Cancelable Templates for Sequence-Based Biometrics with Application to On-line Signature Recognition

Emanuele Maiorana, *Member, IEEE*, Patrizio Campisi, *Senior Member, IEEE*, Julian Fierrez, Javier Ortega-Garcia, *Senior Member, IEEE*, and Alessandro Neri, *Member, IEEE*

Abstract—Recent years have seen the rapid spread of biometric technologies for automatic people recognition. However, security and privacy issues still represent the main obstacles for the deployment of biometric-based authentication systems. In this paper, we propose an approach, which we refer to as BioConvolving, that is able to guarantee security and renewability to biometric templates. Specifically, we introduce a set of noninvertible transformations, which can be applied to any biometrics whose template can be represented by a set of sequences, in order to generate multiple transformed versions of the template. Once the transformation is performed, retrieving the original data from the transformed template is computationally as hard as random guessing. As a proof of concept, the proposed approach is applied to an on-line signature recognition system, where a hidden Markov model-based matching strategy is employed. The performance of a protected on-line signature recognition system employing the proposed BioConvolving approach is evaluated, both in terms of authentication rates and renewability capacity, using the MCYT signature database. The reported extensive set of experiments shows that protected and renewable biometric templates can be properly generated and used for recognition, at the expense of a slight degradation in authentication performance.

Index Terms—Biometrics, cancelable biometrics, hidden Markov model (HMM), security, signature verification, template protection.

I. INTRODUCTION

BIOMETRIC person recognition refers to the use of physiological or behavioral characteristics of people in an automated way to identify them or verify who they claim to be [1]. Biometric recognition systems are typically able to provide improved comfort and security to their users, when compared to traditional authentication methods, typically based on something that you have (e.g., a token) or something that you know (e.g., a password).

Unfortunately, biometrics-based people authentication poses new challenges related to personal data protection, not existing in traditional authentication methods. In fact, if biometric data

Manuscript received November 28, 2008. First published March 22, 2010; current version published April 14, 2010. This paper was recommended by Guest Editor K. W. Bowyer.

- E. Maiorana, P. Campisi, and A. Neri are with the Dipartimento di Elettronica Applicata, Università degli Studi Roma Tre, 00146 Roma, Italy (e-mail: maiorana@uniroma3.it; campisi@uniroma3.it; neri@uniroma3.it).
- J. Fierrez and J. Ortega-Garcia are with the Biometric Recognition Group—ATVS, Escuela Politecnica Superior, Universidad Autonoma de Madrid, 28049 Madrid, Spain (e-mail: javier.ortega@uam.es; julian.fierrez@uam.es).

Digital Object Identifier 10.1109/TSMCA.2010.2041653

are stolen by an attacker, this can lead to identity theft. Moreover, users' biometrics cannot be changed, and they may reveal sensitive information about personality and health, which can be processed and distributed without the users' authorization [2]. An unauthorized tracking of the enrolled subjects can also be done when a cross-matching among different biometric databases is performed, since personal biometric traits are permanently associated with the users. This would lead to users' privacy loss.

Because of these security and privacy issues, there are currently many research efforts toward protecting biometric systems against possible attacks which can be perpetrated at their vulnerable points (see [3]). In essence, the adopted security measures should be able to enhance biometric systems' resilience against attacks while allowing the matching to be performed efficiently, thus guaranteeing acceptable recognition performance.

In this paper, we introduce a novel noninvertible transformbased approach, namely, BioConvolving, which provides both protection and renewability for any biometric template which can be expressed in terms of a set of discrete sequences related to the temporal, spatial, or spectral behavior of the considered biometrics. The proposed approach can be therefore applied to a variety of biometric modalities, for example, speech biometrics [4], where spectral or temporal analysis of the voice signal produces discrete sequences, or to signature [5] and handwriting [4] recognition, where the extracted sequences are related to the pen's position, applied pressure, and inclination. Moreover, when performing gait recognition [6], temporal sequences describing the trajectories of the ankle, knee, and hip of walking people can be considered as templates. A set of discrete finite sequences representing the potentials of brain electrical activity, generated as a response to visual stimuli, can also be employed as a template, when performing brain-activity-based identification [7]. This is also the case when performing iris recognition, since the normalized template can be decomposed into 1-D intensity signals, which retain the local variations of the iris [8].

It is worth pointing out that some methods for the protection of templates extracted from the aforementioned biometrics act on sets of parametric features derived from the originally acquired data, thus limiting the kind of matching which can be performed [9]. Since our BioConvolving approach deals with discrete sequences instead of parametric features, it allows using sophisticated matching schemes such as dynamic time warping (DTW) or hidden Markov models (HMMs).

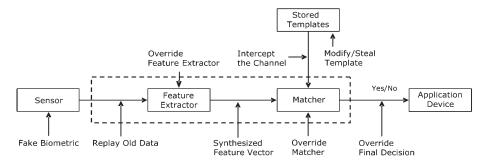


Fig. 1. Points of attack in a generic biometric system (adapted from [3]).

As a proof of concept, we apply the proposed BioConvolving protection scheme to signature biometrics, extending the authors' works presented in [10] and [11]. Specifically, the signature representation here employed comprises a higher number of signature sequences, and a detailed renewability and security analysis is carried out.

Specifically, this paper is organized as follows. In Section II, the different solutions which have been investigated in the recent past to secure biometric templates are analyzed. The proposed approach for the protection of sequence-based biometric templates is illustrated in Section III, and its security analysis is outlined in Section IV. The state of the art on signature-based authentication schemes and on signature template protection approaches is outlined in Section V. The application of the proposed scheme to on-line signature biometrics is presented in Section VI, which details both the employed signature representation and the employed template matcher. The experimental setup is described in Section VII. The authentication and the renewability performances of the proposed protection approach are discussed in Sections VIII and IX, respectively. Eventually, some conclusions are drawn in Section X.

II. BIOMETRIC TEMPLATE SECURITY

A biometric system can be roughly sketched as that in Fig. 1 and consists of a sensor module, a feature extractor module, a matcher, a database, and an application device which is driven by the matcher output. As discussed in [3] and also shown in Fig. 1, eight possible vulnerable points can be identified in a biometric system. Specifically, some attacks can be perpetrated at the sensor level, at the feature extractor level, or against the channel interconnecting two modules, in order to steal or substitute the acquired biometric information with fake or other impostor data.

Other attacks are related to the biometric templates generated by the feature extractor module, which are stored in the database or matched against previously stored templates. The biometric templates are the targets of the attacks either at the database level or at the interconnecting channel level. Finally, the matcher and the output to the device application can be attacked to override the system decision.

The unauthorized access to both raw biometric data and biometric templates is among the most dangerous threats to users' privacy and security. Several techniques for biometric template protection have been studied in the literature. Among them, classical cryptographic techniques [12] can be employed

to secure the transmission of biometric data over reliable but insecure channels and to store data in such a way that they are intelligible only by using a proper cryptographic key. However, when using these techniques, it is necessary to perform the match after decryption, and therefore, no protection is provided during the matching.

On the other hand, data hiding techniques [13] can be used to insert additional information, namely, the watermark, into a digital object. Within this respect, data hiding techniques complement encryption, since the message can remain in the host data even when decryption has been done. The use of data hiding techniques for biometrics protection has already been proposed for fingerprints [14] and signatures [15], among others.

The problem of providing protection to the biometric templates also during the matching process can be solved with new techniques such as *cancelable biometrics*, also known as *anonymous* or *revocable biometrics*. Cancelable biometrics have been introduced in [16], where template protection has been achieved by applying an intentional and repeatable modification to the original biometric template. The transformation must be designed in such a way to satisfy the following properties.

- Renewability: It should be possible to revoke a template and to reissue a new one based on the same biometric data. The new templates should not match with the ones revoked in order to provide *diversity*. This property is needed to ensure the user's privacy.
- 2) Security: It should not be possible, or computationally unfeasible, to obtain the original raw biometric data from the stored secured template. This property is needed to prevent an adversary from reconstructing biometric traits from a single stolen template, as well as from several stolen templates (this is commonly referred to as the record multiplicity attack). In fact, although it was commonly believed that it is not possible to reconstruct the original biometric characteristics from the corresponding extracted template, some concrete counter examples have been provided in the recent literature [17], [18].
- 3) *Performance*: The biometric recognition error rates should not degrade significantly with the introduction of a template protection scheme, with respect to an unprotected approach. Moreover, the recognition performances should not be sensitive to the employed modifications: Even when applying different distortions to the same biometric data, the recognition performances should show a very low variance.

A detailed discussion regarding the requirements of a properly defined cancelable biometrics can also be found in [19].

Designing a template protection scheme that is able to properly satisfy each of the aforementioned properties is not a trivial task, mainly due to the unavoidable intrauser variability shown by every biometric trait. Different solutions have already been proposed for the generation of secure and renewable templates. A recent survey of published methods has been presented in [20], where the authors have classified the existing approaches into two categories: *biometric cryptosystems* and *feature transformation* approaches.

