# On the Dissimilarity Representation and Prototype Selection for Signature-Based Bio-Cryptographic Systems

George S. Eskander, Robert Sabourin, and Eric Granger

Laboratoire d'imagerie, de vision et d'intelligence artificielle, Ecole de technologie supérieure, Université du Québec,1100 rue Notre-Dame Ouest, Room A-3600, Montréal, QC, H3C 1K3, Canada
geskander@livia.etsmtl.ca,
robert.sabourin@etsmtl.ca,eric.granger@etsmtl.ca

**Abstract.** Robust bio-cryptographic schemes employ encoding methods where a short message is extracted from biometric samples to encode cryptographic keys. This approach implies design limitations: 1) the encoding message should be concise and discriminative, and 2) a dissimilarity threshold must provide a good compromise between false rejection and acceptance rates. In this paper, the dissimilarity representation approach is employed to tackle these limitations, with the offline signature images are employed as biometrics. The signature images are represented as vectors in a high dimensional feature space, and is projected on an intermediate space, where pairwise feature distances are computed. Boosting feature selection is employed to provide a compact space where intra-personal distances are minimized and the inter-personal distances are maximized. Finally, the resulting representation is projected on the dissimilarity space to select the most discriminative prototypes for encoding, and to optimize the dissimilarity threshold. Simulation results on the Brazilian signature DB show the viability of the proposed approach. Employing the dissimilarity representation approach increases the encoding message discriminative power (the area under the ROC curve grows by about 47%). Prototype selection with threshold optimization increases the decoding accuracy (the Average Error Rate AER grows by about 34%).

**Keywords:** Dissimilarity-representation, Prototype selection, Bio-Cryptography, Offline signatures.

## 1 Introduction

Bio-cryptographic systems are introduced to replace the traditional usage of simple user passwords by biometric traits like fingerprint, iris, face, signatures, etc., to secure the cryptographic keys within security schemes like encryption and digital signatures [1]. Different than the simple passwords, biometrics provide a more trusted authentication tool. However, their fuzzy nature harden the

classification decision. Similarities between inter-personal traits result in false acceptance and dissimilarities between intra-personal traits result in false rejections.

Robust bio-cryptographic systems operate in the key-binding mode where classical crypto-keys are coupled with the biometric message. For key binding, some encoding schemes like Fuzzy Commitment [2] and Fuzzy Vault (FV) [3] are the most commonly employed. In the enrollment phase, a prototype biometric message encodes the secret key. In the authentication phase, a message is extracted from the query sample to decode the key. The idea behind these schemes is to consider the query biometric message as a noisy version of the encoded message. If the query sample is genuine, the dissimilarity between the encoding and decoding messages is limited, so this noise can be eliminated by the decoder. On the other hand, if the query sample belongs to another person, or if it is a forged sample, the dissimilarity between the two messages is too high to cancel. Accordingly, the secret key will be unlocked only to users who apply similar enough query samples.

Some error correction codes like R-S codes [4] are employed to realize the key binding approach. Practical decoding complexity of such codes need that employed biometric messages should be concise. Also, error correction capacity of such codes can be controlled by adjusting a dissimilarity threshold. The decoder succeeds to unlock the secret, only if the dissimilarity between the prototype and the query message is beyond the threshold. Accordingly, this threshold should be properly adjusted based on the expected dissimilarity ranges. So that, the code can cancel the intra-personal dissimilarities and fails to cancel the inter-personal dissimilarities.

For physiological biometrics like fingerprint and iris, small number of simple features extracted in the spacial domain can be employed to constitute informative encoding messages. This is simply because the intrinsic stability and discriminative nature of such biometrics. On the other hand, for behavioral biometrics like offline signature images, the intra-personal variability and inter-personal similarity are intrinsic properties. Moreover, it is easy to produce forged signature images. Accordingly, discrimination between genuine and forged signatures needs high dimensional feature representation and complicated classifiers [5]. It is a challenging task to produce a concise and informative messages from the signature images, and to use simple classifiers like the bio-cryptographic decoders to differentiate between genuine and forged signatures.

In this paper, design of reliable decoders for offline signature-based bio-cryptography is tackled by employing the concept of dissimilarity-representation [6]. This concept is originally introduced to build classical classifiers, by replacing the feature representation of objects by their dissimilarity to a fixed set of prototypes. Performance of these classifiers relies on the accuracy of the employed dissimilarity measure and how carefully the prototypes are chosen [8]. In literature, dissimilarity measures often composed of graphs, strings, or normalized versions of the raw measurements. However, the dissimilarity approach may also be used on top of a feature representation, where object proximity is

represented by computing the distance between ordinary feature representations in a vectorial space [7].

As most of work on classical offline signature verification is feature-based, where many techniques of feature extraction are already proposed [5], we base our method on top of a feature representation. In such case, the encoding messages are composed of a set of features. The dissimilarity between the prototype and query messages is measured by the distance between the feature vectors that constitute these messages. The rational behind the proposed method is that the overall dissimilarity between two messages is an accumulation of individual dissimilarities between every pair of corresponding elements of the message. So, to increase the separation between the intra-personal and inter-personal dissimilarity ranges, we select features that decrease the intra-personal distances and that increase the inter-personal distances.

The enrolling signature images are first represented as vectors in a high dimensional feature space. This representation is projected on an intermediate space, which we call a "feature-dissimilarity" space, where pairwise feature distances are computed. Boosting feature selection is employed in this intermediate space, producing a compact space with the intra-personal distances are minimized and the inter-personal distances are maximized. Finally, the resulting representation is projected on the dissimilarity space to select the most discriminative prototypes for encoding, along with optimizing the dissimilarity threshold.

