Non-Invertible Online Signature Biometric Template Protection via Shuffling and Trigonometry Transformation

Fahad Layth Malallah
Computer and Embedded System Engineering
UPM / Malaysia

Sharifah Mumtazah bt Syed Ahmad Abdul Rahman,
Senior Lecturer, PhD,
Computer and Embedded System Engineering,
UPM / Malaysia

Wan Azizun bt. Wan Adnan,
Senior Lecturer, PhD,
Computer and Embedded System Engineering,
UPM / Malaysia

Salman bin Yussof
Assoc. Professor,
Systems and Networking
UNITEN / Malaysia

ABSTRACT
This paper describes a novel approach to a cancelable template protection scheme that secures online handwritten signature samples in the reference database of a biometric verification system. We propose a renewable-noninvertible transformation process named Bio-Trigono comprising two consecutive steps. First, a shuffling scheme is applied to a signature sample to attain the renewability property for template protection. This is followed by the deployment of a cosine function for which its periodic characteristic is exploited to achieve a much desired non-invertible property for additional security. The overall template protection scheme was tested rigorously on signature samples of a SIGMA database through an online signature verification system. Its verification utilized the Principal Component Analysis (PCA) for features' extraction and the Artificial Neural Network (ANN) for user reference modeling and classification processes. Results demonstrated an effective cancelable template protection scheme whereby the best averaged error rates were 10.3, 10.5 and 14.1% for un-transformed first and second transformed signature templates, respectively.

General Terms
Pattern Recognition and Security.

Keywords
Artificial neural network, authentication biometrics,cancelable biometrics, principal components analysis, security, signature verification, template protection.

1. INTRODUCTION
Information security aims to maintain confidentiality, integrity and availability of information [1, 2]. User authentication is one of the most important operations in information security and the authentication operation can be based on one of three different modalities: (i) something you know; (ii) something you have; or (iii) something you are [3]. The first modality relies on knowledge factors such as a password. The second refers to ownership factors such as a user ID card and/or security token. The third describes biometric authentication based on inherent factors that specifically pertain to user identity or functions such as fingerprints, facial features, retina scan, signature and voice imprinting [4, 5]. Traditional approaches are based on the first two modalities which, however, have inherited a number of drawbacks. The first modality can be guessed or ascertained (cracked) through dictionaries or brute force attack. The second modality presents risks of loss, theft or duplication. The third modality concerns emerging authentication techniques based on biometrics that are presently hailed as more fool-proof and reliable [6]. The handwritten signature is a behavioral biometric and one of the most accepted since the majority of users is accustomed to writing their signatures; thus, it has played a well established role as a token identification marker for decades. Signatures, therefore, remain identity representations for critical applications such as online legal documents and financial transactions [7]. With the rapid advance of capture devices such as tablets and smart phones, there is huge potential for online signature biometrics whereby signatures are written on electronic devices and verified automatically. Being similar to other biometrics, online signature verification systems require security mechanisms for protection. There are five basic components of a biometric, namely:(1) a device sensor that captures input samples; (2) a feature extractor module that extracts salient features; (3) a template database that stores referenced biometric templates; (4) a matcher module that compares the tested sample with referenced templates; (5) an application device that outputs the biometric decision [8, 9]. However, this scheme has eight possible attack points as illustrated in Figure 1 [10], the first of which includes attacks on the template database, deemed as one of the most damaging attacks on a biometric system. An online signature biometric requires that signature samples are stored in a database for use as a reference model to verify a sample in question. However, this poses a security threat since the database may be vulnerable to security attacks [10]. Specifically, storing signature templates in a database exposes them to three vulnerabilities that can lead to unauthorized access [8, 10]. First of all, the template may be replaced by an imposter's template. Secondly, physical spoofing can be created from the template. Finally, the template can be stolen and replayed to an authentication system.

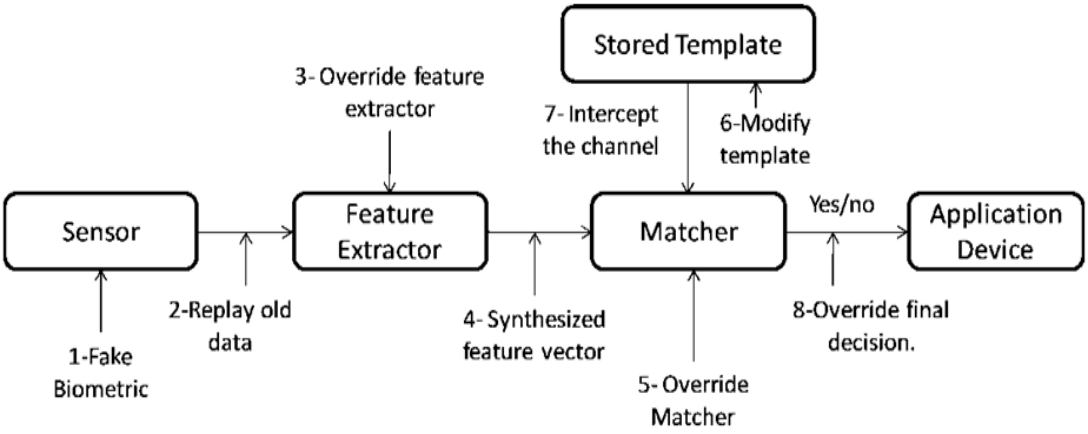

Fig. 1: Attack points in the biometric system, (adapted from [10]).

Unlike passwords or tokens that are easily replaced or reissued should they be compromised, human signatures are established over time and it is very difficult to compel anyone to change his/her signature. Thus, a template protection scheme becomes necessary for online signature biometrics. The ideal protection scheme should be non-invertible. In other words, it would be computationally difficult to reverse any reconstruction of an original template to an artificially transformed substitute, even if secret transformed parameters and stored templates are known. In addition, the proposed algorithm should be secure enough to resist different types of attack such as a brute force attack. Next, the accuracy of an online signature biometric scheme should not be compromised with the deployment of template protection. This includes the requirement for renewable templates that allows template reissuance should compromise occur [10]. These characteristics comprised our prime design objectives for the online signature template protection scheme described in this paper.

This paper is divided into six sections. Signature biometric and related works are presented in Section 2. In Section 3, a framework design for this research is given which comprises a non-invertible template protection scheme called BioTrigono, as well as its security analysis and signature verification protocols. In Section 4, our dataset and experimental trials are presented. Results and discussion follow in Section 5 and finally, a summary and conclusion are presented in Section 6.

2. SIGNATURE BIOMETRICS AND RELATED WORK

Dynamic signature template protection was first proposed by Vielhauer et al.(2002 [11]) where the protection was implemented by using the hash method. Another system was proposed by Feng and Chan (2002[12]) that involved a private key generation from online handwritten signatures. In their work, online signature features were used as a private key for a digital signature cryptosystem named BioPKI. The main purpose of BioPKI was to eliminate the vulnerability of private key storage that resolved key management matters. Another system was proposed by Freire-Santos et al.(2006[13]) where protection was based on a key binding crypto-system using fuzzy vault.

In this work, template protection was implemented through an encoding operation that utilized two values. The first value was a random *k*-bit value used as a secret key, which was then protected by the vault code. The second value was online signature features. Recently, the fuzzy vault was proven vulnerable to the multiplicity attack [14]. This type of adversarial attack is achieved by taking two different vaults computed from the same data and then guessing the genuine points. This particular vulnerability implies the difficulty of satisfying the property of renewability by using the fuzzy vault system. Yip et al. (2006[15]) suggested an online signature template protection system using the salting approach. They derived a secret key by combining actual signature feature coordinates, as well as velocity and acceleration with random numbers. Yet another salting approach was proposed by Freire et al(2008[16]) which implemented an XOR operation between the feature vector (after Feature Extraction, Binarization and Feature Selection) and a random code, which had already passed on to an Error Correction Code (ECC) operation such as the binary BCH code [17]. Later on, a biometric cryptosystem approach using a key binding class based on fuzzy commitment template protection was presented by Maiorana and Campisi (2010 [18]). Their approach managed to achieve the property of renewability. Its strength, as claimed by the author, was in its recognition rate where both unprotected and protected recognition rates were roughly the same. This technique of 'fuzzy commitment' was based on prior work done by Juels and Wattenberg (1999 [19]). Another system was proposed by Maiorana et al. (2010 [20]) where a non-invertible transformation (cancelable template) of online signature templates, called BioConvolving, was presented. The idea for the transformation was first proposed in 2008 [21, 22], but their 2010 effort enhanced the technique by adding the property of renewability. The technique was considered non-invertible because retrieving the original template from the transformed template was just as hard as random guessing. The transformed template was generated using linear convolution among random sequences. As claimed by the author, the security of the transformed template for the BioConvolving technique depended on blind de-convolution [23] to retrieve the original template. Based on our observations, we believed that the primary drawback of the

BioConvolving cancelable transformation system was the length of its transformed template, which was not at the same length of the original one. Another published work presented an hybrid approach to online signature template protection (Rúa 2012 [24]). He combined feature transformation with a biometric cryptosystem. The former exploited the renewability property whereas the latter (based on fuzzy commitment) exploited its strength to provide non-invertibility and thus, manage the intra-class variability of signature samples.