A. Biometric Cryptosystems for Template Protection

Biometric cryptosystems combine cryptographic keys with transformed versions of the input biometrics to generate the secured templates [21]. In this process, some public information, namely, helper data, is generated. Biometric cryptosystems can be further divided into key binding systems, where the helper data are obtained by combining the key with the biometric template, and key generation systems, where both the helper data and the key are directly generated from the biometric template. Two of the most well-known examples of key binding approaches are the fuzzy commitment [22] and the fuzzy vault [23], which represent general schemes that can be applied to different biometrics such as fingerprints or face [24], [25]. Typically, these approaches are able to manage the intrauser variations in biometric data by exploiting the capabilities of error correcting codes. However, it is generally not possible to use sophisticated and dedicated matchers, thus reducing the system matching accuracy. Moreover, it has been proven that the fuzzy vault is vulnerable to the record multiplicity attack [26]: If an adversary has access to two different vaults obtained from the same data, he can easily identify the genuine points in the two vaults. On the other hand, the proposed key generation biometric cryptosystems have been more difficult to implement in practice [27].

B. Feature Transformations for Template Protection

In a feature transformation approach, a function that is dependent on some parameters, which can be used as a key, is applied to the input biometric to generate the protected templates. The employed function can be either *invertible*, resulting in a *salting* approach, whose security is based on the protection of the function parameters, or *noninvertible*, when a one-way function is applied to the template and it is computationally hard to invert the function even if the transformation parameters are known. The use of the methods belonging to the first category typically results in low false acceptance rates; however, if a user-specific key is compromised, the user template is no longer secure due to the invertibility of the transformation. Examples can be found in [28] and [29].

On the contrary, when noninvertible transforms are used, even if the key is known by an adversary, no significant information can be acquired on the template, thus obtaining better security than when using a salting approach, which relies on the key security. Moreover, in contrast with cryptosystem

approaches, the transformed templates can remain in the same feature space of the original ones, being then possible to employ standard matchers to perform authentication in the transformed domain. This guarantees performances that are similar to those of an unprotected approach. In addition to the performance benefits of using standard matchers in the transformed domain, these methods typically result in matching scores which can be fused in multibiometric approaches. Therefore, the use of transform-based approaches for template protection in multibiometrics systems allows using either score-level fusion techniques [30] or decision-level fusion techniques [31], whereas only the latter, which is less effective than the former, can be employed when biometric cryptosystems are considered. Unfortunately, it seems to be difficult to design transformation functions which can satisfy both the discriminability and the noninvertibility properties simultaneously.

The concept of achieving template security through the application of noninvertible transformations has been first presented in [16], where it has been referred to as cancelable biometrics as that in [32], although this expression has been later conceived in a more general sense. One of the first published works including experimental evidence on the feasibility of noninvertible transforms for biometric template protection is [33], where a geometric transform has been employed to protect minutia templates. However, the protection scheme in [33] introduces a significant performance degradation, and the matching score between fingerprints transformed with different keys is relatively high, thus greatly reducing the useful key space. More general geometric transforms (Cartesian, polar, and functional) have been later studied in [34], where better performances have been achieved. However, with reference to the best approach presented in [34], only a small fraction of the data, namely, 8%, is noninvertible in practice [35]. Moreover, all the approaches for template protection in [33] and [34] are vulnerable to a record multiplicity attack: Having access to two or more different transformed versions of the same minutia pattern, it is possible to identify the original positions of the considered minutiae [36].

A registration-free construction of cancelable fingerprint templates has also been proposed in [37]. From each detected minutia, a square patch is extracted and transformed using an orthogonal transformation matrix. The approach presented in [37], being able to withstand also a record multiplicity attack, is more robust than the one proposed in [34], but it exhibits lower verification performances than the one obtained in [34].

Voice-based cancelable templates were proposed in [38], where a noninvertible transformed version of the originally acquired voiceprint is generated. The original biometrics cannot be obtained from the template stored in the server during enrollment, even if the keys employed for transformations are disclosed.

III. GENERATING CANCELABLE SEQUENCE-BASED BIOMETRIC TEMPLATES

The proposed BioConvolving approach provides protection to templates characterized by a set of discrete finite sequences

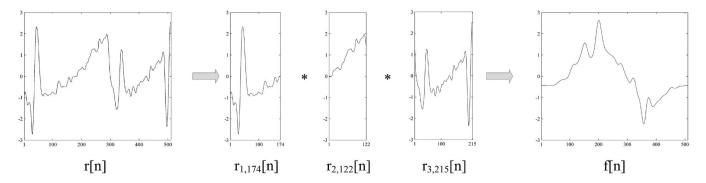


Fig. 2. Baseline approach. Example of a template sequence transformation, where W=3.

extracted from a given biometrics, by applying the transformations defined in Sections III-A and B. The resulting transformed sequences can then be further processed, if the matcher is based on a sequence-based modeling approach (e.g., HMM), or directly stored as templates, if the matcher works directly with sequence-based descriptions (e.g., DTW). Specifically, it is assumed that the proposed transformations can be applied to an original set of sequences \mathcal{R}_F , consisting of F sequences $r_{(i)}[n], i = 1, \dots, F$. The transformed template is indicated as \mathcal{T}_F and consists of F sequences $f_{(i)}[n], i = 1, \dots, F$. In Section III-A, the baseline sequence-based template transform, specifically designed in such a way that it is not possible to retrieve the original data from the transformed ones, is proposed. Moreover, in Section III-B, some alternatives for the protection of sequence-based templates, derived from the baseline approach in Section III-A, will be detailed.

A. Noninvertible Transform: Baseline Approach

Let us consider the set of transformations that are necessary to generate the transformed template, represented by the set of sequences \mathcal{T}_F , by using the original template, given by the set of original sequences \mathcal{R}_F . These transformations are designed in order to satisfy the following properties.

- 1) Each transformed sequence, belonging to the set T_F , must be generated by using at least two sequences, which can be either an original sequence or a segment extracted from an original sequence.
- 2) Each sequence employed in one transformation cannot occur in any other one of the set of transformations employed to generate the transformed template \mathcal{R}_F .

In the baseline implementation, each transformed sequence $f_{(i)}[n], i=1,\ldots,F$, is obtained from the corresponding original sequence $r_{(i)}[n], i=1,\ldots,F$, which represents a generic discrete sequence of length N belonging to the original template, as follows.

A number (W-1) of different integer values d_j between 1 and 99 are randomly selected, ordered in ascending order such that $d_j > d_{j-1}, j = 1, \ldots, W$, and arranged in a vector

$$\mathbf{d} = [d_0, \dots, d_W]^{\mathrm{T}},\tag{1}$$

where d_0 and d_W are set to 0 and 100, respectively. The vector d represents the key of the employed transformation. Then, the

original sequence $r_{(i)}[n]$ is divided into W segments $r_{(i)j,N_j}[n]$ of length $N_j=b_j-b_{j-1}$

$$r_{(i)j,N_j}[n] = r_{(i)}[n+b_{j-1}], \qquad n = 1, \dots, N_j;$$

 $j = 1, \dots, W,$ (2)

where

$$b_j = \left\lceil \frac{d_j}{100} \cdot N \right\rceil, \qquad j = 1, \dots, W. \tag{3}$$

Basically, the sequence $r_{(i)}[n]$ is split into W nonoverlapping parts according to the randomly generated vector \mathbf{d} , as shown in Fig. 2 for the case with W=3. A transformed sequence $f_{(i)}[n], n=1,\ldots,K$, is then obtained through the linear convolution of the sequences $r_{(i)j,N_i}[n], j=1,\ldots,W$, i.e.,

$$f_{(i)}[n] = r_{(i)1,N_1}[n] * \cdots * r_{(i)W,N_W}[n].$$
 (4)

Each transformed sequence $f_{(i)}[n]$ is therefore obtained through the linear convolution of parts of the corresponding original sequences $r_{(i)}[n], i=1,\ldots,F$. Moreover, each original sequence $r_{(i)}[n], i=1,\ldots,F$, undergoes the same decomposition before applying the convolutions. The length of the transformed sequences obtained by means of convolution as that in (4) is equal to K=N-W+1, which is therefore almost the same of the original sequences. A final signal normalization, to obtain zero-mean and unit-standard-deviation transformed sequences, is then applied. Different realizations can be obtained from the same original sequences, simply varying the size or the values of the parameter key d. The complete set of transformed sequences $f_{(i)}[n], i=1,\ldots,F$, is indicated as \mathcal{T}_F . The security analysis of the proposed sequence-based protection scheme is conducted in Section IV.

B. Noninvertible Transform: Extended Approaches

In the previous section, we illustrated how to generate a transformed sequence from an original one. However, as it will be shown in Section IX, when considering the application to the protection of on-line signature templates, the baseline method possesses a low renewability capability. In order to properly address this issue, two additional noninvertible sequence-based approaches, stemming from the approach in Section III-A, are proposed in the following.

1) Mixing Approach: This approach is defined by considering, in addition to the decomposition key \mathbf{d} , a transformation key \mathbf{C} , defined as a matrix of F rows and W columns. Each column of \mathbf{C} is obtained as a scrambled version of the vector $[1,\ldots,F]^{\mathrm{T}}$. An example of a possible matrix \mathbf{C} , for F=7 and W=4, can be

$$\mathbf{C} = \begin{bmatrix} 1 & 4 & 3 & 7 \\ 2 & 7 & 2 & 5 \\ 3 & 1 & 6 & 1 \\ 4 & 2 & 7 & 3 \\ 5 & 6 & 1 & 4 \\ 6 & 5 & 5 & 2 \\ 7 & 3 & 4 & 6 \end{bmatrix} . \tag{5}$$

Each row of the matrix \mathbf{C} , i.e., C[i,j] with $j=1,2,\ldots,W$, is employed to define the combinations that originate the transformed sequences $f_{(i)}[n]$ as follows:

$$f_{(i)}[n] = r_{(C[i,1])1,N_1}[n] * \cdots * r_{(C[i,W])W,N_W}[n]$$
 (6)

with $i=1,\ldots,F$, and where $r_{(i)j,N_j}[n]$ is defined as that in (2). Basically, each transformed sequence $f_{(i)}[n]$ is generated not only from the corresponding original sequence $r_{(i)}[n]$, but the convolutions are performed among segments extracted from different original sequences, thus also defining a feature-level fusion [30] among various sequences.