For proof of concept simulations, the Brazilian signature DB (including genuine and samples with different levels of forgeries) is employed [9]. The impact of proposed dissimilarity representation approach is investigated by analyzing the separation between the intra-personal and the inter-personal dissimilarity distributions. The benefit of prototype selection with optimizing the dissimilarity threshold is tested by its impact on the overall recognition accuracy.

The rest of this paper is organized as follows. The next section provides some background on the dissimilarity representations as applied to bio-cryptographic offline signature based systems. The proposed dissimilarity representation and prototype selection approach for designing signature-based bio-cryptographic systems is illustrated in section 3. The experimental methodology is illustrated in section 4. The experimental results are presented and discussed in section 5.

## 2 Background

Signature Verification systems (SV) are employed to authenticate individuals based on their handwritten signatures. Classical SV systems output a simple acceptance/rejection decision for a query signature sample. On the other hand, signature-based bio-cryptographic systems release a secret cryptographic key only for a user who applies a genuine signature sample. There are two modes of operation for signature-based systems: online and offline. For online systems, users use special devices like special pens and tablets to acquire their signature dynamics such as velocity, pressure, etc. On the other hand, offline signature-

based systems use scanned signature images for the recognition task. Only static information can be acquired from the signature images, producing less informative signals, and hence, a harder pattern recognition task.

Most of work done in the signature verification area applied feature-based pattern recognition approaches, where feature representations are constituted from signature signals. The classifiers are then designed in the feature space. Performance of such systems are basically limited by the quality of employed feature representations.

Handwritten signature images imply high variability between different user samples, and also high similarity between signatures of different users. Accordingly, the feature-based approach succeeds to produce offline SV verification systems, only when high dimensional feature representations and complex classifiers are employed. For a comprehensive review on the different approaches see [5].

For bio-cryptographic systems design, there are some restrictions on the size of the employed feature representations, and on the classification complexity. Accordingly, direct application of the feature-based approach produces inaccurate systems. In literature, few bio-cryptographic implementations are done based on the handwritten signatures. The online signatures produced bio-cryptographic systems with acceptable performance [14], as discriminative features like velocity, pressure, etc, are employed. On the other hand, it is shown that static features extracted from the offline signature images are unstable and they are not discriminant enough to design a bio-cryptographic system [15].

Different than the feature-based approach, the concept of dissimilarity-based classification has been proposed by Elzbieta Pekalska and Robert P.W. Duin., [6]. The rational behind this concept is that modeling the proximity between objects may be more discriminative than modeling the objects themselves. This is because objects belong to a specific class have a shared degree of commonality that could be captured by a dissimilarity value.

We propose that the dissimilarity-based approach can be employed to design reliable key-binding bio-cryptographic systems. In such systems, error correction-based decoders are used. If the dissimilarity between the decoding and the encoding signals is less than a specific threshold, the decoder succeeds to decouple the encoded bio-ctyptographic key. So, functionality of these decoders can be considered as two-class simple thresholding classifiers that operate in the dissimilarity space.

In literature, the concept of dissimilarity representation is not directly employed to design bio-cryptographic systems. However, some authors proposed methodologies to absorb the dissimilarities between encoding and decoding biometric signals, so that they are within the error correction capacity of the decoder. For instance, Fingerprint-based fuzzy vaults are designed by using some minutia points extracted in the spatial space to constitute the encoding message [16]. The dissimilarity between encoding and decoding messages is decreased by aligning the query and the template fingerprints prior to the decoding process. For our proposed method, instead of aligning the dissimilar messages, we design

them in a way that produces similar intra-personal messages and dissimilar inter-personal encoding messages. A preliminary realization of the proposed method is appeared in [17], where a Fuzzy Vault (FV) system based on the offline signature images is proposed. Boosting feature selection (BFS) is employed to select informative representation, so that intra-personal dissimilarities are minimized and inter-personal dissimilarities are maximized. Although produced discriminative representations, this method did not cancel some of the intrinsic fuzziness of the signature signals.

In this paper, we extend the method in [17], so that some of the residual fuzziness of the signature representations is canceled. Inspired by fingerprint alignment technique proposed by Nandakumar et al., [16], we model the representation dissimilarities, and use this information to absorb the residual message fuzziness before sending it to the bio-cryptographic decoders. Moreover, as quality of representation relies mainly on the quality of employed reference signatures (few work is done on selecting a reference subset for classical signature verification systems, e.g., [10].), we extend this idea to the bio-cryptography domain. The designed messages are projected to the dissimilarity space, where each dimension is the message distance to a prototype message. In this space, the most discriminative prototypes are selected, along with optimizing the dissimilarity threshold.

