). The corresponding ROI pair produced in step (d) by image cropping from *A* and *A*' are shown as white squares on Images A and B of Finger 1 in Fig. 4, where they are seen to contain similar geometric patterns.



Fig. 2. *Example fingerprint images with corresponding landmarks for their alignment*

To illustrate interclass and intraclass image differences, four hand vein images from two different hands and four fingerprint images from two different fingers are shown in Figs. 3 and 4, with ROI centre and size superimposed on each image. Compared with distinctive patterns contained in interclass ROIs (in images of different hands or fingers), intraclass ROIs (in Images A and B of the same hand or finger acquired at two separate instants) appear to contain the same and geometrically aligned pattern at a cursory glance. On a close examination of these intra-class images, they are not identical with local image differences which are more apparent in hand vein images. With geometric pre-processing based on affine transformation, the intraclass variations in ROI consist of small linear geometric deformation as a result of inherent alignment errors and uncorrected non-linear geometric deformation. Compared with fingerprint images, hand vein images are seen to have higher intraclass variations and this can be attributed to a

larger effect of non-linear geometric deformation, due to a large coverage area with high curvature at the back of the hand as well as skin elasticity and imaging camera. Furthermore, hand vein images have lower image quality, as reflected by lower contrast between vein and its surrounding tissue, in comparison to high contrast between ridge and valley in fingerprint images.

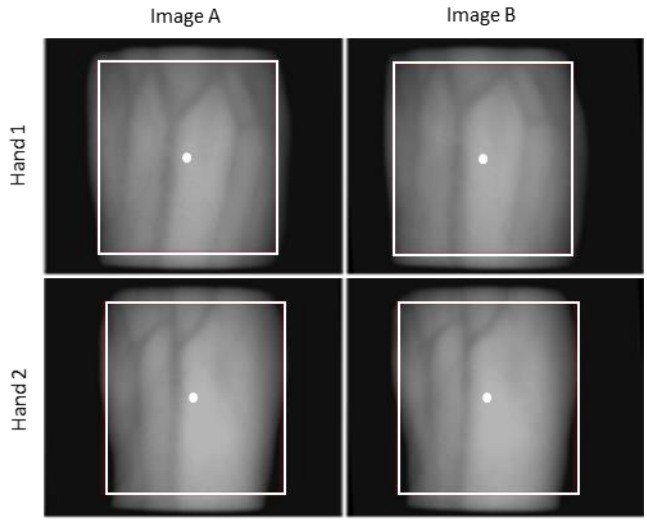


Fig. 3. *Examples of four hand vein images with two images per class*



Fig. 4. *Examples of four fingerprint images with two images per class*

From the overall image characteristics observed between Figs. 3 and 4, one is likely to expect the potential recognition capacity of hand vein to be lower than that of fingerprint. This is because, in comparison to numerous fingerprint papillary ridges densely distributed in a small fingertip area, only a few vein lines sparsely distributed in a relatively large area of dorsal hand are visible, implying a lower information density of hand vein images.

Subsequent to extraction of ROI are median filtering based on a small window of 3×3 pixels to reduce outlier noise in ROI and normalisation of the grey-level range. Outlier noise, if not reduced, can lead to a higher estimate of the potential recognition capacity, since the informativeness of an image block depends on the statistical variability of its binary values among different image classes and the random nature of outlier noise has the potential to increase the statistical variability. For capacity estimation described in the following, the ROI sizes were set to 381×381 pixels for hand vein images to cover most of the dorsal hand part with an area around 2268 mm^2, and 100×100 pixels for fingerprint images to cover a patch around 25 mm^2, respectively.