2) Shifting Approach: Another variation to the approach in Section III-A is obtained by applying an initial shift to the original sequences $r_{(i)}[n], i=1,\ldots,F$. Specifically, a random integer value ϕ is selected in the range [0, 100] and converted to the shift s as

$$s = \left\lceil \frac{\phi}{100} \cdot N \right\rceil,\tag{7}$$

with N being the length of the original sequence, in sample units. Then, each sequence $r_{(i)}[n]$ undergoes the same circular shift ruled by the parameter s, thus obtaining the sequences $z_{(i)}[n] = r_{(i)}[n-s], n=1,\ldots,N$.

The same transformation process described in Section III-A, based on convolutions between segments extracted from the considered sequences, is then applied to the sequences $z_{(i)}[n]$. This modification can also be combined with the extended method presented in Section III-B1, by applying the circular shift before performing the transformations. Obviously, it is also possible to apply different initial shifts to the F sequences before performing the decompositions, in order to further increase the transformation key space. However, in this paper, we only consider the case where the same shift is applied to all the available original sequences.

IV. TRANSFORM INVERTIBILITY ANALYSIS

The analysis of the invertibility, i.e., the possibility of recovering the original sequences from the ones obtained employing the proposed transformation schemes, is investigated in this section. Specifically, this analysis, being related only to the

transformations designed in Section III, does not depend on a specific biometric modality. Furthermore, being the methods in Section III-B derived as extensions of the principal approach described in Section III-A, only the latter one is here analyzed, due to the fact that the security of the extended methods depends on the one provided by the baseline approach.

Having defined the sequence transformation as that in (4), if an attacker gains access to the stored information, he has to solve a *blind deconvolution* problem [39]–[41] to retrieve any information regarding the original sequences. In other words, the security of the proposed sequence-based template protection methods relies on the difficulty in solving a blind deconvolution problem, having no *a priori* knowledge about the original sequences.

The proposed transformation is also robust to the record multiplicity attack, where it is assumed that different transformed templates based on the same original data are available to the attacker. It is worth pointing out that this is a worst case condition because, in real-life applications, the realizations of the original biometrics used in different applications vary depending on the intra-user biometric variability. Under this assumption, we then consider that an attacker has acquired, from two different systems, two different transformed sets of sequences $\mathcal{T}_F^{(1)}$ and $\mathcal{T}_F^{(2)}$, generated from the same original template \mathcal{R}_F by applying different transformation parameters. Considering the simplest case with W=2, the attacker then possesses two transformed instances, namely, $f^{(1)}[n]$ and $f^{(2)}[n]$, of the same original sequences r[n], obtained using the two transformation parameters $d_1^{(1)}$ and $d_1^{(2)}$. Given that

$$\begin{split} r[n] &= r_{1,N_1^{(1)}}^{(1)}[n] + r_{2,N_2^{(1)}}^{(1)} \left[n - b_1^{(1)} \right] \\ &= r_{1,N_1^{(2)}}^{(2)}[n] + r_{2,N_2^{(2)}}^{(2)} \left[n - b_1^{(2)} \right], \end{split} \tag{8}$$

in order to recover the sequence r[n], the attacker should obtain the segments $r_{1,N_1^{(1)}}^{(1)}[n]$ and $r_{2,N_2^{(1)}}^{(1)}[n]$, where $N_1^{(1)}=b_1^{(1)}$ and $N_2^{(1)}=N-b_1^{(1)}$, or the segments $r_{1,N_1^{(2)}}^{(2)}[n]$ and $r_{2,N_2^{(2)}}^{(2)}[n]$, with $N_1^{(2)}=b_1^{(2)}$ and $N_2^{(2)}=N-b_1^{(2)}$, from the available transformed sequences $f^{(1)}[n]=r_{1,N_1^{(1)}}^{(1)}[n]*r_{2,N_2^{(1)}}^{(1)}[n]$ and $f^{(2)}[n]=r_{1,N,2^{(1)}}^{(2)}[n]*r_{2,N_2^{(2)}}^{(2)}[n]$.

Deconvolution problems are typically coped with in the frequency domain, being the convolutions represented by multiplications in the Fourier domain. In order to properly define the discrete Fourier transforms (DFTs) of the segments extracted from r[n], the extended versions $\hat{r}_{h,K}^{(j)}[n], h, j = \{1,2\}$, are generated by applying a zero padding to the respective original traits, until reaching the same length K=N-1 of the convolution products $f^{(1)}[n]$ and $f^{(2)}[n]$. Then, the sequence $\Delta[n], n=1,\ldots,K$, is defined as the difference between $\hat{r}_{1,K}^{(1)}[n]$ and $\hat{r}_{1,K}^{(2)}[n]$, which share a common part that is exactly $r_{1,N_1^{(2)}}^{(2)}[n]$, having assumed that $b_1^{(1)} > b_1^{(2)}$

$$\Delta[n] = \hat{r}_{1,K}^{(1)}[n] - \hat{r}_{1,K}^{(2)}[n], \qquad n = 1, \dots, K.$$
 (9)

The following relations can then be derived for the considered finite sequences:

$$\begin{cases} \hat{r}_{1,K}^{(1)}[n] = \hat{r}_{1,K}^{(2)}[n] + \Delta[n] \\ \hat{r}_{2,K}^{(1)}\left[n - b_1^{(1)}\right] = \hat{r}_{2,K}^{(2)}\left[n - b_1^{(2)}\right] - \Delta[n], \end{cases}$$
(10)

where all the considered shifts are circular shifts. Then, applying the DFT to the *a priori* known sequences $f^{(1)}[n]$ and $f^{(2)}[n]$ and considering the relations between the DFT and the linear convolution of two discrete sequences, it results to (11), shown at the bottom of the page, where the DFT coefficients are indexed with l. Using the relations in (10), the first equation in (11) can be written as

$$\text{DFT} \left\{ f^{(1)}[n] \right\} = \left[\text{DFT} \left\{ \hat{r}_{1,K}^{(2)}[n] \right\} + \text{DFT} \left\{ \Delta[n] \right\} \right]$$

$$\cdot \left[\text{DFT} \left\{ \hat{r}_{2,K}^{(2)} \left[n - b_1^{(2)} \right] \right\} - \text{DFT} \left\{ \Delta[n] \right\} \right]$$

$$\cdot e^{j2\pi(l/K)b_1^{(1)}},$$

$$(12)$$

from which the expressions in (13), shown at the bottom of the page, can be derived.

The resulting system of equations admits ∞^1 possible solutions, which implies that recovering the original segments $\hat{r}_{1,K}^{(2)}[n]$ and $\hat{r}_{2,K}^{(2)}[n]$ is as much hard as random guessing them. The difficulty in reaching a solution for the original sequence observed in our formulation corroborates the difficulty in succeeding in a record multiplicity attack.

V. SIGNATURE BIOMETRICS

A. Signature-Based Authentication

People recognition based on signatures is one of the most accepted biometric-based authentication methods since, being part of everyday life, it is perceived as a noninvasive and non-threatening process by the majority of the users. Furthermore, a signature has a high legal value. On the other hand, this modality is characterized by a high intrauser variability, due to the fact that signatures can be influenced by several physical and emotional conditions, and a small forgery inter-user variability, which must be taken into account in the authentication process. A review of the state of the art covering the literature up to 1993

can be found in [42]. Other survey papers quoting the more recent advances in signature recognition are [43] and [44].

Signature-based authentication can be either *static* or *dynamic*. In the *static* mode, also referred to as off-line, only the written image of the signature, typically acquired through a camera or an optical scanner, is used. In the *dynamic* mode, also called on-line, signatures are acquired by means of a graphic tablet or a pen-sensitive computer display, which can provide temporal information about the signature, such as the pressure, the velocity, the pen tilt signals versus time, etc.

In order to represent the signature, some features must be extracted. Two different kinds of features are typically considered: *parameters* and *functions*. Parametric features can consist of static information, like the height and the width of the signatures, or dynamic information, like signature velocity, acceleration, or pressure. In most comparative studies, the parameters based on dynamic information are typically more discriminative for recognition purposes than those based on static information [45]. On the other hand, sequence-based methods typically use a representation based on various temporal sequences and elastic matching procedures such as DTW, which represents one of the more flexible approaches to manage the signature length variability [46], or statistical recognition approaches such as HMMs [5], [47].