## 3 Proposed Dissimilarity Representation and Prototype Selection Method

Assume an encoding biometric message: $E^p = \{f_i^p\}_{i=1}^t$, where $p$ is the signature prototype used for message extraction, $f_i^p$ is a feature extracted from $p$ to constitute a message element, and $t$ is the message length. In the enrollment phase, $E^p$ is extracted and used to encode a secret cryptographic key $K$. In the authentication time, a decoding query message $E^Q = \{f_i^Q\}_{i=1}^t$ is extracted, where $Q$ is the query signature sample applied to decode the locked key $K$[1]. Assume the dissimilarity between the two messages is $D^{Qp}$. For error correction decoders like the R-S decoders [4], the decoder succeeds to cancel the dissimilarity between $Q$ and $p$, if the dissimilarity (error) $D^{Qp}$ is less than its error correction capacity $\Theta$. Hence, decoder functionality $DF$ can be formulated as follows:

$$DF = \begin{cases} 1 & if \quad D^{Qp} \leq \Theta \\ 0 & if \quad D^{Qp} > \Theta \end{cases} \tag{1}$$

where $\Theta$ is the error correction capacity of the decoder (dissimilarity threshold). Hence, to achieve perfect decoding accuracy, the following condition should be satisfied:

[1] Details of how the crypto-key is encoded/decoded by means of a biometric message is out of the scope of this paper. For more details on this aspect see [3], and [2]

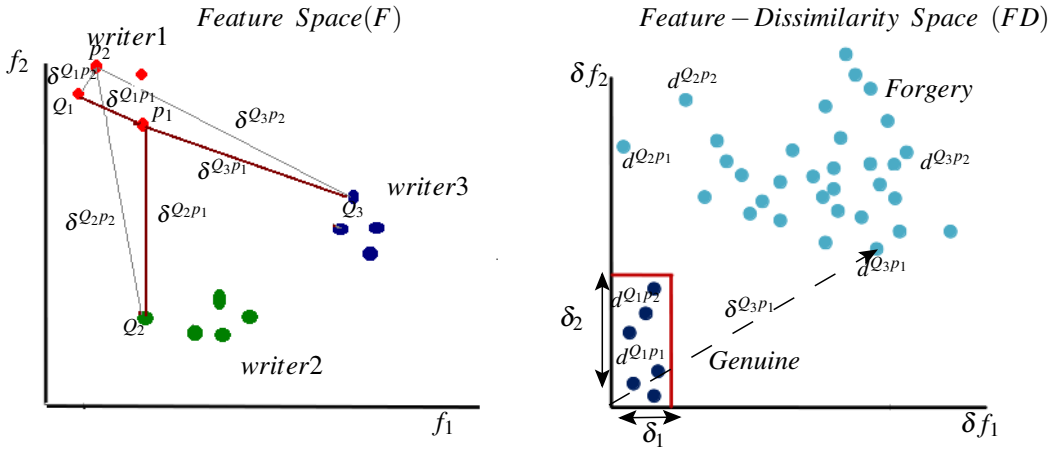


**Fig. 1.** Illustration of feature selection in the original feature space (left) and in the feature-dissimilarity space (right).

$$D^{Qp} \begin{cases} \leq \Theta & if\ Q\ is\ a\ genuine\ sample \\ > \Theta & if\ Q\ is\ a\ forgery\ sample \end{cases} \tag{2}$$

Satisfying the above condition relies on the following design issues:

1. selection of the message elements $\{f_i\}_{i=1}^{t}$.
2. the dissimilarity measure employed to produce the dissimilarity score $D^{Qp}$.
3. selection of the signature prototype $p$ for encoding.
4. the correction capacity of the decoder (dissimilarity threshold) $\Theta$.

In this paper we propose a methodology to optimize these design issues, so high decoding accuracy is achieved. The proposed method consists of two main stages: 1) design of the encoding messages and the dissimilarity measure, and 2) prototype selection and dissimilarity threshold optimization.

### 3.1 Design of the Encoding Messages and the dissimilarity Measure

For a message of length $t$, consider Euclidean distance $\delta^{Q_j p_r}$ between the query message $Q_j$ and the prototype $p_r$:

$$\delta^{Q_j p_r} = \sqrt{\sum_{i=1}^{t} (\delta f_i^{Q_j p_r})^2} \tag{3}$$

where $\delta f_i^{Q_j p_r} = \| f_i^{Q_j} - f_i^{p_r} \|$.

Hence, the overall dissimilarity between messages is an accumulation of the individual dissimilarities between every two corresponding elements of the message. So, to increase the separation between the intra-personal and inter-personal dissimilarity ranges, we select features that decrease the intra-personal distances and that increase the inter-personal distances.

The enrolling signature images are first represented as vectors in a high dimensional feature space $F$. This representation is projected on an intermediate space, which we call a "feature-dissimilarity" space $FD$, where pairwise feature distances are computed. Figure 1 illustrates the transformation from space $F$ to space $FD$. In the left side, signatures of three writers are represented in $F$. For simplicity, only two features $f_1$ and $f_2$ are shown in this figure, while typical representations might have high dimensionality. In this example, we assume that writer 1 is the only authentic person, whose signatures should succeed to decode the cryptographic key $K$. Two signatures are considered as prototypes for this user, $p_1$ and $p_2$. Euclidean distance is employed as a dissimilarity measure. It is clear that a dissimilarity representation that is built on top of this feature representation is discriminative. Distances among intra-personal signatures (like $\delta^{Q_1 p_1}$) are generally smaller than the distances among inter-personal signatures (like $\delta^{Q_2 p_1}$). However, in this space it is not clear which feature is more discriminative. With representations of high dimensionality, high number of system users, unknown forgeries and a small number of training samples, it is not feasible to select the most discriminative features in the feature space $F$.