3. FRAMEWORK DESIGN

The online signature samples utilize for this study comprise time series signals of horizontal $x[t]$ and vertical $y[t]$ coordinates, as well as pen pressure $p[t]$ sampled at time t. These samples were then fed into the system for protection and verification. A block diagram of the system's overall architecture is shown in Figure 2 which illustrates four principle stages. The first is normalization of the signature sample to a fixed or desired sampling length. The second is the proposed non-invertible template protection scheme based on hybrid shuffling and trigonometric transformation

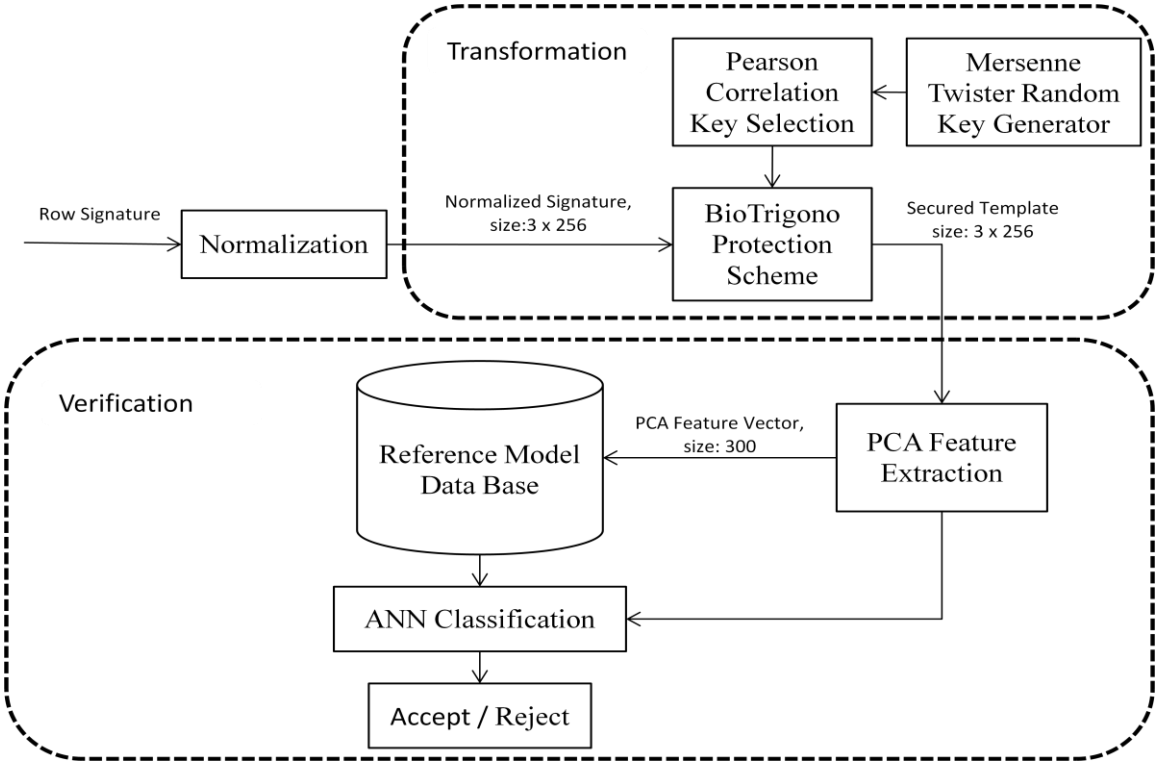

Fig 2: Proposed System for structure normalization, protection and biometric verification.

His proposed feature transformation was based on the Universal Background Model (UBM) [25]. The reason for choosing the UBM framework as a part of an online signature verification system was to provide accurate user characterization despite small enrollment features. The essence of protection comes via the XOR operation between the user template and the error correction code. Finally, a new cryptosystem approach was presented by Maiorana et al. (2012[26]) for online signature template protection using turbo code and modulation constellation. The main purpose of the turbo code was to achieve a high Error Correction Code (ECC) which was then exploited to correct errors of intra-user variability for biometric signatures due to their original use in digital communication by correcting data errors received after transmission. The modulation constellation proved beneficial for soft-decoding modality and resulted in a flexible framework. The protection relies on key binding as a fuzzy commitment approach. Additional details and explanation of signature template protection and verification systems were elaborated by Malallah et al. (2013 [27]).

The third stage is an online signature verification system tasked to classify the queried signature sample as either 'accepted' or 'rejected' for the claimed identity. Here, Principal Component Analysis (PCA) and Artificial Neural Network (ANN) were used as feature extraction and classifier modules, respectively. The fourth stage comprised (i) a Pseudo-Random Number Generator (P-RNG) utilizing the Mersenne Twister (MT) technique; and (ii) Pearson's correlation to select a suitable key for template protection transformation (details on Pearson's correlation are discussed in Section 5.3). The performance of the system was then evaluated based on the False Accept Rate (i.e. the rate whereby forged signatures were accepted by the system), and a False Reject Rate (i.e. the rate whereby genuine signatures were rejected by the system).

3.1 BioTrigono Non-invertible Template Protection Scheme

Since online signatures suffer intra-user variability, normalization with regard to time is crucial. In this research, the desired length was designed to be 256 signal sampling parts for all users in the database as it was close to the average

length of signatures in the SIGMA database of 200 users. This normalization technique is fully explained and given by Malallah et al.(2013 [28]). The main reason for signature normalization is security in order to assign a fixed length user-specific key for protection, which, in turn,

to form a new template. The length of the key was 256 decimal numbers as assumed and described and cited above. A key was then assigned uniquely to a particular user. Four sampling points from each signature signal were combined into one block, (i.e. a block size consists of four sampling points) to achieve higher capability for the renewability property. Figure 4 illustrates the shuffling scheme's implementation.

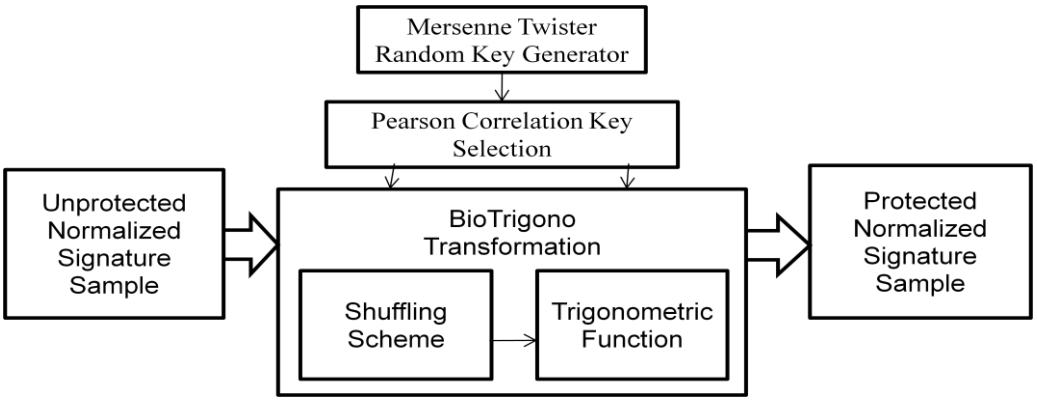

Fig. 3: BioTrigono non-invertible template protection scheme.