B. Signature Template Protection: Related Works

Signature template protection has been first considered in [9] and [48] with a key generation approach which extracts a set of parametric features from the acquired dynamic signatures and applies a hash function to a feature's binary representation, obtained by exploiting some statistical properties of the enrollment signatures. Both methods provide protection for the signature templates, but none of them provides revocability. The fuzzy vault construction has been applied to signature verification in [49], by using a quantized set of maxima and minima of the temporal functions mixed with chaff points in order to provide security. A salting approach has been proposed in [50] as an adaptation of the BioHashing method [28] to signature templates. The fuzzy commitment approach introduced in [22] has also been applied to signature verification in [51] and [52]. In both papers, a practical implementation of fuzzy commitment [25] has been taken into account, and a new user-adaptive

$$\begin{cases}
DFT \left\{ f^{(1)}[n] \right\} = DFT \left\{ \hat{r}_{1,K}^{(1)}[n] \right\} \cdot DFT \left\{ \hat{r}_{2,K}^{(1)}[n] \right\} = DFT \left\{ \hat{r}_{1,K}^{(1)}[n] \right\} \cdot DFT \left\{ \hat{r}_{2,K}^{(1)}\left[n - b_1^{(1)}\right] \right\} \cdot e^{j2\pi(l/K)b_1^{(1)}} \\
DFT \left\{ f^{(2)}[n] \right\} = DFT \left\{ \hat{r}_{1,K}^{(2)}[n] \right\} \cdot DFT \left\{ \hat{r}_{2,K}^{(2)}[n] \right\}
\end{cases}$$
(11)

$$\begin{cases}
DFT \left\{ f^{(1)}[n] \right\} = e^{j2\pi(l/K)b_1^{(1)}} \cdot \left[DFT \left\{ \hat{r}_{1,K}^{(2)}[n] \right\} \cdot DFT \left\{ \hat{r}_{2,K}^{(2)}[n] \right\} \cdot e^{-j2\pi(l/K)b_1^{(2)}} - DFT \left\{ \Delta[n] \right\} \cdot DFT \left\{ \hat{r}_{1,K}^{(2)}[n] \right\} \\
+ DFT \left\{ \Delta[n] \right\} \cdot DFT \left\{ \hat{r}_{2,K}^{(2)}[n] \right\} \cdot e^{-j2\pi(l/K)b_1^{(2)}} - DFT^2 \left\{ \Delta[n] \right\} \right] \\
DFT \left\{ f^{(2)}[n] \right\} = DFT \left\{ \hat{r}_{1,K}^{(2)}[n] \right\} \cdot DFT \left\{ \hat{r}_{2,K}^{(2)}[n] \right\}
\end{cases} (13)$$

error-correcting code selection has also been introduced. The implementation of a security scalable recognition system by exploiting watermarking-based techniques has been studied in [15], [52], and [53]. No template-transformation-based approach has been proposed so far for the protection of signature biometrics.

VI. APPLICATION TO AN ON-LINE SIGNATURE RECOGNITION SYSTEM

The effectiveness of the proposed protection scheme for sequence-based biometrics is here applied to the protection of on-line signature templates. In Section VI-A, it is discussed how to extract a sequence-based template \mathcal{R}_F from an acquired signature, while the employed classifier, based on HMM, is described in Section VI-B.

A. Feature Extraction Stage

During the employed feature extraction stage, the horizontal x[n] and vertical y[n] position trajectories, together with the pressure signal p[n], n = 1, ..., N, are acquired from each online signature through a digitizing tablet. We consider that the signals x[n] and y[n] are already normalized both in position, with respect to their center of mass, and in rotation, with respect to their average path tangent angle. Other four discrete-time sequences are derived from the pair $\{x[n], y[n]\}$, namely, the path tangent angle $\theta[n]$, the path velocity magnitude v[n], the log curvature radius $\rho[n]$, and the total acceleration magnitude a[n]. Specifically, in our experiments, we consider the following set of F = 14 sequences:

$$\mathcal{R}_{14} = \left\{ x[n], y[n], p[n], \theta[n], v[n], \rho[n], a[n], \\ \dot{x}[n], y[\dot{n}], p[\dot{n}], \dot{\theta}[n], \dot{v}[n], \dot{\rho}[n], \dot{a}[n] \right\},$$
(14)

where the upper dot notation denotes the first-order derivative.

B. Signature Modeling

In order to perform signature recognition, a stochastic modelization based on HMMs is applied to the transformed signature templates.

An HMM is characterized by the following elements:

- 1) the number H of hidden states $\{S_1, S_2, \dots, S_H\}$ of the model. The state at discrete time n is indicated as q_n ;
- 2) the state transition probability $A = \{a_{i,j}\}$, where $a_{i,j} =$ $P[q_{n+1} = S_j | q_n = S_i], i, j = 1, \dots, H;$
- the observation symbol probability distributions in each state j, indicated with $\mathbf{B} = \{b_i(\mathbf{o})\}, j = 1, \dots, H$. The observation processes are represented using mixtures of M multivariate Gaussian distributions: $b_i(\mathbf{o}) =$ $\sum_{m=1}^{M} \alpha_{j,m} p_{\mu_{j,m},\Sigma_{j,m}}(\mathbf{o}), j=1,\ldots,H, \text{ where } \mu_{j,m}$ and $\Sigma_{j,m}$ indicate the mean and the diagonal covariance matrix of each Gaussian component, respectively. The coefficients $\alpha_{j,m}$ are selected by respecting the condition of normalization $\sum_{m=1}^{M} \alpha_{j,m} = 1, j = 1, \dots, H;$ 4) the initial state distribution $\pi = \{\pi_j\} = \{p[q_1 = 1], \dots, q_m\}$
- S_i], j = 1, ..., H.

Following the proposed approach, during the enrollment phase, the client model $\lambda = \{\pi, \mathbf{A}, \mathbf{B}\}$ is estimated considering E enrollment signatures of the subject at hand, according to the iterative strategy presented in [5].

The obtained model λ is stored in a database and used in the authentication phase, where a similarity score is calculated as $(1/K) \log P(\mathbf{O}|\lambda)$ using the Viterbi algorithm [54]. Specifically, the Viterbi algorithm is employed to estimate, given an observation sequence O and a model λ , the sequence Q of hidden states corresponding to O. The criterion followed by the Viterbi algorithm is to maximize the probability $P(\mathbf{Q}|\mathbf{O},\lambda)$, which is equivalent to maximizing $P(\mathbf{Q}, \mathbf{O}|\lambda)$. The Viterbi procedure can be efficiently represented by a lattice structure, where each node, at a given instant, represents the hidden state of the model. The computational complexity of the algorithm is reduced when maximizing the log likelihood, with respect to the likelihood of the test sample path given the model. The ratio 1/K is taken into account to normalize to the obtained log likelihood, which decreases when the length of the test signature increases [55].

It is worth pointing out that, when using HMMs for signature recognition, also in an unprotected approach, the client model $\lambda = \{\pi, A, B\}$, instead of the original signature sequences, is stored in the database. However, if an attacker is able to acquire the client HMM, the statistical properties of the client's signatures can be derived from the model and, for example, employed to track the users across multiple databases. Using the proposed protection approach, if an attacker succeeds in acquiring the stored models, he can only retrieve information about the set of transformed sequences T_F , from which it is not possible to get any information about the original sequences $r_{(i)}[n], i = 1, \dots, F$, as discussed in Section IV.

VII. EXPERIMENTAL SETUP

The noninvertible transforms, proposed for the protection of sequence-based biometrics, are tested by verifying both their renewability capabilities and the verification performance achievable in protected systems, with application to on-line signature biometrics. For the experiments, we use the MCYT on-line signature corpus [57]. This database includes signatures from 330 subjects, with 25 genuine signatures per subject. These genuine signatures have been captured in sets of five, allowing some breaks between the different acquisition sets. For each subject, there are also 25 forgeries performed by five different forgers for each subject. Forgers have been asked to reproduce without breaks or slowdowns a signature after having observed the static image of the prototype and after having carried out a training stage, which consists of copying the prototype at least ten times.

In the experiments, we have studied the following key aspects of the proposed template protection approaches:

- 1) Authentication performance
 - a) performance dependence on HMM parameters, for both unprotected and protected systems;
 - b) performance variability with respect to the transformation-defining parameters, for protected systems;

				<u> </u>	
Н	M	Unprotected	Protected Baseline Approach		
		Approach	W=2	W = 3	W = 4
	1	13.38	10.87	13.96	16.85
8	2	9.73	8.67	12.40	15.40
	4	7.84	8.08	11.84	15.78
	8	8.03	7.95	12.36	16.63
	16	8.11	8.53	13.11	18.36
	32	10.73	9.05	16.61	16.94
	1	8.61	8.77	16.95	15.68
16	2	6.47	8.01	12.14	15.56
	4	6.33	8.24	12.51	16.67
	8	6.69	8.97	16.45	19.90
	16	8.23	9.51	19.74	20.36
	32	13.05	20.4	20.60	20.76

c) performance comparison between the baseline approach described in Section III-A and the extended methods in Section III-B;

2) Renewability

a) evaluation of the *diversity* between two templates originated by applying two different transformations on the same original data. The analysis is conducted for the baseline approach described in Section III-A, as well as for the extended methods in Section III-B.

The performance analysis is detailed in Section VIII, while the renewability capabilities of the proposed protection methods are presented in Section IX.