3.2. Block based feature coding

Following the methodology presented in Section 2, the first stage of recognition capacity estimation is block based feature coding after image pre-processing described in Section 3.1, and the first step is to map the texture content in each image block to a corresponding feature value based on a texture descriptor. In the implementation of this step, four GLCMs were computed for each image block in each image using (1) with d=1 at $\theta \in [0°, 45°, 90°, 135°]$ to capture grey-level variations in different directions, and averaged. In order to reduce the matrix size of GLCM and computation, the number of grey-level values was reduced according to image contrast, with 16 grey-level intervals used to represent hand vein images and 2 grey-level intervals to represent fingerprint images. The texture descriptor based on entropy was then used to provide the statistical measure of the averaged GLCM produced for each image block:

$$f(k) = -\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} \overline{P_k}(i,j)\, log_2\, \overline{P_k}\,(i,j) \qquad (6)$$

where $\overline{P_k}(i,j)$ denotes the averaged GLCM. Since $f(k)$ is minimum when all values in $\overline{P_k}(i,j)$ are equal, and high otherwise, a higher texture entropy value is likely to be generated by a relatively inhomogeneous image block containing geometric edges (such as vein-tissue and ridge-valley borders) and vice versa. This is illustrated in Fig. 5, showing the corresponding texture entropy images of Hand 1 Image A ROI shown in Fig. 3 based on 20×20 image blocks and Finger 1 Image A ROI shown in Fig. 4 based on 31×31 image blocks.

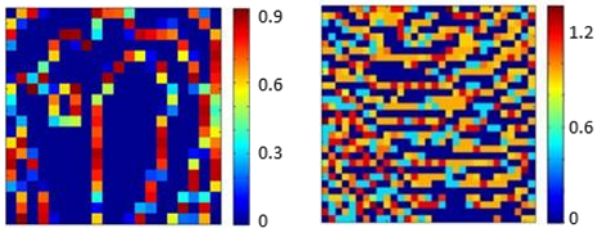


Fig. 5. *Texture entropy images of hand vein* (*left*)*, and fingerprint* (*right*)

For the steps of coefficient quantisation and binary encoding, (3) was applied to the set of texture entropy values generated from each image to produce a threshold, and the threshold was then used in (2) to generate a binary code with each image block represented by either binary 0 or 1, which may be considered as indicating the absence or presence of vein-tissue or ridge-valley borders in an image block. Using the hand vein and fingerprint texture entropy images shown

in Fig. 5 as examples, the resulting binary images produced by this step are illustrated in Fig. 6.

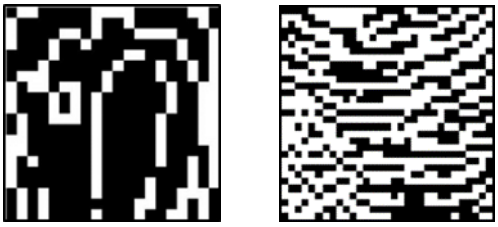


Fig. 6. *Binary coded images of hand vein* (*left*)*, and fingerprint* (*right*)

The quantisation step can be considered as dividing the texture features contained in each image block into two classes corresponding to absence and presence of vein-tissue or ridge-valley borders. With T_q being the threshold produced by (3) to maximise the distance between two possible texture classes, the variance of the texture entropy values obtained from each image, $f(k)$, with respect to T_q should be maximum if the number of image blocks, K, is optimum, and this condition can be expressed as:

$$K_{opt} = \underset{k}{\operatorname{argmax}} \sum_{k=1}^{K} (f(k) - T_q)^2 \qquad (7)$$

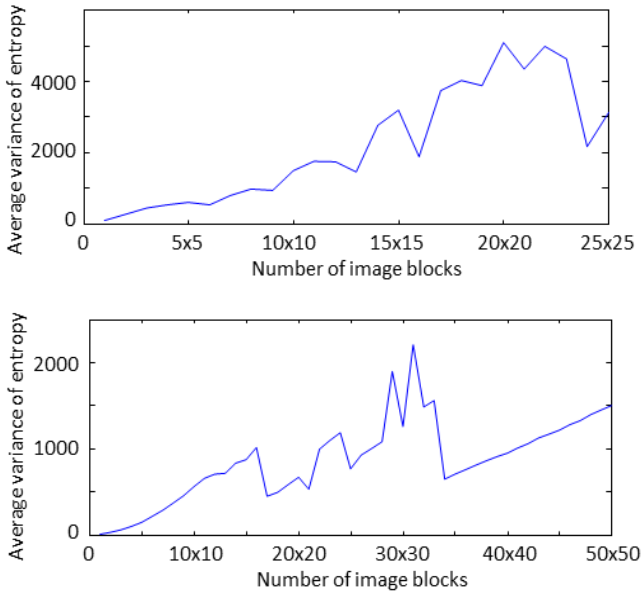


Fig. 7. *Average variance of texture entropy versus number of image blocks for hand vein* (*top*) *and fingerprint* (*bottom*)