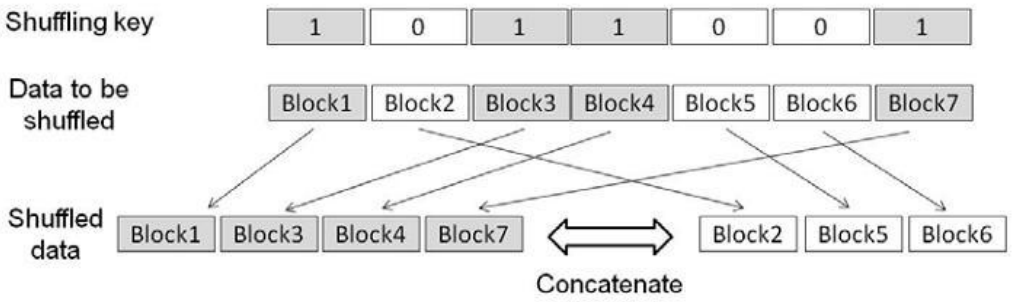

Fig. 4: Shuffling scheme, (adapted from [8]).

supports renewability because it is difficult to assign a key to users with variable signing duration for signature samples. The normalized signature sample was then passed on to the non-invertible transformation process for protection. The proposed transformation method comprised two steps: a Shuffling Scheme and a Trigonometric Function. The former was exploited to expand the renewability property. The latter was employed to make it non-invertible or 'cancelable'. Each step requires a key for implementation. Both steps use the same specific user key. This key was generated randomly using the Mersenne Twister pseudo-random number generator. Figure 3 illustrates the proposed BioTrigono non-invertible transformation scheme.

Shuffling Scheme: This was based on similar cryptography concepts using scrambling and permutation operations on the content of input signature samples to output a scrambled counterpart for security purposes. It decoupled each signal of the online signature template into two groups based on a shuffling key [8, 29, 30]. The first group consisted of odd values while the second group consisted of even values. A concatenation operation was then done between both groups

Trigonometric Transformation: let $b[t]$ be an input vector of float numbers considered as one of the three signals, $x[t]$, $y[t]$ and $p[t]$ from a signature; and let $A[t]$ be its output vector. Let k be a key vector having the same length as the $b[t]$ input vector. The proposed trigonometric transformation equation (1) follows:

$$A[t] = b[t] + 3k\,cos(\,b[t]),\, t = 1 \ldots 256 \quad (1)$$

As derived, this specified equation was based on the need for a non-invertible function, which is characterized as a function that is easy to compute but hard to invert; or a 'one-to-many' function [2, 8, 10]. The cosine trigonometric, being a periodic function, thus fulfills the 'one-to-many' function criterion. Furthermore, it is preferable to use it in a radian angle to obtain a higher number of periods than a degree angle for a specific range. This was achieved by exploiting a cosine function that generated -1 in the following case:$n\pi$,

where n = 1, 3,5,7 ...as shown in Figure 5. (i.e. cosine waveform is a periodic function).

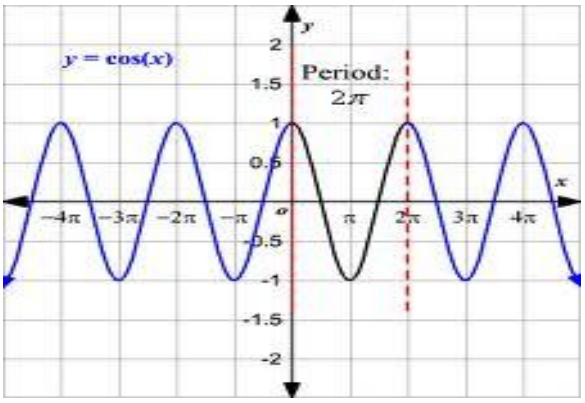

Fig. 5: Cosine waveform in radian angle showing several periods

For example, the inverse of cosine -1 can be $n\pi$, where n = 1 , 3,5,7 ...within a specific period. Hence, it leads to a non-invertible function as there is no exact solution; especially if the range of the input is long such as in our database from 0 to 1024. In addition, the $\cos(b[t])$ is added to $b[t]$ to restore the approximation of an original signature that then maintains an acceptable recognition rate. Furthermore, key k is multiplied by the factor $\cos(b[t])$ to involve key k in the transformation's computation, which is later considered an essential component of renewability. Moreover, the reason of multiplying key k by 3 is to markedly enhance secrecy for worst case scenarios if $k = 1$. In other words, the factor (3) is the smallest factor needed to create a 'one-to-many' solution with the least degradation of the verification rate. For example if we use 20, verification errors will increase.

3.2 BioTrigono Transformation Security Analysis

The non-invertible transformation function is defined in the literature as being easy to compute and hard to invert, or as a 'one-to-many' function where input cannot be computed from the output [2, 10]. We demonstrated that the second part of the proposed BioTrigono transformation, as specified in equation (1), is a non-invertible function when trying to calculate b when A and k are given at specific time t using all possible mathematical methods.

1- Implicit function (algebraic geometry): getting b back from A and k is impossible for two reasons. First, b exists in two places of the proposed equation: b and $3k \times \cos(b)$; which leads to difficulty in extracting the pure formula as an inverse. Second, computing A is one-to-one but computing *b* from both *A* and *k*, is 'one-to-many', equating to a one way function.

2- Substituting $\cos(b[t])$ by exponential function cos(b) = $\frac{1}{2} \times [e^{ib} + e^{-ib}]$: once considering an exponential function in the computation, imaginary components appear which contradictreal part numbers in the research which will not lead to any solution when solving for b .

3- Numerical methods: possible solution with numerical methods by using Newton-Raphson's method or Taylor's methods. These methods include trial and error and require a number of iterations. Furthermore, the termination is based on a specific error. Trying a number of iterations means trying several possible solutions. Hence, this testing method is similar to a brute force attack, discussed later. Furthermore, there is no zero error in numerical methods, which means there is no exact solution. Accordingly, this type of mathematics is absolutely unable to crack the message because some numerical solutions are 'one-to-many' mapping or even 'one-to-null' mappings which lead to no solution.

4- Graphic calculator: a possible method could recover b from A and k if, and only if, $k \leq 1$. Therefore, in the equation's design, the constant number 3 is multiplied by k to ensure that k is always larger than 1, even if P-RNG generates $k = 1$.

The argument here is that BioTrigono is non-invertible as characterized by the 'one-to-many' criterion even for the worst case scenario where P-RNG generates $k = 1$. An example of cracking the equation using the Graphic Calculator is as follows:

Let $k = 1$ as the worst case, and $b[1] = 170$. By substituting them in equation (1), the following results are obtained:

$$A = 170 + 3 \times 1 \times cos(170) \rightarrow \; A = \; 172.8140$$

Now, if an adversary wants to crack 172.8140 to its original message (170) given key ($k = 1$), she/he will apply the graphical method and compute Eq.(2), trying to obtain b as:

$$172.8140 = b + \; 3 \times 1 \times cos(b) \qquad (2)$$

The intruder will then divide equation (2) into two graphs ($Graph_1$and $Graph_2$, as a graphical calculator method), and afterwards draw them together to obtain the intersecting point(s):

$$\text{Graph}_1 = 172.8140 - \text{b}$$
$$\text{Graph}_2 = 3 \times \text{k} \times \cos(\text{b})$$

By guessing, $\text{b} = 1{:}500$ ("trial and error") and by then drawing Graph_1 and Graph_2. The intersecting point(s) is/are considered the solution for b . It is clear in Figure 6 that there are two intersection points (one is true and the other, false) in case $3k \times \cos(b)$. Consequently, two possible results derive from the worst case scenario, which, by itself means one-to-many; thus, leading to a non-invertible function.

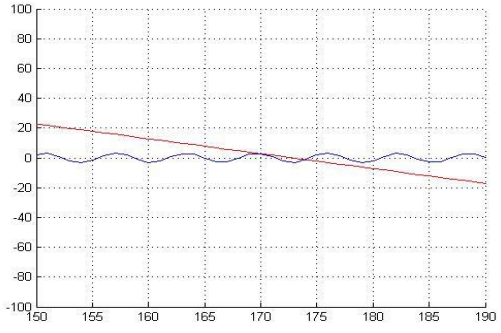

Fig. 6: Two intersection points in case $3k \times \cos(b)$.

On the other hand, Figure 7 illustrates a case without 3 multiplied and with $k \times \cos(b)$, with the worst case being $k = 1$. By using the graphical calculator, the output is one intersection point and the message is easily recovered. Thus, the constant value of 3 in the equation is important to ensure security by crafting BioTrigono as non-invertible.