VIII. AUTHENTICATION PERFORMANCE ANALYSIS

The authentication performances achievable with the proposed protected on-line signature protection methods are here discussed. The system performances are evaluated through the false rejection rate (FRR), the false acceptance rate (FAR) for skilled forgeries (FAR $_{\rm SF}$), the FAR for random forgeries (FAR $_{\rm RF}$), and the equal error rate (EER). These figures of merit are obtained by considering, for each user in the enrollment stage, E=10 signatures taken from the first two acquisition sets of MCYT. The FRR is estimated on the basis of the signatures belonging to the third, fourth, and fifth available acquisition sets. The FAR $_{\rm SF}$ is computed by using the 25 skilled forgeries available for each user. The FAR $_{\rm RF}$ is computed by taking, for each user, one signature from each of the remaining users.

A. Dependence on the HMM Parameters

Within the described experimental setup, the dependence of the authentication performances on the HMM parameters is first discussed. Specifically, the EERs obtained by varying the HMM parameters H and M, considering skilled forgeries,

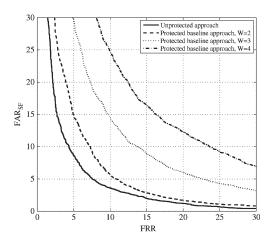


Fig. 3. ROC curves for an unprotected system, and for protected systems with $W=2,\ 3,\ {\rm and}\ 4$ convolved segments, considering skilled forgeries and E=10.

are summarized in Table I, for both unprotected systems and for protected systems employing the baseline approach described in Section III-A, with $W \in \{2,3,4\}$. Specifically, the values of H reported in Table I are $H \in \{8,16\}$, since the best recognition rates are achieved when using, for the HMM modelization, a number of states comprised between 8 and 16, as observed in [5] and [10]. When employing the proposed baseline protection approach, the key vector \mathbf{d} is randomly selected for each considered user, taking the values $d_j, j = 1, \ldots, W-1$, in the range of integers [5, 95]. As described in [34], this reflects how the protected system should be used in a practical implementation, where different transformations are typically used for different individuals.

The best EERs achievable for each configuration (unprotected and protected systems) are highlighted in Table I and are employed to select the best HMM configurations, which are considered in the following to illustrate the performances of the proposed approaches. Specifically, the selected configurations are as follows:

- 1) unprotected approach: H = 16 and M = 4 (EER_{SF} = 6.33%);
- 2) baseline protected approach, with W=2: H=8 and M=8 (EER_{SF} = 7.95%);
- 3) baseline protected approach, with W=3: H=8 and M=4 (EER_{SF} = 11.84%);
- 4) baseline protected approach, with W=4: H=8 and M=2 (EER_{SF} = 15.40%).

The receiver operating characteristic (ROC) curves related to the best authentication rates, achievable using the aforementioned selected configurations, are shown in Fig. 3. From the sketched ROC curves and from the results in Table I, it can be seen that the recognition performances for protected systems worsen when the number W of segments in which the signatures are segmented increases. The loss in performance can be explained as follows. The segmentation of the considered signature time sequences is accomplished by using a set of fixed parameters d_j , $j=1,\ldots,W-1$. They express, in terms of the percentage of the total sequence length, the points where the segmentation has to be done. However, due to the characteristics of signature biometrics, sequences extracted from different

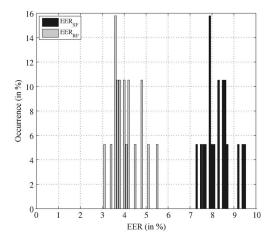


Fig. 4. Normalized histograms of the EERs obtained repeating 20 times the authentication process, for a protected system with W=2.

signatures, also if from the same user, typically have different lengths, which raises an alignment issue. As a consequence, the more separations are performed, the more variable will be the convolutions at the output. The best results are obtained when W=2, due to the fact that only one separation point has to be set in this case. However, the performances achieved with W=3 still remain acceptable, producing an EER for skilled forgeries of about 12%, when taking E=10 signatures for the enrollment. The cited alignment problem can be mitigated by using a dynamic programming strategy, as in the DTW approaches for signature recognition [46], whereas a simple linear correspondence strategy does not represent the best signature alignment approach.

B. Dependence on the Transform Key Vector d

The dependence of the authentication performance on the key d is investigated referring to the baseline approach proposed in Section III-A. More in detail, a protected system, where the signature functions are split into W=2 segments, by means of the key d, is considered. The performance evaluation is made performing 20 times the enrollment and authentication processes over the available test data set, varying at each iteration the transformation parameters d for each user. In Fig. 4, the obtained results are shown, through the normalized histograms of the EERs for both random (EERRF) and skilled forgeries (EERSF), obtained when considering a protected system where E=10 signatures are taken from each user during enrollment. The mean and standard deviation of the obtained EERs are as follows:

- 1) skilled forgeries: mean EER_{SF} = 8.2%, with a standard deviation $\sigma_{\rm EER_{SF}} = 0.7\%$;
- 2) random forgeries: mean EER $_{RF}=4.1\%,$ with a standard deviation $\sigma_{\rm EER_{RF}}=0.5\%.$

As necessary for a properly designed noninvertible transform approach, the variation of the transformation parameters does not result in significant modifications of the matching performances.

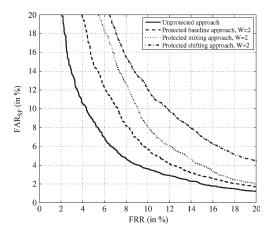


Fig. 5. Performance comparison between the baseline protected system in Section III-A and the extended protection approaches in Sections IX-B and IX-C, considering W=2 convolved segments for template protection.

C. Comparison Between Baseline and Extended Approaches

The proposed approaches for the protection of signature templates are also discussed by comparing the authentication performances achievable when employing the extended transforms described in Section III-B, with those obtained by using the baseline method described in Section III-A. Specifically, only the case where each function is split into W=2 segments is considered.

Fig. 5 shows the performances obtained when considering the mixing and shifting approaches described in Sections III-B1 and III-B2, respectively. The performances of the extended methods are also compared with those related to the use of the baseline protection approach. E = 10 signatures are considered to be taken from each user during enrollment. For all the considered protected approaches, the HMM configuration which gives the best authentication performances for the baseline method is considered (H = 8 and M = 8). The recognition rates shown for the unprotected system are related to the HMM configuration (H = 16 and M = 4) which allows obtaining the best achievable authentication performance. As shown in Fig. 5, systems using the mixing-based protection method, described in Section III-B1, are characterized by almost the same performances of a system using the baseline protection scheme, resulting in an EER of 9.12%. On the other hand, the protection method based on shifting, described in Section III-B2, provides slightly worse results, reaching an EER of about 10.81%.

IX. RENEWABILITY ANALYSIS

The transformations introduced in Sections III-A and B are then analyzed with respect to the *diversity* property, which is a crucial requirement to implement cancelable biometrics. Specifically, it can be noticed that each of the proposed transforms is defined by means of a key or a set of keys and that different transformations can be obtained by varying the employed keys. Moreover, two transformed templates, generated from the same original data, are as more different as more distant the respective transformation keys are. With the space of possible keys finite, the number of possible instances, which can be generated from the same data and which are

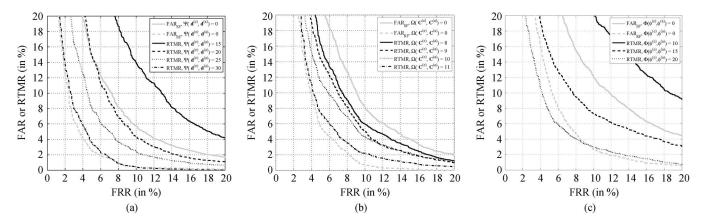


Fig. 6. Renewability analysis of the proposed approaches. The HMM configuration is H=8 states and M=8 Gaussian components per state. W=2 segments convolved for template protection. (a) Baseline approach (Section III-A). (b) Mixing-based approach (Section III-B1). (c) Shifting-based approach (Section III-B2).

distant enough from each other to properly respect the diversity requirement, is necessarily limited. The capability of the baseline approach described in Section III-A in generating multiple templates from the same original data is discussed in Section IX-A. Then, the renewability of the mixing and of the shifting approaches introduced in Section III-B is analyzed in Sections IX-B and C, respectively.

A. Baseline Approach

Considering the baseline approach detailed in Section III-A, the key of the employed transformation is represented by the vector \mathbf{d} , which specifies how to decompose the originally acquired functions into W parts, before performing the proposed transformation given by (4). In the considered experiments, for the sake of simplicity, the values which each element $d_j, j=1,\ldots,W-1$, of a key vector \mathbf{d} can assume are restricted to the range [5, 95] and taken at a distance of 5 one from the others, to guarantee a minimum distance among the different signal decomposition lengths. With these constraints, the total number of allowed vectors \mathbf{d} is limited to $N_D=(95-5)/5+1=19$, when W=2, and to $N_D=(19\times18)/2=171$, when W=3. However, in order to be compliant with the diversity property, the actual number of transformations, which can be used in different systems, has to be further reduced.