Based on the images of 10 individual hands and 10 individual fingerprints, the average variance of texture entropy against the number of image blocks for hand vein and fingerprint are shown in Fig. 7. From the highest average variance in Fig. 7, the optimum numbers of image blocks are seen to be 20×20 blocks for hand vein, and 31×31 blocks for fingerprint. Converting these numbers to image block sizes gives 19×19 pixels per block for hand vein, and 3×3 pixels per block for fingerprint. Hence, for the given image resolution and texture feature descriptor, the biometric information density of hand vein based on two texture classes per block is around 40 times lower, compared with fingerprint.

From the classification perspective, K_{opt} should also produce a good separation between the interclass binary codes representing different classes and intraclass binary codes representing the same class. Hence, Hamming distance can be used to show the goodness of K_{opt} by measuring the dissimilarity among interclass and intraclass binary codes. For two binary codes denoted by $q_a(k)$, and $q_b(k)$, the normalised Hamming distance is given by:

$$HD = \frac{1}{K} \sum_{k=1}^{K} q_a(k) \otimes q_b(k) \qquad (8)$$

where $\otimes$ denotes the exclusive-OR operation. By using 200 interclass image pairs and 100 intraclass image pairs, the computed *HD* distributions of hand vein and fingerprint are shown in Fig. 8 for two different selections in the number of image blocks.

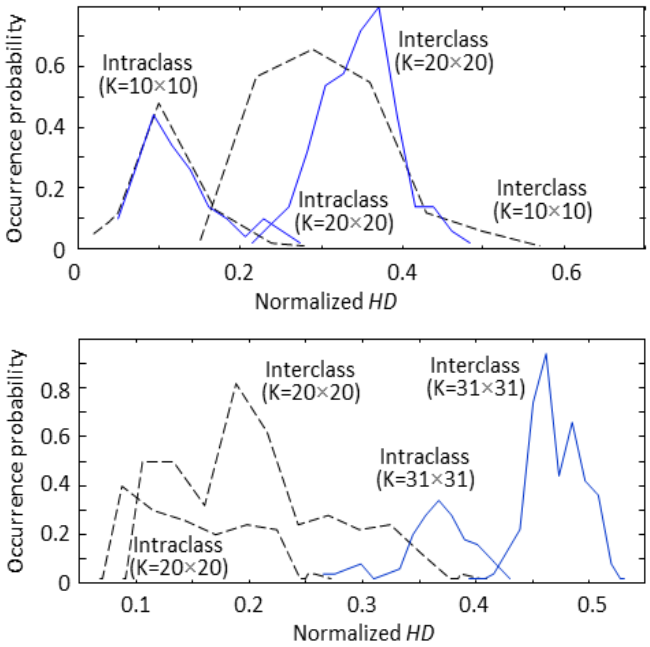


Fig. 8. *Interclass and intraclass HD distributions for hand vein* (*top*) *and fingerprint* (*bottom*)

For hand vein images, the interclass and intraclass *HD* distributions were computed based on the image partitions of K_{opt}=20×20 blocks with 19×19 pixels per block and K=10×10 blocks with 38×38 pixels per block. It is seen from the top figure in Fig. 8 that the interclass *HD* distance values for the optimum image partition with K_{opt}=20×20 blocks are higher, as shown by its average interclass distance of 0.35 (corresponding to $\bar{H}$ =139 bits), compared with 0.31 (corresponding to $\bar{H}$=31 bits) for K=10×10 blocks, indicating a much lower correlation among the binary codes of hand veins in different hand classes. The overlap between the interclass and intraclass *HD* distributions is also less for the optimum image partition with K_{opt}=20×20 blocks, as shown by a narrower interclass *HD* distribution with the variance value of 0.048, compared with 0.07 for K=10×10 blocks, indicating a much lower error in distinguishing between

binary codes of hand veins in different hand classes and binary codes of hand veins in the same hand class.