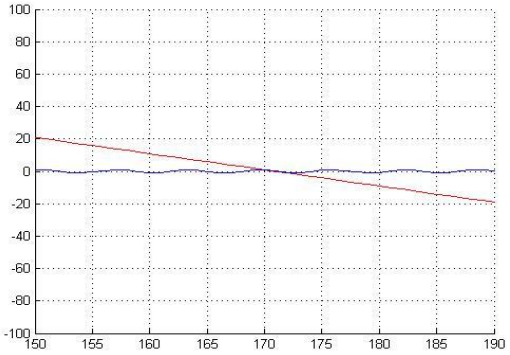

Fig.7: One intersection point in case $k \times \cos(b)$

Brute force attack is a type of "trial and error" attack that can be considered a numerical method test. Using this approach, an attacker tries every possible key until she/he discovers the correct one [10, 31]. No encryption algorithm is entirely safe from the brute force method. However, if the number of possible keys is high enough, program cracking becomes very difficult for the brute force approach because the longer the password or key, the more difficult it becomes to crack. In particular, if an attacker wants to crack a signature template to its original form using the brute force method, the operation proves extremely complex due to the extended time interval required to try all possible keys. Such a time interval can be estimated as follows:

The template consists of 300 features as a signature represented vector. If the attacker uses a graphical calculator and the signature was transformed by using a key, which by chance and for illustration purposes consists of 1 for the entire key vector as $k = [1_1, 1_2, 1_3, \ldots, 1_{256}]$ (very unlikely to happen); there will be 2-to-1 mapping or two solutions for each trajectory (one of which is true) as explained in Figure 6. Both choices can be expressed by one bit (either "0" or "1"). Once there are 300 *Xs* (feature vector length will be explained later) two choices, there are then 2^{300} possible choices. Computing all possible permutations by a workstation, 2^{300} results in 2×10^{90} possible templates; one of which is the original template. Let say an attacker uses China's Tianhe-2 supercomputer—the fastest supercomputer available as of June 2013), capable of performing 33.86×10^{15} floating point operations per second [32]—then the number of years required to try all possible template combinations is given by Eq.(3) below:

$$time\ per\ year = \frac{2^{300}}{33.86 \times 10^{15} \times (3600 \times 24 \times 30 \times 12)} \quad (3)$$

This equates to 1.9×10^{66} years. Consequently, the required time to crack the genuine or close to the genuine template by brute force attack is exceptionally long.

3.3 Signature Verification

The online signature verification system comprises two modules: Principal Component Analysis (PCA) for feature extraction, and Artificial Neural Network (ANN) for classification.

3.3.1 Principal Component Analysis (PCA)

One of the main reasons of using feature extraction in a computer vision or pattern recognition system is to improve the accuracy of the recognition rate [33]. For the purposes of this research, feature extraction was used to transform signature signals from the original time series domain to another domain which then maximized variance while decreasing correlation between genuine and forged signature samples [34]. Principal Component Analysis (PCA) was used to improve the recognition rate due to its ability to transform data sets (signatures in this case) from a correlated domain to another domain characterized as highly uncorrelated among original data sets [34, 35].

For this work, the input for PCA were signature samples represented by three time series signals based on horizontal $x[t]$ and vertical $y[t]$ signals as trajectories derived from the signature samples; and these were complemented by pen pressure $p[t]$ signal values. After PCA implementation, three component vectors ($comp$) from PCA output were then combined into one vector to represent a signature sample. Furthermore, Eigen values (Eig_{vlu}) and Eigen vectors (Eig_{vec}) were also added to the signature feature vector to consolidate the represented feature vector of the signature sample. The latter was then passed on to ANN for final classification using the equation (4):

$$No.ofFeatures = Eig_{vec} + Eig_{vlu} + 3comp \quad (4)$$

Each of the three ($comp$) has 256 features. By combining them with three (Eig_{vlu}) and nine (Eig_{vec}) to one vector, the number of features representing the vector is now 780. However, by doing this, the signature feature vector becomes too long, making it unpractical for classification as too problematic for the speed of ANN training and testing. Therefore, feature selection based on equal segments was employed to reduce the feature vector length while maintaining recognition accuracy. For our purposes, feature selection was done empirically. Selection was implemented for each component vector by dividing the vector ($comp$) into eight segments (*seg_xx*) where each segment size had 32 features. This was accomplished by taking the 1st (seg_11), 4th (*seg_14*) and 8th (seg_18) segments from among the eight segments from the first ($comp$) vector, and doing likewise for the second and third ($comp$). The reason being that all three segments were equivalent to first, middle and last partitions of the signature. Finally the length of each signature feature vector was computed using Eq.(5):

$$Final\ No.of\ Features = Eig_{vec} + Eig_{vlu} + 3\ (3\ seg\_xx) \quad (5)$$

Since each (*seg_xx*) consists of 32 features, the feature vector length for each signature sample equaled 300 features. Table 1 shows selected features for feature vectors from each signature sample, where each underlined cell was considered as included in the final signature's representative vector. Eventually, the length of the final feature vector was 300 floating point numbers.

Table 1. Explanation of the representative vector construction for each signature sample consisting of underlined features

Eigen-Vector	Eigen-Value	Three Component Vectors (seg_(row, column))							
c11 c12 c13	Highes t-v1	seg _11	seg _12	seg _13	seg _14	seg _15	seg _16	seg _17	seg _18
c21 c22 c23	Middle -v2	seg _21	seg _22	seg _23	seg _24	seg _25	seg _26	seg _27	seg _28
c31 c32 c33	Lowest -v3	seg _31	seg _32	seg _33	seg _34	seg _35	seg _36	seg _37	seg _38

(i.e.) *c21*: first value of the second Eigen vector; *v1*: first Eigen value, *seg_38*: 8th segment of the 3rd component vector. The feature vector for each signature sample consisted of the underlined values in Table 1. The order was: *c11, c12, c13, c21, c22, c23, c31, c32, c33, v1, v2, v3, seg_11, seg_14, seg_18, seg_21, seg_24,seg_28, seg_31, seg_34, seg_38*. Each *cxx* and *vx* was one floating point number, while *seg_xx* held 32 floating point numbers.

3.3.2 Artificial Neural Network (ANN)

We employed the Multi-layer perceptron (MLP) in this research. It is a feed-forward artificial neural network model that maps sets of input data into a set of target outputs [36, 37]. MLP is a multiple layer system (input layer, hidden layer(s) and output layer) where each layer contains several nodes (cell). Each node is stimulated according to an activation function. Every node is connected to subsequent nodes in the next layer (full connection), but no connections exist for nodes in the same layer. The training type of MLP is a supervised learning technique called back-propagation of the training network [38]. MLP is therefore able to recognize data that are linearly or not linearly separable [39].

In the network construction, the number of nodes in the input layer is the same as the number of input features of the represented vector. In the current case, represented signature features were 300 real numbers (as discussed in the PCA section). Therefore, the input layer consisted of 300 nodes. The number of output nodes must be able to identify the general category of the state of the system [40]. The proposed verification system had only one output node, that being whether or not the signature sample was accepted as the claimed identity.

The characteristics of the proposed ANN classifier were as follows:

1- Layers: 2 hidden layers; the first and the second hidden layers comprised 80 and 40 nodes, respectively, with the final output layer having one node.

2- Training algorithm: the Scaled Conjugate Gradient (SCG) algorithm was used because it was proven valid by Moller (1993[41]) and thus suitable for the efficient training of a high number of features.

3- Activation Function: The tangent sigmoid function (to activate -1 and +1 threshold output) was used. Accordingly, score results from -1 to 0 are classified as 'forged signature', while results from 0 to +1 validate the signature sample as 'genuine'.

4- Number of trained iterations was 150: This was set experimentally. Using a larger number of iterations may cause an overtraining problem that undermines generalization.

5- Learning rate was 0.3.

All ANN specifications cited above were chosen empirically after intensive MATLAB trials. Figure 8 illustrates the finalized ANN topology according to the proposed network's structure.

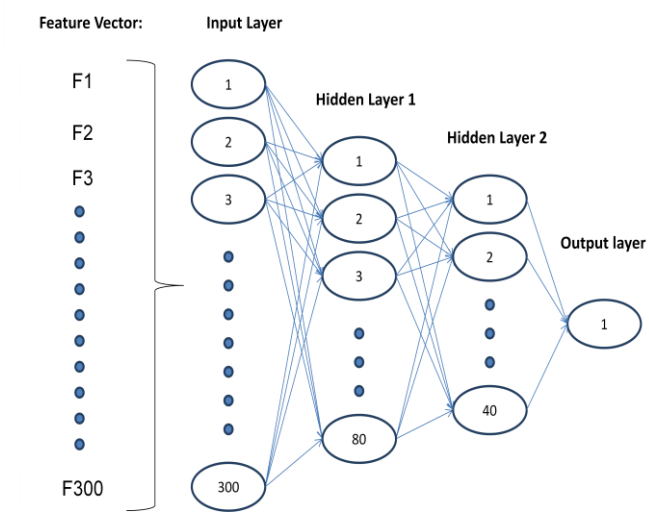

Fig. 8: Artificial neural network structure for the proposed recognition system.