Accordingly, we project this representation on a feature-dissimilarity space $FD$, as shown in the right side of Figure 1. In this space, distance between each corresponding features, for each pair of signatures, is computed and used as new set of features $\{\delta f_i\}_{i=1}^{t}$. So, dimensionality of the $F$ and $FD$ spaces is equal. A distance $\delta^{Q_j p_r}$ between a query $Q_j$ and a prototype $p_r$ is mapped from $F$ to $FD$ as a point $d^{Q_j p_r}$:

$$d^{Q_j p_r} = \{\delta f_i^{Q_j p_r}\}_{i=1}^{t} \tag{4}$$

where, $\delta^{Q_j p_r}$ is represented by the distance from the origin point to $d^{Q_j p_r}$. Here, the impact of every individual feature on the signature dissimilarities is clear. It is obvious that $f_2$ is more discriminative than $f_1$. For all genuine query samples like $Q_1$, $\delta f_2^{Q_1 p_r} < \delta_2$ and for all forgery query samples like $Q_2$ and $Q_3$, $\delta f_2^{Q_j p_r} > \delta_2$. On the other hand, $f_1$ is less discriminant. For the forgery query $Q_2$, $\delta f_1^{Q_2 p_1} < \delta_1$, same as that for the genuine sample $Q_1$. Accordingly, it is easier to rank and select features in the $FD$ space, as the impact of the individual features on the overall dissimilarity is clear in this space. Moreover, the multi-class problem with few training samples per class in $F$ space is transformed to a two-class problem in $FD$ space, with more training samples per class.

Ranking and selecting the most discriminant features in the $FD$ space, produces encoding/decoding messages with low dissimilarities between intra-personal instances and with high dissimilarities between inter-personal instances.

However, some of the intrinsic fuzziness of the signature signal will not be canceled through this feature selection approach. To alleviate that, we propose an adaptive distance measure that is computed in the $DF$ space, and absorbs some of the residual fuzziness. For a feature representation $F = \{f_i\}_{i=1}^{t}$, the feature dissimilarity vector $\Delta = \{\delta_i\}_{i=1}^{t}$ is learnt in $FD$ space, where $\delta_i$ discriminates between the intra-personal and the inter-personal dissimilarities for a feature $f_i$. Based on this modeled dissimilarity, we replace the Euclidean distance measure ($\delta^{Q_j p_r}$) by an adaptive dissimilarity measure:

$$D^{Q_j p_r} = \sum_{i=1}^{t} (D_i^{Q_j p_r}), \quad where \ D_i^{Q_j p_r} = \begin{cases} 0 & if \ (\delta f_i^{Q_j p_r} < \delta_i) \\ 1 & otherwise \end{cases} \tag{5}$$

Employing this adaptive distance measure absorbs some of the intrinsic feature variability and increases its discriminative power. For instance, according to Eq.5, distances among the genuine query and its prototypes $D^{Q_1 p_r} = 0$. Moreover, most of the distances between the unauthorized queries and the genuine prototypes $D^{Q_j p_r} = 2$, $\vee j \in [2, 3]$. Hence, some of the variability of the dissimilarity values is canceled.

Ranking the features $\{f_i\}_{i=1}^{t}$ and learning the dissimilarity vector $\{\delta_i\}_{i=1}^{t}$ in the $FD$ space is a general approach, that can be achieved by employing different feature selection methods. However in this paper, this concept is realized by employing a two-step boosting feature selection (BFS) method [12], for fast searching in high dimensional spaces. Decision-stumps ($DS$) [19], that are single-split single-level classification trees, are trained through a boosting process [18]. Training of a $DS$ is equivalent to selection of a single feature that discriminats between two classes based on a splitting threshold. If the BFS runs in the $FD$ space, a $DS_i$ at a learning iteration $i$, locates the best dissimilarity feature $\delta f_i$, that splits the two classes around a splitting dissimilarity threshold $\delta_i$.

In the first step, a development database ($DevDB$) containing samples of simulated users, is used for training. The reason is that the signature samples of real users are not enough for feature selection in high dimensional spaces. Then, population-based representation is produced by running a BFS process in a $DF$ space, generated by multi-feature representations extracted from the $DevDB$ database. This approach is employed by Rivard et al., to design a writer-independent (WI) classical offline signature verification system [11]. However, the produced population-based spaces have high dimensionality. This is not suitable for encoding bio-cryptographic systems, as the encoding/decoding messages should be concise.

In the second step, the exploitation database ($ExpDB$), containing samples of the real users, is used for training. Signature samples are represented in the population-based space defined through the first step, and additional BFS process runs in this user-based space. Recently, we employed this approach to adapt WI systems to specific writers [13]. Reliable writer-dependent (WD) systems are achieved based on concise and discriminative user-based feature spaces. In this

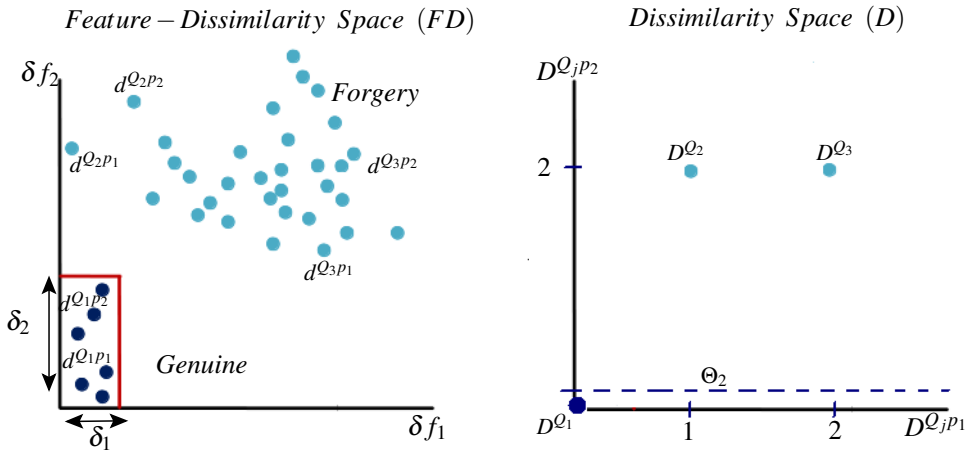


**Fig. 2.** Illustration of the transformation from the feature-dissimilarity space (left) to the dissimilarity space (right).

paper, a similar two-step BFS process is employed, however, the user-based BFS step is employed in a $FD$ space, in order to model the feature dissimilarity vector $\Delta = \{\delta_i\}_{i=1}^{t}$.