In order to support this analysis with experimentations, a distance measure Ψ between two key vectors, namely, $\mathbf{d}^{(1)}$ and $\mathbf{d}^{(2)}$, is introduced as follows:

$$\Psi\left(\mathbf{d}^{(1)}, \mathbf{d}^{(2)}\right) = \sum_{i=1}^{W-1} \left| d_i^{(1)} - d_i^{(2)} \right|. \tag{15}$$

Considering the entire MCYT database, each user is enrolled taking into account his first E=10 signatures, to which the baseline transformation process in Section III-A is applied. Specifically, the transformations employed during enrollment are ruled by a key vector $\mathbf{d}^{(e)}$. The remaining signatures of each user are employed to estimate the FRR, after being transformed according to key vectors $\mathbf{d}^{(a)}$ which are identical to those employed during enrollment $(\mathbf{d}^{(a)} = \mathbf{d}^{(e)})$ and $\Psi(\mathbf{d}^{(e)}, \mathbf{d}^{(a)}) = 0$. Moreover, the FAR related to skilled and random forgeries is computed by transforming the available signatures according

to key vectors $\mathbf{d}^{(a)}$ which are the same of those employed during enrollment $(\mathbf{d}^{(a)} = \mathbf{d}^{(e)})$ and $\Psi(\mathbf{d}^{(e)}, \mathbf{d}^{(a)}) = 0$. The obtained ROC curves are shown in Fig. 6(a). Additionally, in order to evaluate the renewability capacity of the proposed approach, the genuine signatures of each user are transformed according to key vectors $\mathbf{d}^{(a)}$ having a distance $\Psi(\mathbf{d}^{(e)}, \mathbf{d}^{(a)}) \in$ $\{15, 20, 25, 30\}$ from the ones employed during enrollment. The obtained templates are then matched against those stored during enrollment, and the resulting matching statistics are indicated as renewable template matching rate (RTMR) in Fig. 6(a), where they are plotted versus the obtained FRR. It is worth pointing out that, in order to properly satisfy the diversity property, different templates, generated from the same data but using different keys, should not match between themselves. This means that transformed templates generated from the same signature should behave like signatures produced by different users. The diversity requirement is then fulfilled when the pseudo-ROC curves related to the RTMR are close to the ROC curves regarding the FAR obtained when considering random forgeries. As shown in Fig. 6(a), the desired condition can be met only for key vector distances $\Psi(\mathbf{d}^{(e)}, \mathbf{d}^{(a)}) > 30$. This implies that a maximum number of $\Gamma = \lfloor (95-5)/30 \rfloor + 1 = 4$ different key vectors d can be properly considered in a template protection scheme. The obtained results show that the available key space, for a system employing the baseline approach described in Section III-A, is very small and therefore not suitable for real-world signature verification systems. The extended approaches presented in Section III-B provide a higher dimensionality key space, being thus more suitable for the system deployment in real-world applications.

B. Mixing Approach

In Section III-B1, it has been shown how to transform an original signature employing two transformation keys: the decomposition vector **d**, used to define the decomposition points, and the scrambling matrix **C**, which defines the original functions whose selected segments generate the transformed sequences, according to (6). In order to evaluate the renewability capacity of the approach described in Section III-B1, the maximum number of scrambling matrices which can be

properly employed to transform the original signature representations, while keeping fixed the decomposition vector d, will be estimated. As defined in Section III-B1, a scrambling matrix C consists of F rows and W columns. The total number of matrices which can be defined is then equal to $(F!)^{(W-1)}$, which corresponds to 14! = 87178291200 when considering F = 14 and W = 2. However, among all the possible scrambling matrices, only those which allow fulfilling the diversity property can be employed.

Given two generic matrices $C^{(1)}$ and $C^{(2)}$, let us define the distance

$$\Omega\left(\mathbf{C}^{(1)}, \mathbf{C}^{(2)}\right)$$
 = number of different rows between $\mathbf{C}^{(1)}$ and $\mathbf{C}^{(2)}$. (16)

Following the approach illustrated in Section III-B1, two transformations obtained by using the same decomposition vector d, while employing two distinct scrambling matrices $C^{(1)}$ and $C^{(2)}$, produce more distinct templates as the distance $\Omega(\mathbf{C}^{(1)}, \mathbf{C}^{(2)})$ increases. Considering the entire MCYT database, each user is then enrolled by using his first E = 10 signatures, to which the transformation process in Section III-B1 is applied. Specifically, the transformations employed during enrollment are ruled by a decomposition vector d and a scrambling key matrix $C^{(e)}$. The remaining signatures of each user, after being transformed using the same keys d and $C^{(e)}$ applied during enrollment, are employed to estimate the FRR. Moreover, the FAR related to skilled and random forgeries is computed by transforming the available signature according to the decomposition vector **d** and to the same scrambling matrix $\mathbf{C}^{(a)} = \mathbf{C}^{(e)}$ employed during enrollment $(\Omega(\mathbf{C}^{(e)}, \mathbf{C}^{(a)}) =$ 0). The RTMR related to the use of the mixing approach is computed by transforming the genuine signatures of each user according to the same decomposition key d employed during enrollment, but with different scrambling keys $C^{(a)}$, characterized by distances $\Omega(\mathbf{C}^{(e)}, \mathbf{C}^{(a)}) \in \{8, 9, 10, 11\}$ from $\mathbf{C}^{(e)}$.

The matching statistics obtained for a system with E=10 and W=2 are reported in Fig. 6(b). Specifically, the renewability property of the mixing approach is verified by comparing the ROC curve where the FAR for random forgeries is taken into account with the pseudo-ROC curves where the RTMR for different distances $\Omega(\mathbf{C}^{(e)},\mathbf{C}^{(a)})$ is considered. The obtained performances show that the use of different scrambling matrices between enrollment and authentication, when keeping fixed the decomposition keys, allows obtaining matching rates which are similar to those associated with the use of random forgeries, but only when $\Omega(\mathbf{C}^{(e)},\mathbf{C}^{(a)}) \geq \Xi = 11$ (over F=14 considered functions).

Therefore, the total number of scrambling matrices which can be considered still satisfying the diversity property, guaranteed by a distance $\Omega(\mathbf{C}^{(e)},\mathbf{C}^{(a)}) \geq \Xi = 11$, has an upper bound that is equal to $(F!/(\Xi-1)!) = 24\,024$. Moreover, keeping in mind that, as explained in Section IX-A, $\Gamma=4$ distinct decomposition vectors can be defined for each scrambling matrix \mathbf{C} , the total number of renewable templates which can be properly generated, following the approach in Section III-B1 with W=2, is $4\cdot24\,024=96\,096$.

C. Shifting Approach

In this section, we verify how the renewability property of the baseline approach in Section III-A is improved when using the method described in Section III-B2, which employs a decomposition vector **d** and a shifting parameter ϕ as transformation keys. Following an approach that is similar to the one employed in Sections IX-A and B, each user available in the entire MCYT database is enrolled by using his first E = 10 signatures, which are then transformed according to the transformation keys d and $\phi^{(e)}$. Then, the remaining genuine signatures of each user are transformed using the same decomposition key d employed during enrollment, but with a different initial shift, indicated as $\phi^{(a)}$, to determine the RTMR that is used to analyze the renewability capacity of this approach. The values of the shifts are taken in the range between 0 and 95, considering only multiples of five: In this way, 20 different possible values are taken into account. Having defined a distance between the shifting parameters taken during enrollment and verification as

$$\Phi\left(\phi^{(e)}, \phi^{(a)}\right) = \left|\phi^{(e)} - \phi^{(a)}\right|,\tag{17}$$

Fig. 6(c) shows the RTMR statistics obtained by considering the same decomposition keys during enrollment and verification, at an increasing distance $\Phi(\phi^{(e)}, \phi^{(a)})$ between the employed shifting parameters. A comparison with the FAR performances obtained considering skilled and random forgeries, transformed with the same transformation keys d and $\phi^{(e)}$ employed in enrollment, is also given. The obtained experimental results show that the RTMR pseudo-ROC curves, related to the use of different shifting parameters for the enrollment and the authentication stage, are similar to the ROC curve obtained when random forgeries are taken into account when the distance $\Phi(\phi^{(e)},\phi^{(a)})$ is equal or greater than the 20% of the signature length N. This implies that the number of values ϕ which can be properly considered is limited to $\Upsilon = 5$. Applying the modification described in Section III-B2 to the baseline approach in Section III-A, we obtain an increase of the number of templates that can be generated by a factor of five, thus obtaining a number of $\Gamma \cdot \Upsilon = 4 \cdot 5 = 20$ templates. Obviously, this number is still too small for a practical application. However, if the considered modification is applied in conjunction with the method described in Section III-B1, it is possible to properly produce renewable templates with an upper limit of $\Gamma(F!/(\Xi -$ 1)) $\Upsilon = 96\,096 \cdot 5 = 480\,480$ discriminable templates.

In conclusion, although with the proposed approaches, it is not possible to obtain an infinite number of discriminable templates, almost 500 000 templates can be generated from a single original signature, properly fulfilling the diversity requirement. It is also worth pointing out that, having the possibility of managing almost 500 000 different templates, a user could issue a new biometric template each hour, for 60 years.

X. DISCUSSION AND CONCLUSION

The security and privacy issues probably represent the most important problems that have to be tackled during the design of a biometric-based automatic recognition system. In this paper, we have proposed template protection methods which can be applied to any biometrics represented by a sequence. The basic idea of the proposed BioConvolving protection approach relies on the use of a convolution-based noninvertible transformation, applied to the segments in which a sequence is split according to a transformation key. The security of our approaches relies on the difficulty in solving a blind deconvolution problem. A baseline approach, together with two extended versions of the baseline method, has been introduced.

As a proof of concept, the proposed protection approaches have been applied to an on-line signature-based authentication system, where HMMs are employed for template matching. An analysis of the security and renewability properties of the proposed methods has been extensively carried out. Specifically, the following can be observed.