Input values for ANN are bipolar (-1 or +1), or binary (0 or +1), or continuous real numbers within a given range [42]. It is preferable to map the original input range to (-1 ~ +1) to obtain a higher verification rate because the target range is wider than (0 ~ +1). Here is a schematic illustration of the initial type of input used for this research.

One problem that might occur during neural network training is called overtraining (over-fitting) [36] a problem that affects system Generalization. The term 'Generalization' measures the capability of ANN to recognize samples from outside of the training samples [43]. Hence, in order to improve generalization for our proposed recognition system, overtraining (over-fitting) of ANN was necessarily avoided. Avoiding overtraining requires the cessation of neural network training at the very point when overtraining begins. Generally, ANN training stops in the following cases: (i) if the number of iterations exceeds certain epochs; (ii) if the performance function drops below goal; (iii) if the magnitude of the gradient is less than min-grade; (iv) if training time exceeds the set time; and (v) if the validation error exceeds the set number of validation errors, which is related to the early stopping technique [36]. In this research, improvement of generalization was done by using the interleave division method [36] which cycles samples between training, validation and test sets according to percentages. The rates of division were 70% for training, 15% for validation, and 15% for testing. The maximum error number for validation which then stopped the training was set to 6 by default.

4. DATASET AND EXPERIMENT

The proposed transformation and verification methods were tested with the SIGMA signature database [44] and constructed by taking signature samples from Malaysians. The database held 6,000 genuine and more than 2,000 skillfully forged signature samples. It contained two modes: online and

offline signature. Online signatures were captured as a series of coordinates using tablet devices used as an electronic pen tracer. Offline signatures refer to static images written on paper and scanned as a digital image. Online signatures in SIGMA database were captured using the Wacom Intuos3 A4 digitizing tablet, which has the following specifications: tablet resolution is 5080 dpi; surface area is A4 size; and sampling frequency is 200 points per second with 1,024 levels of pressure sensitivity. The online signal was represented by $x[t]$ and $y[t]$ coordinates along with pen pressure $p[t]$. Intra-user variability for the signature biometric was taken into consideration when compiling the SIGMA database. This consideration was implemented by requiring a subject to provide at least ten signature samples on three different days chosen at random. Intra-user variability is also a crucial element when studying human signatures since a signature is affected by emotions, writing posture and health. The total duration for data collection took three months and involved 213 contributors. Most contributors provided thirty samples as genuine signatures. Ten forged signature samples were provided (by a third party) for each user as skilled forged signature samples. For each subject, the forger was given sufficient genuine signature samples and ample time to practice the forging. Forged samples were manually assessed for resemblance to original signatures before acceptance into the SIGMA database.

However, we only utilized samples from 200 individuals. This was mainly because several subjects did not provide a complete set of thirty genuine signature samples, or we lacked a completed set of ten forged signature samples. Each signature was represented by three time series signals based on horizontal $x[t]$ and vertical $y[t]$ signals as trajectories of the signature sample, in addition to pen pressure $p[t]$ signal values. The experiment was conducted on an online signature template both with and without template protection. What follows is the experimental protocol:

1- From the SIGMA database [44], a training matrix was built comprising ten genuine samples and ten forged samples (five were skillfully forged and five were randomly forged samples) for each of 200 users. Each signature sample was represented by 300 features. Thus, the training matrix size was $[300 \times 20]$, i.e. 300 features for each sample with 20 samples for each user. Separate training was done by ANN for each user.

2- Testing ANN was done by extracting FAR and FRR for each user, separately. The testing matrix was built the same way as the training matrix with a size of $[300 \times 20]$, but different signature samples were taken from each user (20 signature samples per user), where the first ten were genuine and the next ten were forged samples, exactly as stated above.

3- In the training target for ANN, a sign $+1$ was assigned to the first ten samples of the training matrix, while -1 was assigned to the second ten samples to indicate and learn that the first set was genuine and the second were forged samples.

4- To compute the ROC curve for verification, the threshold was varied from +1 to -1 with 0.1 intervals $(+1 : 0.1 : -1)$.

5- FRR was computed by seeking test results for the first ten samples. If any of the first ten samples had a sign less than threshold, the tendency to generate False Rejection (FR) was indicated by increasing the FR counter by one $(FR = FR + 1)$ since they were supposed to be accepted (signs larger than threshold) but were wrongly rejected. On the other hand, if any of the second group had a sign greater than threshold it was considered a False Accept (FA) and then indicated by incrementing the FA counter by one $(FA = FA + 1)$. The FRR and FAR were then computed using Eqs.(6) and (7):

$$FAR = \frac{FA}{10} \times 100\% \quad (6)$$

$$FRR = \frac{FR}{10} \times 100\% \quad (7)$$

The accuracy for each user was computed by Eq.(8):

$$User_{Accuracy}\% = 100 - \frac{FAR+FAR}{2} \quad (8)$$

Considering the full complement of signatures in the SIGMA database, the average for all 200 individual accuracy scores was computed by Eq. (9):

$$AVR_{Accuracy} = \frac{1}{200} \sum_{u=1}^{200} User_{Accuracy}[u] \quad (9)$$

Now, Table 2 summarizes all parameters used in the experiments.

Table 2. Experiment Details for Accuracy Computation.

Sample Types		Training Matrix	Testing Matrix	Error Type	Error Calculation	1_user Accuracy %
Genuine		10	10	False Reject (FR)	FRR=(FR/10) * 100%	Accuracy= 100 – ((FAR + FRR)/2)
Forged	Skilled	5	5	False Accept(FA_S)	FAR=(FA/10) * 100%	
	Random	5	5	False Accept(FA_R)		

5. RESULT AND DISCUSSION

Results are now presented and include signature transformation results (Section 5.1), signature verifications using the SIGMA database (Section 5.2), and renewability analyses (Section 5.3).

5.1 Signature Transformation

The output from the shuffling scheme was passed on to the trigonometric function to generate the non-invertible (cancelable) transformed template of the online handwritten signature which we called the BioTrigono function. Figure 9 depicts the first output version of the signature transformation using the BioTriono.

It is clear that the protected signature (right side) differs from the un-protected one (left side). Moreover, the signal should be different in terms of matching from the previous version to achieve the renewability property. Later on, in cases where the first transformation was compromised, the proposed BioTrigono technique was able to generate a different version from the previous one in which both were generated from the same original signature. Figure 10 shows the signature by drawing x versus y. It can be seen that the first transformed template (on the left) was totally different from its second template (on the right).

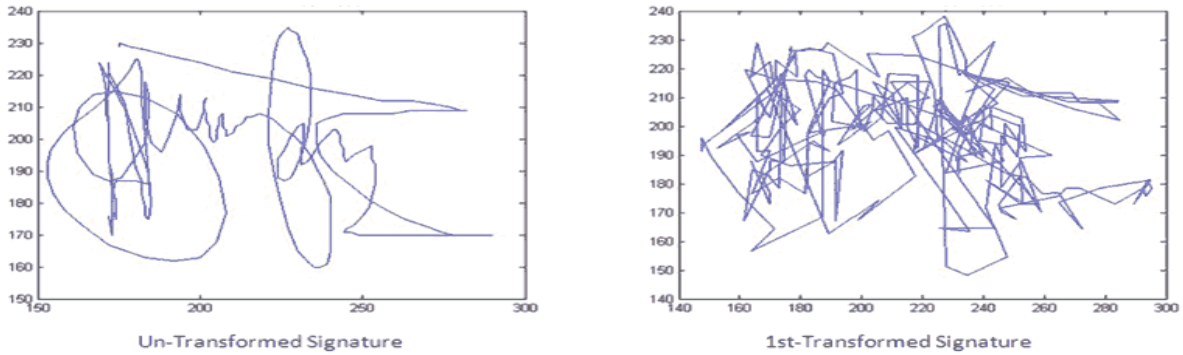

Fig. 9: First transformed signature as x in terms of y.