Similarly, for fingerprint images, the interclass and intraclass *HD* distributions were computed based on the image partitions of K_{opt}=31×31 blocks with 3×3 pixels per block and K=20×20 blocks with 5×5 pixels per block. As shown in the bottom figure of Fig. 8, the same two observations apply, namely, higher interclass *HD* distance values, and less overlap between the interclass and intraclass *HD* distributions for the optimum image partition with K_{opt}=31×31 blocks. While the average interclass distance for K_{opt}=31×31 blocks is 0.472 (corresponding to $\overline{H}$=453 bits), significantly higher than 0.202 (corresponding to $\overline{H}$=81 bits) for K=20×20 blocks, the interclass variance for K_{opt}=31×31 blocks is 0.024, significantly lower than 0.064 for K=20×20 blocks.

With each ROI image divided into K_{opt} blocks and each block coded as 0 or 1, the binary coded images obtained for Hand 1 Image A and Finger 1 Image A shown in Figs. 3 and 4 are illustrated in Fig. 6.

3.3. Estimation of potential recognition capacity

After block-based feature coding of a set of biometric images to their corresponding binary codes, the second stage of the proposed method is to estimate the potential recognition capacity based on the statistical variability of the resulting binary codes. With K_{opt}=20×20 and K_{opt}=31×31 as the optimum numbers of blocks for hand vein and fingerprint images, respectively, Fig. 9 illustrates the statistical variability of each image block based on the occurrence probability of binary 1 derived from 80 different classes.

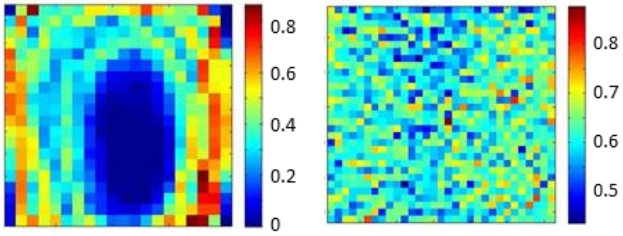


Fig. 9. *Occurrence probability of binary 1 in each image block for hand vein* (*left*)*, and fingerprint* (*right*)

For an image block to be informative, the binary values assigned to the same block based on different individuals should be uncorrelated, giving a high statistical variability with an equal occurrence probability of 0.5 for both binary values of 0 and 1. Based on the statistical variability of blocks shown in Fig. 9 for hand vein images, the informativeness for image blocks in the middle area of the image as well as at four corners are low due to its near to zero probability of being coded as binary one. While the former could be attributed to a high probability of encountering a skin tissue area without vein in the same image block among different individuals, the latter could be attributed to a high probability of encountering background pixels outside dorsal hand area (see hand vein ROI images shown in Fig. 3 with a small number of dark background pixels appeared at corners). Unlike hand vein images, the statistical variabilities of most image blocks are high throughout the fingerprint images as shown in Fig. 9, and this is seen as a result of its dense and high contrast ridge-valley patterns.

In order to estimate the potential recognition capacity, an informativeness threshold is needed in (4) to exclude those image blocks with low statistical variability. However, there exists a dilemma in setting an appropriate informativeness threshold. While too high an informativeness threshold value is likely to result in an underestimate of recognition capacity due to an increase in the number of informative image blocks being excluded, too low an informativeness threshold value is likely to result in an overestimate of recognition capacity due to an increase in the number of non-informative image blocks being included. In this work, by treating binary representation of each biometric unit as a Bernoulli trial with two possible outcomes [38], this dilemma is solved by defining the informativeness threshold as the standard deviation of the binary occurrence probability from the ideal value of 0.5 in order for the statistical variability of an image block to be accepted as random or informative.

Let M represent the number of different image classes available for learning the statistical variability. With these images acquired from different hand or fingerprint classes, the binary values assigned to the same image block based on different images can be treated as independent, meaning that the occurrence of a binary value from one image has no influence on the occurrence of other binary values from other images. If the binary value resulted from each image is assumed to have an identical and random Bernoulli distribution with the same variance, the composite standard deviation of M binary values from the expected mean based on the variance sum law is given by

$$\sigma = \sqrt{\frac{p_1(1-p_1)}{M}} \qquad (9)$$

where p_1 denotes the occurrence probability of binary 1 of an image block, and equals 0.5 for an informative image block. By the law of large numbers, increasing M reduces σ.