- The baseline protection approach presented in Section III-A introduces only a slight loss of performance in terms of EER, with respect to an unprotected system. Moreover, the authentication performances achievable with the protected system present a slight dependence on the transformation parameters.
- 2) The authentication performance obtained using the mixing approach in Section IX-B is similar to the one achievable with the baseline method. On the other hand, the incorporation of the shifting approach in Section IX-C leads to a small degradation in the recognition rates.
- 3) The baseline method has not enough renewability capability to be used in a practical protected on-line signature-based authentication system. On the other hand, the mixing method has been shown to properly satisfy the highlighted desirable properties to protect the user's privacy. Therefore, it can be deployed in a real-world scenario.
- 4) The shifting approach can be applied to significantly increase the renewability of the mixing approach, at the cost of a small degradation in the recognition performance.

Finally, we would like to stress that the proposed BioConvolving approach can be applied to any other biometrics for which a sequence-based representation is feasible. Moreover, being able to provide a score as output of the recognition process, the proposed methods can be employed in order to construct protected multibiometrics systems, where score-level fusion is used to combine different biometric modalities, while keeping secret the original biometric data.

REFERENCES

- A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 4–20, Jan. 2004.
- [2] S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: Security and privacy concerns," *IEEE Security Privacy*, vol. 1, no. 2, pp. 33–42, Mar./Apr. 2003.
- [3] N. Ratha, J. H. Connell, and R. M. Bolle, "An analysis of minutiae matching strength," in *Proc. Int. Conf. Audio- Video-Based Biometric Person Authentication*, Halmstad, Sweden, Jun. 2001, pp. 223–228.
- [4] A. Humm, J. Hennebert, and R. Ingold, "Combined handwriting and speech modalities for user authentication," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 39, no. 1, pp. 25–35, Jan. 2009.
- [5] J. Fierrez, J. Ortega-Garcia, D. Ramos, and J. Gonzalez-Rodriguez, "HMM-based on-line signature verification: Feature extraction and sig-

- nature modeling," *Pattern Recognit. Lett.*, vol. 28, no. 16, pp. 2325–2334, Dec. 2007.
- [6] I. Bouchrika and M. S. Nixon, Model-Based Feature Extraction for Gait Analysis and Recognition. Berlin, Germany: Springer-Verlag, Jun. 2007, pp. 150–160.
- [7] R. Palaniappan and D. P. Mandic, "Biometrics from brain electrical activity: A machine learning approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 738–742, Apr. 2007.
- [8] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Local intensity variation analysis for iris recognition," *Pattern Recognit.*, vol. 37, no. 6, pp. 1287–1298, Jun. 2004.
- [9] C. Vielhauer, R. Steinmetz, and A. Mayerhöfer, "Biometric hash based on statistical features of online signatures," in *Proc. ICPR*, 2002, vol. 1, pp. 123–126.
- [10] E. Maiorana, M. Martinez-Diaz, P. Campisi, J. Ortega-Garcia, and A. Neri, "Template protection for HMM-based on-line signature authentication," in *Proc. Workshop Biometrics CVPR Conf.*, Anchorage, AK, Jun. 2008, pp. 1–6.
- [11] E. Maiorana, P. Campisi, J. Ortega-Garcia, and A. Neri, "Cancelable biometrics for HMM-based signature recognition," in *Proc. IEEE 2nd Int. Conf. BTAS*, Washington, DC, 2008, pp. 1–6.
- [12] A. Menezes, P. van Oorschot, and S. Vanstone, Handbook of Applied Cryptography. Boca Raton, FL: CRC Press, 1996.
- [13] I. Cox, M. Miller, J. Bloom, M. Miller, and J. Fridrich, *Digital Water-marking and Steganography*. San Mateo, CA: Morgan Kaufmann, 2007.
- [14] A. K. Jain and U. Uludag, "Hiding biometric data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 11, pp. 1494–1498, Nov. 2003.
 [15] E. Maiorana, P. Campisi, and A. Neri, "Biometric signature authentication
- [15] E. Maiorana, P. Campisi, and A. Neri, "Biometric signature authentication using Radon transform-based watermarking techniques," in *Proc. IEEE Biometric Symp.*, Baltimore, MD, Sep. 2007, pp. 1–6.
- [16] N. K. Ratha, J. H. Connell, and R. Bolle, "Enhancing security and privacy of biometric-based authentication systems," *IBM Syst. J.*, vol. 40, no. 3, pp. 614–634, Mar. 2001.
- [17] A. Adler, "Can images be regenerated from biometric templates?" in *Biometrics Consortium Conference*, Arlington, VA, USA, Sep. 2003.
- [18] R. Cappelli, A. Lumini, D. Maio, and D. Maltoni, "Fingerprint image reconstruction from standard templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 9, pp. 1489–1503, Sep. 2007.
- [19] K. H. Cheung, "Use of intelligent system techniques for storage and retrieval of biometrics data with application to personal identification," Ph.D. dissertation, Hong Kong Polytechnic Univ., Hong Kong, 2005.
- [20] A. K. Jain, K. Nandakumar, and A. Nagar, "Biometric template security," EURASIP J. Adv. Signal Process., vol. 2008, p. 113, Jan. 2008.
- [21] U. Uludag, S. Pankanti, and A. K. Jain, "Biometric cryptosystems: Issues and challenges," *Proc. IEEE*, vol. 92, no. 6, pp. 948–960, Jun. 2004.
- [22] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *Proc. 6th ACM Conf. Comput. Commun. Security*, Singapore, Nov. 1999, pp. 28–36.
- [23] A. Juels and M. Sudan, "A fuzzy vault scheme," *Designs Codes Cryptogr.*, vol. 38, no. 2, pp. 237–257, Feb. 2006.
- [24] K. Nandakumar, A. K. Jain, and S. Pankanti, "Fingerprint-based fuzzy vault: Implementation and performance," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 4, pp. 744–757, Dec. 2007.
- [25] M. Van der Veen, T. Kevenaar, G.-J. Schrijen, T. H. Akkermans, and F. Zuo, "Face biometrics with renewable templates," in *Proc. SPIE Secu*rity, Steganography, Watermarking Multimed. Contents, 2006, vol. 6072.
- [26] W. J. Scheirer and T. E. Boult, "Cracking fuzzy vault and biometric encryption," in *Proc. IEEE Biometric Symp.*, 2007, pp. 1–6.
- [27] Y. Sutcu, Q. Li, and N. Memon, "Protecting biometric templates with sketch: Theory and practice," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 503–512, Sep. 2007.
- [28] A. B. J. Teoh, D. C. L. Ngo, and A. Goh, "Random multispace quantization as an analytic mechanism for BioHashing of biometric and random identity inputs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 1892–1901, Dec. 2006.
- [29] M. Savvides, B. V. K. Vijaya Kumar, and P. K. Khosla, "Cancelable biometric filters for face recognition," in *Proc. Int. Conf. Pattern Recog.*, 2004, pp. 922–925.
- [30] A. A. Ross, K. Nandakumar, and A. K. Jain, Handbook of Multibiometrics. Berlin, Germany: Springer-Verlag, 2006.
- [31] V. Chatzis, A. G. Bors, and I. Pitas, "Multimodal decision-level fusion for person authentication," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 29, no. 6, pp. 674–680, Nov. 1999.
- [32] R. M. Bolle, J. H. Connell, and K. K. Ratha, "Biometrics perils and patches," *Pattern Recognit.*, vol. 35, no. 12, pp. 2727–2738, Dec. 2002.
- [33] R. Ang, R. Safavi-Naini, and L. McAven, "Cancelable key-based finger-print templates," in *Proc. 10th Australian Conf. Inf. Security Privacy*, Jul. 2005, pp. 242–252.