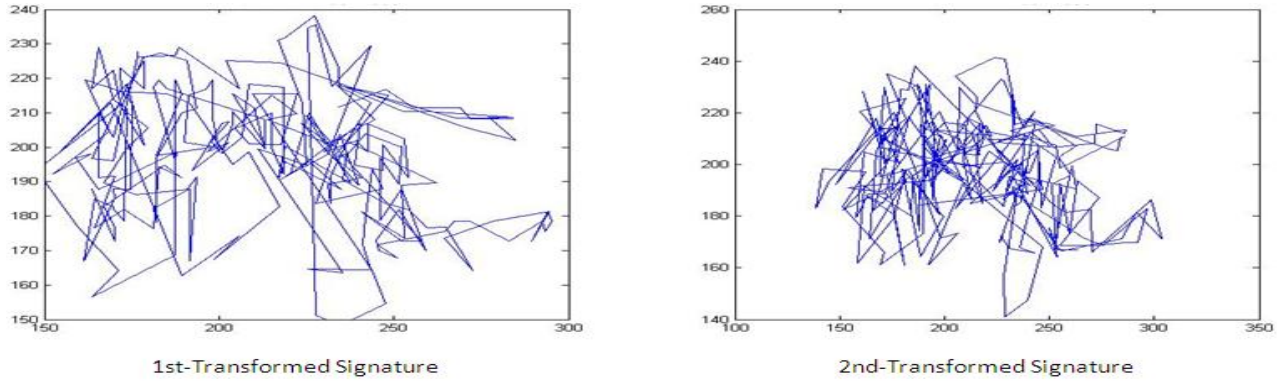

Fig. 10: Second transformed signature as x in terms of y.

The transformation from the first version to the second version was different not only in signature visualization but also in verification so that it fulfilled the renewability property. More details on renewability are given in section 5.3.

In terms of security analysis, let us say that numerical or statistical methods for this transformation ($b+3k\ cos(b)=A$) might be efficient for message extraction, which is similar to or not exactly the same as the original because biometric measurements are not exactly reproducible. Nevertheless, this transformation remained secure since this similar extracted message will not be definitely recognized with the same recognition rate as genuine signatures as the recognition error rate will be increased. Accordingly, the system administrator must adjust the threshold or level of security for the recognition rate in order to avoid undesired access. Moreover, skilled forger signature samples, which are close to genuine signature samples, were also employed and considered in this evaluation to avoid such threats.

5.2 Signature Verification

For implementation, three verification experiments were performed to estimate the error rate of the signature verification operation. Each of the three following experiments was conducted for different signature templates:

1- Original templates (before transformation).

2- First protected templates (after first transformation).

3- Second protected templates (after second transformation).

The best results were achieved when the threshold was 0, -0.1 or -0.2 as shown in Table 3. For example, the average error rates for FRR and FAR at a threshold of -0.2 were 10.3% for the first experiment (original templates); 10.5% for the second

experiment (first protected templates); and 14.1% for the third experiment (second protected templates) considered as the first renewable template. More details regarding FRR and FAR for other threshold values are listed in Table 3.

ROC (Receiver Operating Characteristic) curves for all three experiments are depicted in Figure 11. The first curve (asterisk) represents the un-protected template; the second curve (square) represents the first protected template; the third curve (circle) represents the second protected template after renewability was done for the first transformation. ROC results showed no significant degradation between the unprotected and the first protected template as the absolute difference was 0.2%. When comparing the first and second transformed template in terms of verification accuracy, there was a slight degradation of 3.6%. It is worth mentioning, however, that the limitation of the cancelable transformation type of biometric template protection is the verifiable degradation of the transformed template, while its advantage is improved security [8, 10, 20]. Furthermore, enrolled data are totally different from authenticated data for FRR and FAR evaluation. Verification error depends on the secured template's degree of complexity. In other words, more secrecy in the template results in a greater verification error rate.

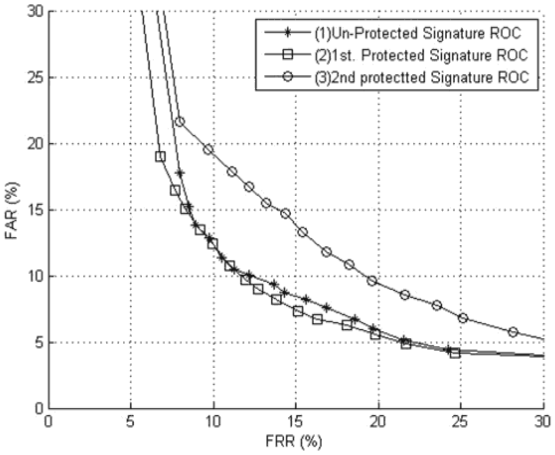

Fig. 11: Receiver operator characteristics (ROC) curves for unprotected, 1st protected and 2nd protected SIGMA database

Table 3. Result from experiments 1, 2 and 3 reporting FAR and FRR errors for several thresholds

Experiment (1), (2) and (3)									
	(1)- Un-protected			(2)- 1st-Protected			(3)- 2nd-Protected		
Threshold	FRR%	FAR%	E_AVR(1)	FRR%	FAR%	E_AVR(2)	FRR%	FAR%	E_AVR(3)
-0.2	11.9	8.7	10.3	11.35	9.7	10.525	15.4	12.95	14.175
-0.1	13.5	9.05	11.275	13.45	8.2	10.825	16.15	12.75	14.45
0	13.9	9.5	11.7	13.65	8.3	10.975	18.4	9.8	14.1
0.1	17.3	8.15	12.75	14.15	8.4	11.275	20.85	9.45	15.15
0.2	15	8.15	11.575	16.65	6.3	11.475	20.5	9.05	14.775
-0.2	100 - 10.3= 89.7%			100-10.5= 89.5%			100- 14.2=86.8%		

5.3 Signature Renewability

The proposed transformation achieves the renewability property by estimating a suitable correlation between the untransformed template and the transformed template by using Pearson's correlation coefficient metric (*r*), which is a measurement of the linear correlation between two variables, x and y. The range of r lies between +1 and -1 where +1 indicates full similarity between variables x and y, while -1 indicates a definite difference between variables x and y [45, 46]. The formula for Pearson's correlation(*r*) is given in (10):

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \qquad (10)$$

Where $\bar{x}$ and $\bar{y}$ are means of x and y vectors, respectively, and n is the length of the vector.

The main reason for using Pearson's correlation(*r*) metric between templates is due to distance metrics between keys, as cited by Maiorana [20], which can lead to unstable results in renewability analysis in the case of a proposed transformation. This is because the transformation is a random shuffling followed by a random amplitude of the cosine waveform. In addition, the employed user key for the template is also long as it has the same length as the signature template (256 sampling parts). Therefore, the difference between values for key_1 and key_2 is not sufficiently accurate to supply an indication as to whether or not there is template renewability. Hence, it became necessary to establish a method that predicted the existence or non-existence of the renewability property. This was done by designing and testing a correlation threshold (*r*) between untransformed and transformed

signature templates applicable to the SIGMA database. Renewability analysis was performed to prove that a reissued signature template using a specific key did not match another reissued signature template using an additional key, even though both templates originated with the same signer. If both templates did not match each other, the renewability property was achieved. In order to test whether the reissued templates matched or not, the following steps were performed according to the renewability test provided by Maiorana [20]:

1- Transform all SIGMA database signatures (genuine and forged signatures) using key_1, name them Transformed SIGMA Database A (TSDB_A) and then extract two ROC curves. One curve's FAR estimates random forged signatures and the other curve's FAR estimates skilled forged signatures.

2- Renew by transforming TSDB_A to TSDB_B using key_2, which has a distance, 299, (any distance might be used) from key_1. Similarly, TSDB_A is transformed to TSDB_C and TSDB_D using different keys; let's say key_3 and key_4, respectively. Table 4 summarizes key types, distances, and their transformed database names.

Table 4. Databases types, keys and distances (key distances were randomly derived).

No.	Database	Key used	Key distance
1	TSDB_A	Key_1	308 from '0'
2	TSDB_B	Key_2	299 from key_1
3	TSDB_C	Key_3	307 from key_1
4	TSDB_D	Key_4	283 from key_1

3- Next, extract the Renewable Template Matching Rate for the TSDB_B ($RTMR_B$) ROC curve using only genuine signatures from both TSDB_A and TSDB_B. Notice that FAR is estimated from TSDB_A, and FRR is estimated from TSDB_B, considering TSDB_B contains currently validated (new) genuine signatures while TSDB_A contains genuine signatures that are no longer valid because they were transformed (reissued). Similarly, the ROCs of both $RTMR_C$ and $RTMR_D$ are computed. Table 5 shows details of RTMR ROCs for all four databases.