In experimental validation, four different sets of hand vein images acquired at different times from the same 80 individuals were used to test the proposed method to see if similar recognition capacity values could be produced. Based on K_{opt} found for hand vein images, each ROI image was divided into a total of 20×20 blocks per image. With the statistical variability of an informative image block allowed to lie within one standard deviation, substituting M=80 and p_1=0.5 in (9) gives σ=0.056, and using it as the informativeness threshold in (4) results in four consistent set of results shown in Table 1, where K' denotes the number of informative image blocks obtained using (4), and $\overline{H}$ denotes the average of the Hamming distance values obtained from 3160 possible image pairs in each dataset. Using the minimum value of K' in Table 1, the percentage of the informative image blocks with respect to K_{opt} is 14.25%, and the potential recognition capacity for hand vein images is estimated to be around 10^{15} for an area of 2268 mm^2 by applying (5).

Table 1 Number of informative image blocks and average *HD* for hand vein

Dataset	1	2	3	4
K'	58	59	58	57
$\bar{H}$	143	143	143	142

Based on the same setting except K_{opt} that is set to 31×31 blocks per image, Table 2 shows the results obtained by applying the proposed method to four different sets of fingerprint images acquired at different times from the same 80 individuals. Although the number of informative image blocks is seen to be variable against consistent values for average Hamming distance from Table 2, it is not seen to invalidate the proposed method, as it is a consequence of using a small image block size of 3×3 pixels, which increases the susceptibility of the binary representation to small image difference caused by geometric distortion of the same fingerprint, ROI extraction error, even random noise remained after filtering. Using minimum value of K' in Table 2, the percentage of the informative image blocks with respect to K_{opt} is around 13.42%, similar to that for hand vein images, and the potential recognition capacity for fingerprint is estimated to be around 1.51×10^{36} for an area of 25 mm^2 by applying (5).

Table 2 Number of informative image blocks and average *HD* for fingerprint

Dataset	1	2	3	4
K'	134	154	129	147
$\bar{H}$	455	453	451	452

There exists a relationship between the potential recognition capacity and the probability of random correspondence (PRC). While the probability of an individual fingerprint can be equated to the inverse of the potential recognition capacity, random correspondence for a pair of fingerprints can be considered as the probability of drawing the same fingerprint twice with replacement. If all possible fingerprints can be assumed to be independent and identically distributed, the equivalent PRC value for the estimated fingerprint recognition capacity can be approximated as the product of two individual fingerprint probabilities, or $1/(1.51\times10^{36})^2 = 4.39\times10^{-73}$, thereby enabling a direct comparison with the estimates of the previous work to show its compatibility. The available PRC values for comparison can be divided into two groups, with one from non-generative modelling approaches and the other from generative modelling approaches. For the non-generative modelling approaches, a wide range of the PRC values has been produced. For fingerprints which are divided into 24 regions and have 36 minutiae on average, the PRC values range from 1.45×10^{-11} based on the early model in 1890s [4, 7] to 1.2×10^{-80} based on the latter model in 1980s [7, 13]. From which, the equivalent PRC value of the proposed method is seen to be more comparable with that from the latter non-generative model. For the generative modelling approaches based on the positions and orientations of minutiae, the PRC values between two fingerprints with each containing 46 minutiae were reported to lie from 5.86×10^{-7} for matching 12 out of 46 minutiae to 1.33×10^{-77} for matching all 46 minutiae [13]. From which, the equivalent PRC value of the proposed method is seen to be more comparable with the lower end of the PRC value involving a large number of features, and it is consistent with the use of a large the number of image blocks in the proposed method for estimation of recognition capacity. Another factor affecting the PRC values is the number of feature types used, and the PRC values for fingerprints have been shown to be significantly smaller in the generative modelling approach including not only minutiae but also ridges and pores [23]. This should be applicable to the proposed method since its equivalent PRC value is likely to be reduced by increasing the number of texture classes in each image block.