- [34] N. Ratha, S. Chikkerur, J. H. Connell, and R. M. Bolle, "Generating cancelable fingerprint templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 561–572, Apr. 2007.
- [35] T. E. Boult, W. J. Schreirer, and R. Woodworth, "Revocable fingerprint biotokens: Accuracy and security analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2007, pp. 17–22.
- [36] F. Quan, S. Fei, C. Anni, and Z. Feifei, "Cracking cancelable fingerprint template of Ratha," in *Proc. ISCSCT*, Dec. 2008, vol. 2, pp. 572–575.
- [37] S. Chikkerur, N. K. Ratha, J. H. Connell, and R. M. Bolle, "Generating registration-free cancelable fingerprint templates," in *Proc. IEEE BTAS Conf.*, Sep. 29–Oct. 1, 2008, pp. 1–6.
- Conf., Sep. 29–Oct. 1, 2008, pp. 1–6.
 [38] W. Xu, Q. He, Y. Li, and T. Li, "Cancelable voiceprint templates based on knowledge signatures," in *Proc. Int. Symp. Electron. Commerce Security*, Aug. 2008, pp. 412–415.
- [39] A. Cichocki and S. Amari, Adaptive Blind Signal and Image Processing. New York: Wiley, 2002.
- [40] P. Campisi and K. Egiazarian, Blind Image Deconvolution: Theory and Applications. Boca Raton, FL: CRC Press, 2007.
- [41] Y. He, K. H. Yap, L. Chen, and L. P. Chau, "A novel hybrid model framework to blind color image deconvolution," *IEEE Trans.* Syst., Man, Cybern. A, Syst., Humans, vol. 38, no. 4, pp. 867–880, Jul. 2008.
- [42] F. Leclerc and R. Plamondon, "Automatic signature verification: The state of the art 1989–1993," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 8, no. 3, pp. 643–659, 1994.
- [43] G. Dimauro, S. Impedovo, M. G. Lucchese, R. Modugno, and G. Pirlo, "Recent advancements in automatic signature verification," in *Proc. 9th Int. Workshop Frontiers Handwriting Recog.*, Oct. 26–29, 2004, pp. 179–184.
- [44] J. Fierrez and J. Ortega-Garcia, "On-line signature verification," in *Handbook of Biometrics*, A. K. Jain, A. Ross, and P. Flynn, Eds. Berlin, Germany: Springer-Verlag, 2008, pp. 189–209.
- [45] C. Vielhauer and R. Steinmetz, "Handwriting: Feature correlation analysis for biometric hashes," *EURASIP J. Appl. Signal Process.*, vol. 2004, no. 4, pp. 542–558, Jan. 2004.
- [46] A. Kholmatov and B. Yanikoglu, "Identity authentication using improved online signature verification method," *Pattern Recognit. Lett.*, vol. 26, no. 15, pp. 2400–2408, Nov. 2005.
- [47] L. Yang, B. W. Widjaja, and R. Prasad, "Application of hidden Markov models for signature verification," *Pattern Recognit.*, vol. 28, no. 2, pp. 161–170, 1995.
- [48] H. Feng and C. W. Chan, "Private key generation from on-line handwritten signatures," *Inf. Manage. Comput. Security*, vol. 10, no. 4, pp. 159–164, 2002
- [49] M. Freire-Santos, J. Fierrez-Aguilar, and J. Ortega-Garcia, "Crypto-graphic key generation using handwritten signature," in *Proc. SPIE*, *Defense Security Symp.*, *Biometric Technol. Human Identification*, 2006, vol. 6202, pp. 225–231.
- [50] W. K. Yip, A. Goh, D. C. L. Ngo, and A. B. J. Teoh, "Generation of replaceable cryptographic keys from dynamic handwritten signatures," in *Proc. ICB*, 2006, pp. 509–515.
- [51] E. Maiorana, P. Campisi, and A. Neri, "User adaptive fuzzy commitment for signature templates protection and renewability," *J. Electron. Imag.*, vol. 17, no. 1, p. 011 011, Mar. 2008.
- [52] P. Campisi, E. Maiorana, and A. Neri, "On-line signature based authentication: Template security issues and countermeasures," in *Biometrics: Theory, Methods, and Applications*, N. V. Boulgouris, K. N. Plataniotis, and E. Micheli-Tzanakou, Eds. New York: Wiley, 2009.
- [53] E. Maiorana, P. Campisi, and A. Neri, "Multi-level signature based biometric authentication using watermarking," in *Proc. SPIE Defense Security, Mobile Multimedia/Image Process. Military Security Appl.*, 2007, vol. 6579, pp. 657 90J.1–657 90J.12.
- [54] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989
- [55] B. L. Van, S. Garcia Salicetti, and B. Dorizzi, "On using the Viterbi path along with HMM likelihood information for online signature verification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 5, pp. 1237–1247, Oct. 2007.
- [56] J. Hennebert, R. Loeffel, A. Humm, and R. Ingold, "A new forgery scenario based on regaining dynamics of signature," in *Proc. IEEE ICB*, 2007, pp. 366–375.
- [57] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J.-J. Igarza, C. Vivaracho, D. Escudero, and Q.-I. Moro, "MCYT baseline corpus: A bimodal biometric database," *Proc. Inst. Elect. Eng.—Vision, Image Signal Process.*, vol. 150, no. 6, pp. 395–401, Dec. 2003.



Emanuele Maiorana (S'06–M'08) received the Laurea degree (*summa cum laude*) in electronic engineering and the Ph.D. degree in telecommunication engineering from Università degli Studi "Roma Tre," Roma, Italy, in 2004 and 2009, respectively.

From September 2004 to November 2005, he was with the Communications and High Tech Workgroup, Accenture Consulting Workforce. From October 2007 to March 2008, he was a Visiting Researcher with the Biometric Recognition Group, ATVS, Universidad Autonoma de Madrid, Madrid, Spain. He is

currently a Postdoctoral Researcher with Università degli Studi Roma Tre. His research interests are in the areas of digital signals, image processing, textures, biometrics, and security of telecommunication systems.

Dr. Maiorana is the recipient of the Lockheed Martin Best Paper Award for the Poster Track at the 2007 IEEE Biometrics Symposium and of the Honeywell Student Best Paper Award at the 2008 IEEE Second International Conference on Biometrics: Theory, Applications and Systems.



Patrizio Campisi (M'00–SM'08) received the Ph.D. degree in electronic engineering from Università degli Studi "Roma Tre," Roma, Italy.

He held invited visiting positions at the University of Toronto, Toronto, ON, Canada, in 2000, at the Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana–Champaign, Urbana, in 2003, at Ecole Polytechnique de l'Università de Nantes, Nantes, France, in 2006, 2007, 2009, and 2010, and at Universidade de Vigo, Vigo, Spain, in 2010. He is currently an Associate

Professor with the Dipartimento di Elettronica Applicata, Università degli Studi Roma Tre. He is currently involved in many European Union (EU) projects on biometrics. His research interests include secure biometric authentication, secure multimedia communications, data hiding for digital right management and quality assessment, blind image deconvolution, and data equalization. He is the Coeditor of the book entitled *Blind Image Deconvolution: Theory and Applications* (CRC Press, 2007).

Dr. Campisi is a corecipient of a best student paper award at the 2006 IEEE International Conference on Image Processing and at the 2008 IEEE Second International Conference on Biometrics: Theory, Applications and Systems and of a best paper award at the 2007 IEEE Biometrics Symposium. He is the General Chair of the 12th ACM Workshop on Multimedia and Security on September 9-10, 2010, in Roma. He is the Italian delegate for the European COST 2101 Action "Biometrics for Identity Documents and Smart Cards" (2006-2010) and the Coordinator, at Università degli Studi Roma Tre, of the EU-FP7 thematic network "Biometrics European Stakeholders Thematic Network" (October 2009-October 2011). He is a member of IEEE's Certified Biometrics Professional Learning System Committee and of the IEEE Systems, Man, and Cybernetics Society Technical Committee on Information Assurance and Intelligent Multimedia-Mobile Communications (2007-present). He is an Associate Editor of IEEE SIGNAL PROCESSING LETTERS (December 2008-December 2010), of the International Journal of Digital Crime and Forensics (January 2009-present), and of Advances in Multimedia by Hindawi (January 2009-present).



Julian Fierrez received the M.Sc. and Ph.D. degrees in electrical engineering from Universidad Politecnica de Madrid, Madrid, Spain, in 2001 and 2006, respectively.

Since 2002, he has been with Universidad Autonoma de Madrid, Madrid, where he currently holds a Marie Curie Postdoctoral Fellowship, part of which has been spent as a Visiting Researcher at Michigan State University, East Lansing. His research interests include signal and image processing, pattern recognition, and biometrics.

Dr. Fierrez was the recipient of the Best Poster Award at the 2003 Audioand Video-Based Biometric Person Authentication Conference, of the Rosina Ribalta Award to the best Spanish Ph.D. proposal in 2005, of the Motorola Best Student Paper at the 2006 International Conference on Biometrics, and of the 2006 EBF European Biometric Industry Award.



conference papers.

Javier Ortega-Garcia (M'96-SM'08) received the Ph.D. degree (cum laude) in electrical engineering from Universidad Politecnica de Madrid, Madrid, Spain, in 1996.

He is the Founder and Codirector of the Biometric Recognition Group, ATVS. He is currently a Full Professor with Escuela Politecnica Superior, Universidad Autonoma de Madrid, Madrid. His research interests are focused on biometrics signal processing. He has published over 150 international contribu-

tions, including book chapters, refereed journal, and Dr. Ortega-Garcia has chaired "Odyssey 2004: The Speaker and Language Recognition Workshop," cosponsored by ISCA and IEEE, and cochaired

"ICB 2009," the Third IAPR International Conference on Biometrics.



Alessandro Neri (M'82) received the Ph.D. degree in electronic engineering from the University of Rome "La Sapienza," Rome, Italy, in 1977.

In 1978, he joined the Research and Development Department, Contraves Italiana S.p.A., where he gained expertise in the fields of radar signal processing and applied detection and estimation theory, becoming the Chief of the Advanced Systems Group. In 1987, he joined the INFOCOM Department, University of Rome La Sapienza, as an Associate Professor of signal and information theory.

In November 1992, he joined the Department of Electronic Engineering, Università degli Studi "Roma Tre," Rome, as an Associate Professor of electrical communications, where he became a Full Professor of telecommunications in September 2001. Since 1992, he has been responsible for the coordination and management of research and teaching activities in the telecommunication field with Università degli Studi Roma Tre. Since 1998, he has also been responsible for planning and designing activities related to university campus telecommunication systems and services. His research activity has mainly been focused on information theory, signal theory, and signal and image processing and their applications to telecommunication systems, remote sensing, and biometrics. His current research is focused on third- and fourth-generation cellular systems and multimedia communications.