### 3.2 Prototype Selection and Dissimilarity Threshold Optimization

The aforementioned approach enlarges the separation between the dissimilarity distributions of the genuine and impostor encoding messages. However, the distributions differ based on the prototype used for the dissimilarity computations (Eq.5). To get the best possible dissimilarity representation, we propose a prototype selection method.

To this end, the user-based representation, produced through the two-step BFS process, is projected from the $FD$ space to a dissimilarity space $D$. Consider the available set of $R$ prototypes $P = \{p_1, p_2, ..., p_R\}$. The adaptive dissimilarity distance for a query $Q_j$ is computed for every prototype $p_r \in P$, according to Eq.5. This operation produces a dissimilarity vector $D^{Q_j}$ in the dissimilarity space, where

$$D^{Q_j} = \{D^{Q_j p_1}, D^{Q_j p_2}, ..., D^{Q_j p_R}\}. \tag{6}$$

Figure 2 illustrates the transformation between the $FD$ and $D$ spaces. In the left side, distances between prototype and query messages are represented in the $FD$ space. It is obvious that different prototypes produce different distance values, where significant variability exists for the genuine and the forgery classes.

Also, in this space, it is not clear which prototype is the most informative. For space $D$ shown in the right figure, it is obvious that some variability is absorbed through employing the adaptive dissimilarity measure. For instance, $D^{Q_j} = \{0, 0\}$ for all genuine queries (see Eq.5 and Eq.6), as feature dissimilarities $\delta f_1^{Q_j p_r} < \delta_1$ and $\delta f_2^{Q_j p_r} < \delta_2$, for the genuine queries. Also, for most of the forgery queries, $D^{Q_j} = \{2, 2\}$, as $\delta f_1^{Q_j p_r} > \delta_1$ and $\delta f_2^{Q_j p_r} > \delta_2$ for the forgery queries.

Moreover, the dissimilarity space representation provides easier way to rank prototypes according to their discriminative power. For instance, $p_2$ is more discriminative than $p_1$, as for all forgery queries, $D^{Q_j p_2} = 2$. While for $Q_2$, $D^{Q_2 p_1} = 1$ (as $\delta f_1^{Q_2 p_1} < \delta_1$). So, measuring the dissimilarity relative to $p_2$ results in more isolated clusters.

Finally, in the $D$ space, we optimize the dissimilarity threshold ($\Theta$). In the illustrated example, if the selected prototype is $p_2$, then any $\Theta_2 < 2$ is discriminant. For $p_1$, any $\Theta_1 < 1$ is discriminant. Selection of prototypes with higher margin between clusters, provides wider range for selecting the dissimilarity threshold $\Theta$. This results in more flexibility for parameter setting of the bio-cryptographic decoder and hence, higher security and recognition accuracy can be achieved [16].

Based on the proposed method, the decoding functionality $DF$ formulated by Eq.1 can be reformulated as:

$$DF_r(Q_j) = sign(\Theta_r - D^{Q_j p_r}). \tag{7}$$

where $r$ is the index of the selected prototype $p_r$, $Q_j$ is the query encoding message, $\Theta_r$ is the dissimilarity threshold associated with this prototype, and $D^{Q_j p_r}$ is the dissimilarity value computed according to Eq. 5.

The prototype selection method can be realized by various feature selection techniques (with considering prototypes as features), however, we realized it through employing the BFS approach [12].

## 4 Experimental Methodology

### 4.1 Database

The Brazilian database [9] is used for proof-of-concept simulations. It contains 7,920 samples of signatures that were digitized as 8-bit grayscale images over 400X1000 pixels at resolution of 300 dpi. This DB contains three types of signature forgery: random, simple and simulated. Random forgeries do not know neither the signerś name nor the signature morphology. It can also happen when a genuine signature presented to the system is mislabeled to another user. For simple forgery, the forger knows the writerś name but not the signature morphology. He can only produce a simple forgery using a style of writing of his

liking. Simulated forgeries have access to a sample of the signature. A forger can therefore imitate the genuine signature.

The signatures were provided by 168 writers and are organized as follows: the first 60 writers have 40 genuine signatures, 10 simple forgeries and 10 simulated forgeries per writer, and the other 108 have only 40 genuine signatures per writer. The experimental database is split into two sets: a development dataset ($DevDB$) composed of the last 108 writers, and an exploitation dataset ($ExpDB$) composed of the first 60 writers. Set $DevDB$ is used for the population-based BFS step as illustrated in Section. 3.1.

Set $ExpDB$ is split into two subsets: the reference subset ($R$) contains the first 30 genuine signatures, and the query subset ($Q$) contains the rest 10 genuine samples, 10 simple and 10 simulated forgeries. The subset $R$ is used for the user-based BFS step as illustrated in Section. 3.1, and for the prototype selection and dissimilarity threshold optimization as illustrated in Section. 3.2. Both subsets of $ExpDB$ are used for evaluating the method performance.

### 4.2 Feature Extraction

Extended-Shadow-Code (ESC) [20], and Directional Probability Density Function (DPDF) [21] are employed. Features are extracted based on different grid scales, hence a range of details are detected in the signature image. A set of 30 grid scales is used for each feature type, producing 60 different single scale feature representations. These representations are then fused to produce a feature representation of huge dimensionality $(30, 201)$ [11].