In further experimental validation, potential recognition capacity is estimated for degraded biometric images, thereby showing its usefulness as a figure of merit for performance comparison. Simple moving average filters with different window sizes were applied to each hand vein and fingerprint image after pre-processing before using them for recognition capacity estimation. As the window size increases, the image quality becomes poorer with increasingly blurred vein-tissue and ridge-valley borders, and the potential recognition capacity should decrease as a result of more image blocks becoming less informative.

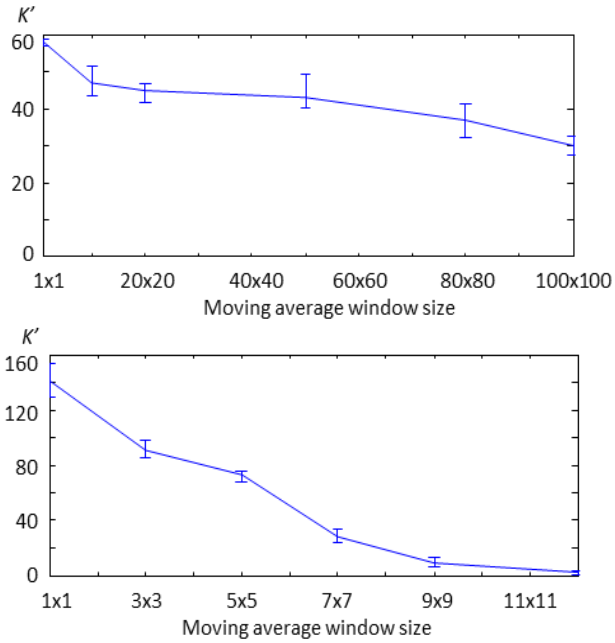


Fig. 10. *Number of informative image blocks versus average window size for hand vein* (*top*) *and fingerprint* (*bottom*)

Using the same four sets of hand vein and finger print images, moving average filters were applied to each set consisting of 80 individuals, and Fig. 10 shows the decline in the number of informative images blocks for hand vein and fingerprint images against moving average window size, where the line represents the average, and the whiskers represent the maximum and minimum values. Due to lower biometric information density as a result of a larger image block size, the decline in the potential recognition capacity for hand vein is seen to be at a slower rate compared with that for fingerprint. In particular, when the image blurring area was set to equal the block size of K_{opt}, the percentage of the informative image blocks with respect to K_{opt} was found to be

reduced by around 3.75% (from 14.25% to 10.5%) for degraded hand vein images, smaller than 4.58% (from 13.43% to 8.95%) for degraded fingerprint images. These correspond to a reduction of the potential recognition capacity to around 10^{10} for hand vein images, and around 10^{23} for fingerprint images. Although these reductions in recognition capacity were produced as a result of uniform blurring of hand vein and fingerprint images by applying moving average filters, it can be related to the practical scenario of image capture with a defocusing camera.

4. Conclusions

By treating each biometric image as consisting of a number of constituent biometric components based on the concept of block-based feature coding, this paper presents a method to provide a coarse estimate of the potential recognition capacity for a practical biometric system, with solutions to the problems of image partition based on separation distance between texture classes and informativeness of image blocks based on the statistical variability of their binary representation. Although the paper focuses on a simple implementation of the proposed methodology based on a common image texture descriptor and the smallest number of texture classes per image block, the approach is applicable for estimation of potential recognition capacity based on other types of feature descriptors and can be extended to multiple texture classes per image block.

Experimental validation involves the use of NIR dorsal hand vein images acquired from an in-house bespoke system and fingerprint images from a commercially available system. The work has led to several findings:

- The biometric information density of hand vein images is around 40 times lower than that of fingerprint images.
- The percentage of informative image blocks with respect to the optimum image blocks is similar for both hand vein and fingerprint images, and less than 15%.
- The potential recognition capacity for hand vein images is around 10^{15} for an area of 2268 mm^2 which has not been reported before.
- The potential recognition capacity for fingerprint images is around 10^{36} for an area of 25 mm^2 which is in good agreement with the previous estimates.
- The estimated recognition capacity decreases as the image quality decreases, and the estimation is likely to be more consistent for large image blocks with low information density.