Table 5. Renewable template matching rate (RTMR) extraction

No.	RTMR ROC	Used DB for Estimating	
		FAR	FRR
1	$RTMR_B$	TSDB_A	TSDB_B
2	$RTMR_C$	TSDB_A	TSDB_C
3	$RTMR_D$	TSDB_A	TSDB_D

4- In order to discover any mismatch, the FAR axis in the randomly forged ROC curve (in step 1) should have the same characteristics of ROC curves $RTMR_B$, or $RTMR_C$, or $RTMR_D$ from step 2, when considering the FAR axis [20].

Figure 12 shows the results of our analysis. Five ROCs are depicted that illustrate the ability of the proposed transformation to achieve renewability. The first ROC FAR_{RF} (FAR_random_forgery) is the relationship between FAR and FRR where the FAR estimation was derived from randomly forged signatures that were transformed by using key_1 (for both enrollment and authentication). The FRR estimation was derived from genuine signatures that were transformed by using key_1 as well (for both enrollment and authentication). The second ROC FAR_{SF} (FAR_skilled_forgery) was estimated in the same manner as the first ROC with the only difference being that it included skilled forged signatures instead of randomly forged signatures. Results showed an obvious degradation in the error rate when skilled forged signatures were included compared to randomly forged signatures. This was most likely due to the fact that skilled forged signatures are more difficult to recognize.

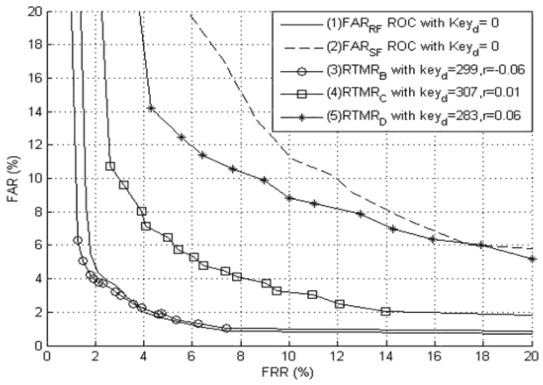

Fig. 12: Renewability Analyses in Terms of ROC

The third ROC ($RTMR_B$) (distance from key_1 was 299) had the best characteristics demonstrating renewability as its curve more closely approximated the FAR_{RF} curve. The fourth ROC ($RTMR_C$) renewability characteristics were less robust than $RTMR_B$ but better than $RTMR_D$.

Key distances used for this experiment were chosen randomly and we noted an unstable relationship between key distances and ROC renewability. As such, it was unsuitable to rely on key distance as a metric. Therefore, Pearson's correlation (r) among templates was employed to obtain the more rigid factor (r). Nevertheless, to achieve renewability, Pearson's correlation with the SIGMA database had to satisfy the following condition: $r \leq -0.06$. In other words, for any key used to transform the signature template, this condition ($r \leq -0.06$) should be met. This factor was determined by taking the best ROC, $RTMR_B$ (Figure 12) and computing the correlation for each individual in the SIGMA database between TSDB_A and TSDB_B, after which we averaged(r) for 200 individuals. The outcome was ($r = -0.06$), the most suitable correlation in terms of renewability among transformations C and D. Furthermore, in case a cancelable transformation was required for any user, the condition ($r \leq -0.06$) was made mandatory and consequently considered a general case requirement to achieve the renewability property.

Another example is presented in Figure 13 using a different transformed key and dissimilar distances, provided the proposed condition ($r \leq -0.06$) was met. For this example, a correlation (r) test was performed on a randomly chosen individual (# 58, sample 1, SIGMA database) where (r) was set to -0.08 to construct yet another transformed SIGMA database named $TSDB_{TEST}$.

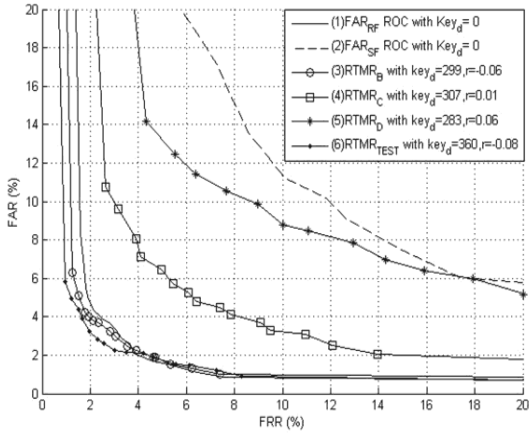

Fig. 13: $RTMR_{TEST}$ ROC with *(r)* set to -0.08 for a sample signature.

This $RTMR_{TEST}$ ROC followed the same procedure as all previous ROCs. The obvious observation was that the $RTMR_{TEST}$ ROC held a lower error rate as it lies lower than the FAR_{RF} ROC curve, meaning that a better renewability property was achieved. Nevertheless, it is not advisable to use a correlation (r) lower than -0.06 by a large value because it will lead to degradation in the verification rate due to the high complexity of the transformed template.

6. SUMMARY AND CONCLUSION

We implemented online signature template protection by designing and testing a proposed algorithm that includes both template protection (non-invertible transformation) and a verification system. Before the protection operation, the length of the signature template sample is normalized. The reason for length normalization is to set a key for each user as an essential requirement for transformation and renewability operations. The length of normalization is 256 signal sampling parts for all individuals as it approximates the average length (261 signal sampling parts) of all 200 individuals in the SIGMA database. The transformation is implemented by a Shuffling scheme followed by a Trigonometric function using the cosine waveform as a one way function to achieve the cancelable property. The recognition system is designed with PCA as the feature extractor followed by the ANN classifier. The protected feature vector of the signature sample comprises 300 features taken after PCA results and selection, after which this vector is stored as a reference model in the system database. The classifier recognizes genuine signatures for both original and transformed signatures with nearly the same recognition rates: 10.3% before transformation and 10.5% after transformation. This transformation accomplishes template renewability provided it satisfies *Pearson's correlation* condition, $r \leq -0.06$, whereas r is computed from both vectors that are the previous transformed template and the transformed template. Based on these cited operations, our objectives were met. The first is being to secure stored signature templates against attackers. The second is being to solve the problem of using one version of the signature for many applications (more than one database) without cross matching. The third achievement concerns cases where the signature template is compromised. Using the proposed algorithm, the administrator can perform the same action as in password authentication where the template is renewed by doing a transformation with a different key.

For future research, BioTrigono transformation could be implemented as a template protection technique for other biometric modalities such as the facial recognition system to guarantee human privacy.

7. ACKNOWLEDGE

This work was supported by Ministry of Higher Education of Malaysia which was made possible through the grant of Exploratory Research Grant Scheme (ERGS).

8. REFERENCES

[1] L. O'Gorman, "Comparing passwords, tokens, and biometrics for user authentication", in *Proc.* IEEE, Vol. 91, No. 12, Dec. 2003, pp. 2021–2040.

[2] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems", IBM Syst. J., vol. 40, no. 3, pp. 614–634, 2001.

[3] Y. Sutcu, H. T. Sencar and N. Memon, "A Secure biometric authentication scheme based on robust hashing", in *Proc*. 7th workshop on Multimedia and security (MM-Sec'05), New York, USA, Aug. 2005, pp. 111-116.

[4] L.Beaugé and A.Drygajlo, "Fully featured secure biometric smart card device for fingerprint-based authentication and identification", in *Proc*. 12th ACM workshop on Multimedia and security(MM&Sec'10), Roma, Italy, Sept. 2010, pp. 181-186.

[5] P. Briggs and P. Olivier, "Biometric daemons: authentication via electronic pets", in *Proc.* CHI, Florence, Italy, 2008.

[6] S.M.S. Ahmad, B. M. Ali and W.A.W. Adnan ,"Technical issues and challenges of biometric applications as access control tools of information security", International journal of innovative computing, information and control, vol. 8, vo. 11, pp.7983-7999 Nov. 2012.

[7] M. Faundez-Zanuy, "Signature recognition state-of-the-art", IEEE aerospace and electronic systems magazine , vol. 20, no.7, pp. 28–32, Jul. 2005.

[8] S. G. Kanade, D. Petrovska-Delacr´ etaz, and B. Dorizzi, "Cancelable biometrics for better security and privacy in biometric systems", Springer-Verlag Berlin Heidelberg, Part III, CCIS 192, pp. 20–34, 2011.

[9] E. Maiorana, "Biometric template protection for signature based authentication systems", PhD dissertation, University Roma Tre, Rome, Italy, 2009.