### 4.3 Design of Encoding Messages and Dissimilarity Measure

The two-step BFS process is implemented as illustrated in section 3.1. First, the ($DevDB$) is used for the population-based BFS phase. We followed the same experimental settings as in the system in [11]. This phase produced a population-based representation ($PR$) of dimensionality $L = 555$. Second, the reference subset ($R$) is used for the user-based BFS phase. For each user in $ExpDB$, the signatures in $R$ are used to represent the genuine class, and some signatures from the $DevDB$ are used to represent the forgery class. Then, signatures of both classes are represented in the $PR$ space of $L$ dimensionality. This representation is then transformed to the $FD$ space, where the user-based BFS step runs for $t$ boosting iterations. The process outputs the message elements $\{f_i\}_{i=1}^{20}$, along with their dissimilarities $\Delta = \{\delta_i\}_{i=1}^{20}$, that are used for computing the adaptive dissimilarity measure defined by Eq. 5.

### 4.4 Prototype Selection and Dissimilarity Threshold Optimization

The thirty signatures in the reference subset ($R$) are used as a prototype set $P = \{p_r\}_{r=1}^{30}$. To constitute the dissimilarity space $D$, the adaptive dissimilarity value is computed for every signature in $R$ against all of the thirty signatures (Eq.6).

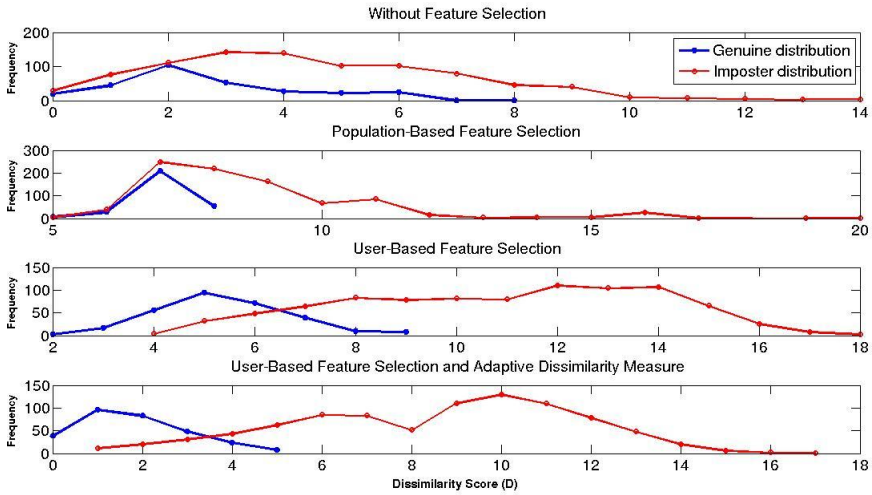


**Fig. 3.** Dissimilarity score distribution for a specific user.

To constitute the forgery class, samples from $DevDB$ are chosen randomly, and dissimilarities between them and the prototypes are computed. BFS runs in this dissimilarity space, to select the best prototype of $p_r$ with the associated threshold $\Theta_r$.

### 4.5 Performance Measures

To assess the impact of the proposed dissimilarity representation approach on the separability of the genuine and impostor clusters, we use the Hellinger distance. Assuming normal distributions G and I for the genuine and impostor classes, respectively. the squared Hellinger distance between them is give by:

$$H^2(G,I) = 1 - \sqrt{\frac{2\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2}}\, e^{-\frac{1}{4}\frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}}. \tag{8}$$

where, $\mu_1$, $\mu_2$ and $\sigma_1$, $\sigma_2$ are the mean and variance values for $G$ and $I$, respectively.

To measure the clusters separability for the different types of forgeries, we report $H_{random}$, $H_{simple}$ and $H_{simulated}$, where the parameters $\mu$ and $\sigma$ of the impostor cluster $I$ are computed each time, based on the dissimilarities against samples of a specific type of forgeries. Also, we report $H_{all}$, where the distribution parameters are computed according to dissimilarities of all forgery types.

Also, as the recognition accuracy of bio-cryptographic decoders relies on the dissimilarity ranges separability and on the employed dissimilarity threshold, we measure the recognition errors for all of the dissimilarity scores and use them to generate ROC curves. A ROC curve plots the False Accept Rate (FAR) against the Genuine Accept Rate (GAR) for all possible thresholds (all generated

dissimilarity scores). FAR for a specific threshold is the ratio of forgery samples with a dissimilarity score smaller than this threshold. GAR is the ratio of genuine samples with a dissimilarity score smaller than the threshold.

In order to have a global assessment on the quality of encoding messages representation, we compute and average the area under the ROC curves (AUC), for all users in the $ExpDB$ subset. High AUC indicates more separation between the dissimilarity score distributions for the genuine and impostor classes.