However, it is important to recognise that these results represent crude estimates, since they were obtained from a relatively small dataset based on two texture classes which were assumed to be statistically independent and uniformly distributed. From repeated experiments on hand vein and fingerprint images, it is observed that consistent results have been produced for large image blocks with low information density, but this has not been the case for small image blocks with high information density. Therefore, larger datasets are needed in order to establish a higher level of statistical confidence with more consistent results, especially if the proposed method is to be applied to biometric modalities containing feature rich patterns. Furthermore, it should be noted that the recognition capacity estimates are closely related to the choice of the informativeness threshold value. In the proposed method, the informativeness threshold value is set as a function of the dataset size, and the use of a relatively small dataset may result in a higher value for informativeness threshold with more image blocks treated as informative, resulting in overestimates of the potential recognition capacity.

Nevertheless, the proposed methodology is seen to offer an alternative approach for estimation of the potential recognition capacity of biometric systems with data-driven advantages. In addition to using large datasets to refine recognition capacity estimation and application of the proposed method to other texture-based biometrics, such as face, finger knuckle and palmprint, it will be interesting to explore the possible use of the proposed method as a performance indicator for comparative assessment of individual factors affecting texture-based biometrics, and to investigate the relationship of the proposed method with other information theoretic approaches.

5. Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61673021 and Beijing Natural Science Foundation under Grant KZ201410009012.

6. References

[1] Corcoran, P.M.: ‘Biometrics and consumer electronics, A brave new world or the road to dystopia?’, IEEE Consumer Electron. Magazine, 2013, 2, (2), pp. 22-33

[2] Alonso-Fernandez, F., Fierrez, J. and Ortega-Garcia, J.: ‘Quality measures in biometric systems’, IEEE Security & Privacy, 2012, 10, (6), pp. 52-62

[3] Unar, J.A., Seng, W.C. and Abbasi, A.: ‘A review of biometric technology along with trends and prospects’, Pattern Recognition, 2014, 47, pp. 2673-2688

[4] Galton, F.: ‘Finger prints’, (McMillan, 1892)

[5] Stoney, D.A.: ‘Measurement of fingerprint individuality’, in Lee, H. C. and Gaensslen, R. E. (Eds.) ‘Advances in fingerprint technology’ (CRC Press, 2001)

[6] Abraham, J., Champod, C., Lennard C. and Roux, C.: ‘Modern statistical models for forensic fingerprint examinations: A critical review’, Forensic Science International, 2013, 232, pp. 131–150

[7] Maltoni, D., Maio, D., Jain, A. K. and Prabhakar, S.: ‘Handbook of fingerprint recognition’, (Springer, 2009)

[8] Roxburgh, T.: ‘Galton’s work on evidential value of fingerprints’, Indian Journal of Statistics, 1933, 1, (1), pp. 50-62

[9] Osterburg, J., Parthasarathy, T., Raghaven, T. and Sclove, S.: ‘Development of a mathematical formula for the calculation of fingerprint probabilities based on individual

characteristics', Journal American Statistic Association, 1977, 72, (360), pp.772-778

[10] Henry, E.: 'Classification and uses of finger prints', (Routledge, 1900)

[11] Gupta, S. R.: 'Statistical survey of ridge characteristics', Int. Criminal Police Review, 1968, 5, (2018), pp. 130-134

[12] Roxburgh, T.: 'On evidential value of fingerprints', Indian Journal of Statistics, 1934, 1, pp. 189-214

[13] Trauring, M.: 'Automatic comparison of finger-ridge patterns', Nature, 1963, 197, (871), pp. 938-940

[14] Stoney, D. A.: 'A quantitative assessment of fingerprint individuality', PhD thesis, (University of California, Davis, 1985)

[15] Pankanti, S., Prabhakar, S. and Jain, A. K.: 'On the individuality of fingerprints', IEEE Trans. Pattern Analysis and Machine Intelligence, 2002, 24, (8), pp. 1010-1025

[16] Chen, J. and Moon, Y. S.: 'A minutiae-based fingerprint individuality model', IEEE Conf. Computer Vision & Pattern Recognition, Minneapolis, USA, 2007, pp. 1–7

[17] Chen, J. and Moon, Y. S.: 'The statistical modelling of fingerprint minutiae distribution with implications for fingerprint individuality studies', IEEE Conf. Computer Vision & Pattern Recognition, Anchorage, USA, 2008, pp. 449–455