[10] A. K. Jain, K. Nandakumaand A. Nagar, "Biometric template security", EURASIP J. Adv. Signal Process," vol. 2008, no. 1, pp. 1-17, 2008.

[11] C. Vielhauer, R. Steinmetza and A. Mayerhofer, "Biometric hash based on statistical features of online signatures", in *Proc.* International conference on pattern recognition (ICPR), vol. 1, 2002. pp.123–126.

[12] H. Feng, C. C. Wah, "Private key generation from on-line handwritten signatures", Information management and computer security, vol.10, no.4, pp: 159–164, 2002.

[13] M. Freire-Santos, J. Fierrez-Aguilara and J. Ortega-Garcia, "Cryptographic key generation using handwritten

signature", in *Proc.* Biometric technologies for human identification III, Orlando, FL, vol. 6202, Apr. 2006, pp. 225–231.

[14] W. J. Scheirer and T. E. Boult, "Cracking fuzzy vault and biometric encryption", in *Proc*. IEEE Biometric Symp., 2007, pp. 1–6.

[15] W.K. Yip, A. Goh, D.C.L. Ngo, and A.B.J. Teoh, "Generation of Replaceable Cryptographic Keys from Dynamic Handwritten Signatures", International Conference in Biometrics (ICB06), Hong Kong, China, January 5-7, 2006, pp. 509–515.

[16] M. R. Freire, J. Fierrez and J. Ortega-Garcia, "Dynamic signature verification with template protection using helper data", in *Proc.* IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2008, pp.1713-1716.

[17] M. Purser, "Introduction to error correcting codes", Artech House, Boston,MA,1995.

[18] E. Maiorana and P. Campisi, "Fuzzy commitment for function based signature template protection", IEEE signal process. Lett., vol. 17, no. 3, pp. 249–252, Mar. 2010.

[19] A. Juels and M. Wattenberg, "A fuzzy commitment scheme", in *Proc.* ACM Conf. computer and communications security (ACM CCS '99), Singapore, Nov. 1999. pp. 28–36.

[20] E. Maiorana, P. Campisi, J. Fierrez, J. Ortega-Garcia and A.Neri, "Cancelable templates for sequence-based biometrics with application to on-line signature Recognition", IEEE Transaction on system, man and cybernetics-part A: system and human, vol. 40, no.3, pp. 525–538, May. 2010.

[21] E. Maiorana, M. Martinez-Diaz, P. Campisi, J. Ortega-Garcia and A. Neri, "Template Protection for HMM-based On-line Signature Authentication", in *Proc.* IEEE conference on computer vision and pattern recognition workshops (CVPRW), Anchorage, USA, June 2008, pp.1-6.

[22] E. Maiorana, P. Campisi, J. Ortega-Garcia, and A. Neri, "Cancelable biometrics for HMM-based signature recognition", in *Proc*. 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, (BTAS) Washington, DC, 2008. pp. 1–6.

[23] A. Cichocki and S. Amari, "Adaptive blind signal and image processing", Wiley, New York, 2002.

[24] E. A. Rúa, E. Maiorana, J. L. A. Castro and P. Campisi, "Biometric template protection using universal background mode: an application to online signature", IEEE transaction on information forensics and security, vol. 7, no. 1, pp. 269- 282, Feb, 2012.

[25] M. Martinez-Diaz, J. Fierrez, J. Ortega-Garcia ,"Universal Background Models for Dynamic Signature Verification", in *Proc.* IEEE, BTA, Crystal City, USA, Sept. 2007 pp.1 – 6.

[26] E. Maiorana, D. Blasi, P. Campisi, "Biometric template protection using turbo codes and modulation constellations", in *Proc.* information forensics and security (WIFS), Tenerife, Spain, Dec. 2012, pp. 25-30.

[27] F. L. Malallah, S. M. Syed Ahmad, S. Yussof, W. A. Wan Adnan, V. Iranmanesh and O. A. Arigbabu, "A Review of Biometric Template Protection Techniques for Online Handwritten Signature Application", International Review on Computers and Software (I.RE.CO.S.), Vol. 8, n. 12, Dec. 2013.

[28] F. L. Malallah, S. M. S. Ahmad, W. A.W. Adnan, V. Iranmanesh, S. Yussof, "Online signature template protection by shuffling and one time pad schemes with neural network verification", in *Proc.* 2nd International conference on computer science & computational mathematics (ICCSCM). Kuala Lumpur, Malaysia, 2013, pp. 43-50.

[29] J. Wen, M. Severa, W. Zeng M. H. Luttrell and W. Jin, "A Format-compliant configurable encryption framework for access control of video", IEEE Transaction on circuit and system for video technology, vol. 12, no. 6, pp. 545-557, June, 2002.

[30] S. Kanade, D.Petrovska-Delacrétaz, and B.Dorizzi,"Multi-biometrics based crypto-biometric session key generation and sharing protocol", in *Proc.*13th ACM multimedia workshop on Multimedia and security (MM&Sec '11), New York, NY, USA 2011, Pp.109-114.

[31] R.M. Bolle, J. H. Connell and N. K. Ratha, "Biometric perils and patches", Pattern recognition, vol. 35, no.12, pp.2727 – 2738, 2002.

[32] Davey Alba , "China's tianhe-2 caps top 10 supercomputers", IEEE Spectrum, Retrieved June 19, 2013, (Accessed 27 August, 2013).

[33] J. Fortuna and D. Capson, " Improved support vector classification using PCA and ICA feature space modification", Pattern Recognit., vol.37, no.4, pp.1117 – 1129, 2004.

[34] C. M. Bishop, "Pattern recognition and machine learning", in Information Science and Statistics, Springer, 2006.pp.225-233.

[35] J. Mohamad-Saleh and B. S. Hoyle ,"Improved neural network performance using principal component analysis on Matlab", International journal of the computer, the internet and Management vol.16, no.2, pp.1-8, May, 2008.

[36] H. Demuth, M. Beale and M. Hagan ," Neural Network Toolbox™ 6 User's Guide", by The MathWorks, Inc., 2009.

[37] H. Krishna, J. An and L. Zheng ,"A Neural network approach to classify inversion regions of high mobility ultralong channel single walled carbon nanotube field-effect transistors for sensing applications," in *Proc.* IEEE 5th International nanoelectronics conference (INEC), 2013, pp.85-88.

[38] C. Pratola, F. D. Frate, G. Schiavon and D. Solimini ,"Toward fully automatic detection of changes in suburban areas from VHR SAR images by combining multiple neural-network models", IEEE Transaction on geosciences and remote sensing, vol.51, no. 4, pp.2055-2066, April, 2013.

[39] N. Kamaruddin and A. Wahab ," Emulating human cognitive approach for speech emotion using MLP and GenSofNN", in *Proc.* IEEE 5th International conference

on information and communication technology for the Muslim world (ICT4M), Rabat, 2013, pp.1–5.

[40] M. R. G. Meireles, P. E. M. Almeida and M. G. Simõe ," A comprehensive review for industrial applicability of artificial neural networks", IEEE Trans Ind Electron, vol. 50, no. 3, pp.585-601, Jun. 2003.

[41] M. F. Moller , "A scaled conjugate gradient algorithm for fast supervised learning", Neural networks, vol.6, no.4, pp.525-533, 1993.

[42] T. Munakata, "Fundamentals of the new artificial intelligence—beyond traditional paradigms", Berlin, Germany: Springer-Verlag, 1998.

[43] M. Adya and and F. Collopy ,"How effective are neural networks at forecasting and prediction a review and evaluation", Journal of Forecasting, vol. 17, issues 5-6, pp.481-495,Dec.1998.

[44] S.M. Syed Ahmad, A. Shakil, A.R. Ahmad, M.A. M. Balbed and R. Md. Anwar, " SIGMA – A Malaysian signature's database", in *Proc.* AICCSA 2008, IEEE/ACS International conference on computer systems and applications, Doha, Qatar, 31 March – 4 Mar.2008, pp.919-920.

[45] V. Iranmanesh, S.M. Syed Ahmad, Adnan W. A. W. and F.L. Malallah, "Online signature verification using neural network and pearson correlation features", IEEE Conference on Open Systems (ICOS), Kuching, Malaysia, 2013, pp. 18–21.

[46] J. Wang and N. Zheng ," A novel fractal image compression scheme with block classification and sorting based on pearson's correlation coefficient ", IEEE Transaction on image processing, vol. 22, no.9, pp. 3690-3702, Sept. 2013.