To assess the impact of the prototype and threshold selection step, we compute the recognition rates. Decoder outputs are estimated by employing Eq. 7 for the selected prototypes and thresholds. By comparing the decoder outputs to the actual class labels, we compute the average error rate ($AER_{all}$), where

$$AER_{all} = (FRR + FAR_{random} + FAR_{simple} + FAR_{simulated})/4 \qquad (9)$$

False Reject Rate ($FRR$) is the ratio of genuine queries that produce '0' decoding outputs, $FAR_{random}$, $FAR_{simple}$ and $FAR_{simulated}$ are the ratio of random, simple, and simulated forgeries respectively that produce '1' decoding outputs. The error rates are also computed when no prototype selection step is employed and for a fixed threshold $\Theta = 6$. [2]

## 5 Experimental Results

The power of the proposed method for designing the encoding messages and employing the adaptive dissimilarity measure is assessed by its impact on the separability of the genuine and impostor dissimilarity distributions. Figure 3 illustrates the impact of each step of the proposed method for a specific user of the $ExpDB$ dataset. It is obvious that, when no feature selection is employed to constitute the encoding message, the genuine and impostor distributions are overlapped. Running BFS based on population signature samples increases the separation between the two distribution. Running the user-based BFS step enhanced the separability. Employing the adaptive distance measure, increased the stability of the genuine class. For instance, the maximum dissimilarity score for the genuine class is decreased from 9 to 5. However, this impact differs for the different forgery types. For instance, in Figure 4, it is clear that while the random forgery class distribution is significantly separated, the simulated forgery distribution still has significant class overlap.

To asses the average performance of the proposed method, the average Hellinger distance is computed over the 60 Users, and for the different types of forgeries. Table 1 shows the results of this analysis. It is obvious that each processing step increased the distances between the genuine and impostor distributions, for

[2] $\Theta = 6$ is equivalent to encoding a crypto-key of $128 - bits$ by a biometric message of length $t = 20$, by implementing the FV key-binding scheme [3]. Also, for technical issues, the message elements $\{f_i\}_{i=1}^{t}$ are quantized in 8-bit words before computing the dissimilarities.

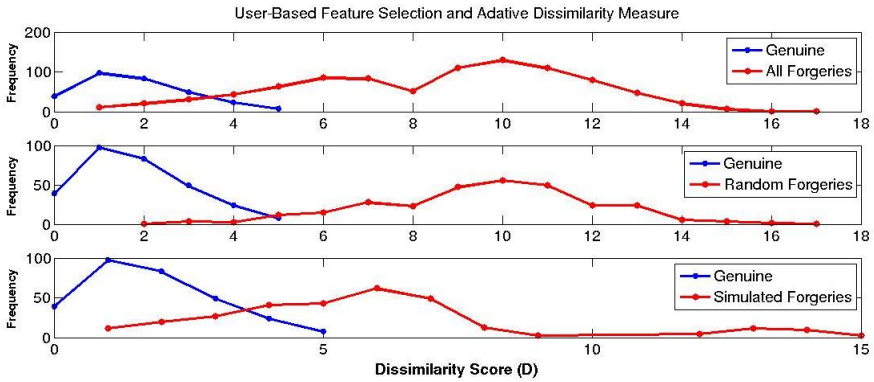


**Fig. 4.** Dissimilarity score distribution for different forgery types.

**Table 1.** Average Hellinger distance over all Users for the different design scenarios

| Design Aspect | Without Feature Selection | Population-based Feature Selection | User-based Feature Selection | User-based Feature Selection with Adaptive Distance Measure |
|---|---|---|---|---|
| Average $H_{random}$ | 0.2976 | 0.6093 | 0.6617 | 0.7398 |
| Average $H_{simple}$ | 0.2519 | 0.5531 | 0.6011 | 0.6951 |
| Average $H_{simulated}$ | 0.1466 | 0.4395 | 0.4786 | 0.5907 |
| Average $H_{all}$ | 0.2496 | 0.5590 | 0.5923 | 0.6617 |
| Average $AUC$ | 0.6577 | 0.7724 | 0.9328 | 0.9700 |

all types of forgeries. Average distance of the all forgeries distributions $H_{all}$ is increased from 0.2496 to 0.6617. Also, the average AUC is increased by about 47% (from 0.6577 to 0.9700).

The dissimilarity scores reported above are averaged for all prototypes in the subset $R$. However, class separation differs for the different prototypes. For instance, Figure 5 shows distributions of the best and worst prototypes for a specific user. For the worst prototype, a dissimilarity threshold $\Theta = 4$ results in $FRR = 10\%$, $FAR_{random} = 10\%$ and $FAR_{simulated} = 30\%$. For the best prototype, $FRR = 0\%$, $FAR_{random} = 0\%$ and $FAR_{simulated} = 20\%$.

The overall impact of running the prototype selection and threshold optimization step is investigated by computed the recognition error rates for both cases. Tabel 2 shows that $AER$ is decreased by about 34% (from 11.15% to 7.32%), through employing this selection step.

## 6 Conclusions and Future Work

In this paper, a methodology for designing bio-cryptographic systems based on the dissimilarity representation approach, is proposed. Separation between gen-

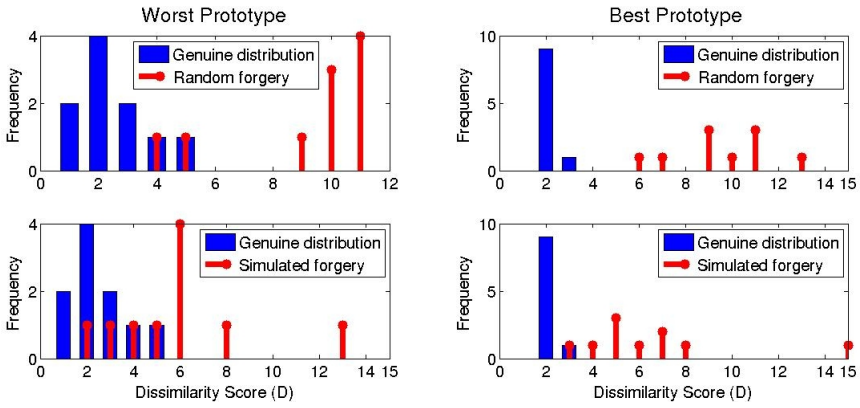


**Fig. 5.** Dissimilarity score distributions for different prototypes.

**Table 2.** Impact of the Prototype Selection on Average Error Rate over all Users

| Design Aspect | Without Prototype Selection | With Prototype Selection |
|---|---|---|
| Average $FRR$ | 5.25 | 4.83 |
| Average $FAR_{random}$ | 2.74 | 0.6 |
| Average $FAR_{simple}$ | 3.49 | 1.5 |
| Average $FAR_{simulated}$ | 33.14 | 22.33 |
| Average $AER$ | 11.15 | 7.32 |