[18] Lim, C. and Dass, S. C.: 'Assessing fingerprint individuality using EPIC: a case study in the analysis of spatially dependent marked processes', Technometrics, 2011, 53, pp. 112–124

[19] Dass, S. C., Zhu, Y. and Jain, A. K.: 'Statistical models for assessing the individuality of fingerprints', Proc. IEEE Workshop on Automatic Identification Advanced Technologies, Buffalo, USA, 3–9

[20] Zhu, Y., Dass, S. C. and Jain, A. K.: 'Statistical models for assessing the individuality of fingerprints', IEEE Trans. Inf. Forensics and Security, 2007, 2, (3), pp. 391–401

[21] Su, C. and Srihari, S. N.: 'Probability of random correspondence for fingerprints', in Geradts, Z. J. M. H., Franke, K. Y. and Veenman C. J. (Eds.), Computational Forensics, Springer Lecture Notes in Computer Science, 2009, 5718, pp. 55–66

[22] Su, C. and Srihari, S. N.: 'Evaluation of rarity of fingerprints in forensics', Advances in Neural Information Processing Systems, 2010, 23, pp. 1207–1215

[23] Chen, Y. and Jain, A. K.: 'Beyond minutiae: A fingerprint individuality model with pattern, ridge and pore features', in Tistarelli, M. and Nixon, M. S. (Eds.), Advances in Biometrics, Springer Lecture Notes in Computer Science, 2009, 5558, pp. 523-533

[24] Unnikrishnan, M.: 'How is the individuality of a face recognized?', Journal of Theoretical Biology, 2009, 261, pp. 469-474

[25] Daugman, J.: 'The importance of being random: statistical principles of iris recognition', Pattern Recognition, 2003, 36, pp. 279-291

[26] Daugman, J.: 'Probing the uniqueness and randomness of IrisCodes: Results from 200 billion iris pair comparisons', Proc. IEEE, 2006, 94, (11), pp. 1927-1935

[27] Daugman, J.: 'Information theory and the IrisCode', IEEE Trans. Inf. Forensics and Security, 2016, 11, (2), pp. 400–409

[28] Willems, F., Kalker, T., Goseling, J. and Linnartz, J.-P.: 'On the capacity of a biometrical identification system', IEEE Int. Symp. Information Theory, Yokohama, Japan, 2003, pp. 82

[29] Schimid, N.A. and Nicolo, F.: 'On empirical recognition capacity of biometric systems under global PCA and ICA encoding', IEEE Trans. Inf. Forensics and Security, 2008, 3, (3), pp. 512-528

[30] Bhatnagar, J. and Kumar, A.: 'On estimating performance indices for biometric identification', Pattern Recognition, 2009, 42, pp. 1803-1815

[31] Adler, A., Youmaran, R. and Loyka, S.: 'Towards a measure of biometric feature information', Pattern Analysis & Applications, 2009, 12, (3), pp. 261-270

[32] Haralick, R., Shanmugam, K. and Dinstein, I.: 'Textural features for image classification', IEEE Trans. Systems, Man and Cybernetics, 1973, 3, (6), pp. 610-621

[33] Otsu, N.: 'A threshold selection method for grey-level histograms', IEEE Trans. Systems, Man, and Cybernetics, 1979, 9, (1), pp. 62-66

[34] Wang, Y., Xie, W., Yu, X. and Shark, L.-K.: 'An automatic physical access control system based on hand vein biometric identification', IEEE Trans. Consumer Electronics, 2015, 61, (3), pp. 320-327

[35] 'Fingerprint readers' http://www.crossmatch.com/, accessed 23 March 2017

[36] Wang, Y., Zhang, K. and Shark, L.-K.: 'Personal identification based on multiple keypoint sets of dorsal hand vein images', IET Biometrics, 2014, 3, (4), pp. 234-245

[37] Wang, Y., Duan, Q., Shark, L.-K. and Huang, D.: 'Improving hand vein recognition by score weighted fusion of wavelet-domain multi-radius local binary patterns', International Journal of Computer Applications in Technology, 2016, 54, (3), pp.151-160

[38] Papoulis, A. and Pillai, S. U. 'Probability, random variables and stochastic processes', (McGraw-Hill, 2002)