uine and impostor distributions is increased through maximizing the distance between the individual elements of the encoding messages. Some of the intrinsic variability of the messages is absorbed by employing an adaptive dissimilarity measure. A prototype selection and dissimilarity threshold optimization method is proposed, to enhance the recognition performance. Future work will employ the proposed method to build a complete signature-based bio-cryptographic system.

## Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada and BancTec Inc.

## References

1. U. Uludag, S. Pankanti, S. Prabhakar and A.K. Jain., Biometric Cryptosystems: Issues and Challenges. *Proceedings of the IEEE*, vol.92, issue.6, pp.948-960, 2004.
2. A. Juels and M. Wattenberg., A Fuzzy Commitment scheme. *Sixth ACM Conference on Computer and Communications Security*, pp.28-36, ACM Press. 1999.

3. A. Juels and M. Sudan., A Fuzzy Vault scheme. *In proc. IEEE int. Symp. Inf. Theory*, Switzerland, pp.408, 2002.
4. Berlekamp and Elwyn R., Algebraic Coding Theory. *McGraw-Hill*, New York, NY, USA, 1968.m
5. D. Impedovo and G. Pirlo., Automatic signature verification: the state of the art. *IEEE Transactions on SMC, Part C: Applications and Reviews*, vol.38, no.5, pp.609-635,2008.
6. Elzbieta Pekalska , Robert P.W. Duin. Dissimilarity representations allow for building good classifiers. *PR Letters*, vol.23, no.8, pp.161-166, 2002.
7. Robert P.W. Duin , Marco Loog, Elzbieta Pekalska , and David M.J. Tax. Feature-Based Dissimilarity Space Classification. *Proceedings of the 20th International conference on Recognizing patterns in signals, speech, images, and videos (ICPR'10)*, pp.46-55, 2010.
8. Elzbieta pekalska, Robert P.W. Duin, Pavel Paclk Prototype selection for dissimilarity-based classifiers. *PR*, vol.39, pp.189208, 2006.
9. C. Freitas, M. Morita, L. Oliveira, E. Justino, A. Yacoubi, E. Lethelier, F. Bortolozzi, and R. Sabourin., Bases de dados de cheques bancarios brasileiros. *XXVI Conferencia Latinoamericana de Informatica*, Mxico, 2000.
10. Dimauro, G. ; Guerriero, A. ; Impedovo, S. ; Pirlo, G. ; Salzo, A. ; Sarcinella, L. Selection of reference signatures for automatic signature verification. *Proceedings of the Fifth International Conference on Document Analysis and Recognition (ICDAR '99.)*, pp.597-600, 1999.
11. Rivard, D, Granger, E and Sabourin, R., Multi-Feature extraction and selection in writer-independent offline signature verification. *IJDAR*, vol.16, no.1, pp.83-103, 2013.
12. K. Tieu and P. Viola., Boosting image retrieval. *International Journal of Computer Vision*, vol.56, no.1, pp.17-36, 2004.
13. Eskander, G.S., Sabourin, R. and Granger, E., Adaptation of writer-independent systems for offline signature verification. *The 13th International Conference on Frontiers in Handwriting Recognition (ICFHR-2012)*, pp.432-437, Bari, Italy, 2012.
14. M. Freire-Santos, J. Fierrez-Aguilar and J. Ortega-Garcia., Cryptographic key generation using handwritten signatures. *proc of SPIE*, vol.6202, pp.225-231, 2006.
15. Manuel Freire-Santos, J. Fierrez-Aguilar, M. Martinez-Diaz and J. Ortega-Garcia., On the applicability of off-line signatures to the Fuzzy Vault construction. *proc of ICDAR2007*, Curitiba, Brazil, 2007.
16. K. Nandakumar A. K. Jain and S. Pankanti., Fingerprint based Fuzzy Vault: Implementation and Performance. *IEEE Transactions on IFS*, vol.2, no.4, pp.744-757, 2007.
17. Eskander, G.S., Sabourin, R. and Granger, E., Signature based Fuzzy Vaults with boosted feature selection. *IEEE Workshop on Computational Intelligence and Identity Management (SSCI-CIBIM 2011)*, pp.131-138, Paris, 2011.
18. R. Schapire., The boosting approach to machine learning: An overview. *Proc. MSRI Workshop on Nonlinear Estimation and Classification*, 2002.
19. W. Iba and P. Langley., Induction of one-level decision trees. *Proc of the Ninth International Machine Learning Conference*, Scotland, pp. 233-240, 1992.
20. R. Sabourin and G. Genest., An Extended-Shadow-Code based Approach for Off-Line Signature Verification. *Proc of the 12th international conference on PR*, Jerusalem, vol.2, pp.450-453, 1994.
21. J. Drouhard, R. Sabourin and M. Godbout., A neural network approach to off-line signature verification using directional pdf. *PR*, vol.29, no.3, pp.415-424, 1996.