Multimodal authentication based on random projections and distributed coding

Voloshynovskyy, Svyatoslav; Koval, Oleksiy; Pun, Thierry

Multimodal Authentication Based on Random Projections and Source Coding

Sviatoslav Voloshynovskiy
CVML-SIP, University of Geneva
7, route de Drize,
1227, Geneva, Switzerland
svolos@cui.unige.ch

Oleksiy Koval
CVML-SIP, University of Geneva
7, route de Drize,
1227, Geneva, Switzerland
koval@cui.unige.ch

Thierry Pun
CVML-SIP, University of Geneva
7, route de Drize,
1227, Geneva, Switzerland
pun@cui.unige.ch

ABSTRACT

In this paper, we consider an authentication framework for independent modalities based on binary hypothesis testing using source coding jointly with the random projections. The source coding ensures the multimodal signals reconstruction at the decoder based on the authentication data. The random projections are used to cope with the security, privacy, robustness and complexity issues. Finally, the authentication performance is investigated for both direct and random projections domains. The asymptotic performance approximation is derived and compared with the exact solutions. The impact of modality fusion on the authentication system performance is demonstrated.

Categories and Subject Descriptors

I.4.7 [**Feature Measurement**]: [feature representation, projections]

General Terms

Performance, Security, Verification.

1. INTRODUCTION

Recently, reproduction technologies have performed an impressive evolution allowing to reproduce not only two-dimensional (2D) graphics but also three-dimensional (3D) relief structures, synthesize voice, images, 3D shapes and even recreate human facial expressions. Besides the obvious advantages, these tools offer at the same time unprecedent possibilities for the people targeting illegal actions ranging from the gaining access to various services, facilities and infrastructures to the counterfeiting of legal documents, certificates, IDs or even physical items such as luxury goods, art objects etc.

MM-SEC'08, September 22-23, 2008, Oxford, UK.

The classical protection mechanisms are based either on proprietary techniques (material science), which use added features that are either rare in nature or difficult to clone, or shared secret (password, pin, key) or accompanying ID documents that all together do not completely satisfy the requirements to the reliability and security. The proprietary technology secrets are often relatively quickly become outdated due to the technological development, can be disclosed, guessed, reversely engineered or maliciously intercepted in various protocols while the ID documents can be stolen, lost or sometimes relatively easily faked. The attempts to apply classical crypto-based techniques to the protection of humans and physical items fail due to the inherently noisy nature of measurements.

That is why reliable and secure authentication of humans and physical objects preserving the privacy is an emerging and challenging problem for various applications. Fortunately, both humans and physical items possess forensic features that are a sort of unique fingerprints, which exist in a single copy, can not be easily reproduced or cloned using even the most advanced existing or envisaged future equipment of the counterfeiters or the genuine manufacturers. The forensic features are formed by either nature or during involved manufacturing processes and have essentially random character. These features can be used for the authentication purposes. A particular selection of forensic features depends on various factors that include their *universality* (do all humans or items have them?), *distinctiveness* (can all humans/items be distinguished based on them?), *permanence* (how permanent are the features?), *collectability* (how well can the features be captured and quantified?), etc.

Unfortunately, not all forensic features perfectly meet the above requirements [1]. Most of the forensic features are facing the fundamental trade-off problem between the performance and security. A possible solution to this trade-off is based on the multimodal authentication that recalls the necessity to use several forensic features of the same object for reliable and secure authentication [2].

At the same time, the presence of multiple features creates a freedom for their fusion at various structural levels: *sensor level*, *feature level*, *match score level*, *rank level* and *decision level*. Besides the existing multibiometric authentication and identification system designs where fusions are mostly performed at match score or decision levels [3, 4], it is evident that such an approach is suboptimal in terms of the

attained performance accuracy due to the data processing inequality [5] that suggests to combine the available biometric data at sensor level. However, due to various practical constraints, not so many cases are known where the fusion is performed optimally from an information-theoretic perspective [6]. Therefore, it is important to establish the theoretical limits on performance under optimal feature fusion.

Moreover, most of biometrics features belonging to the same person are statistically independent considering them from the signal processing perspectives. For example, the samples of facial photo, iris, fingerprint and voice of person pronouncing his name shown in Figure 1 can be considered to be statistically independent.

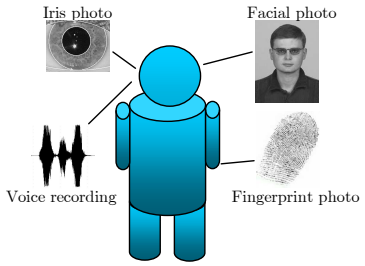


Figure 1: Example of biometrics belonging to the same person: facial photo, fingerprint, iris and voice.

Similarly, the statistical features of various items are unique. Moreover, several samples taken from the same item in different locations using even the same acquisition technique might also demonstrate the statistical independence. An example of document authentication based on microstructure images of document carrier (paper) scanned at 1200 dpi is shown in Figure 2. These samples can be considered as different modalities representing the same physical item. Obviously, other forensic features potentially acquired using different imaging techniques can be considered for the fusion.

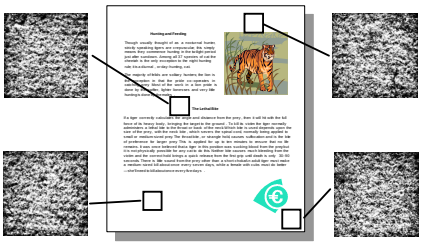


Figure 2: Example of microstructure image of paper acquired at 1200 dpi by an optical scanner.

That is why the problem of optimal fusion of multiple modalities is of great importance for both human and item authentication. Moreover, quite often the multimodal features are of high dimensionality and the dimensionality reduction techniques are applied to reduce the complexity and memory storage requirements. Additionally, the dimensionality reduction is considered in the scope of uninvertible transformation for the security/privacy enhancement known as *cancelable biometrics* [7]. However, it is not often clearly understood the impact of this transformation on both performance and security. Finally, both fusion at different consequent levels and dimensionality reduction unavoidably lead to the loss of information due to the above mentioned data processing inequality and the special care should be taken about the optimal design of multimodal authentication system with the reduced size or compressed features.

Besides the different nature and applications, the authentication of humans and physical items have a lot in common and in the following we will refer to it as a generic authentication of objects, where under the object we will understand the humans and physical items. From the technical point of view, the main challenge consists in providing reliable authentication based on noisy multimodal observations that are different from those acquired at the enrollment stage. Obviously, the traditional cryptography-based authentication will produce a negative result even if a single bit is altered that is not suitable for this protocol. Additionally, the security leakages about the authentication protocol might cause an appearance of a number of attacks targeting to trick the authentication (including impersonation and physical attacks). To resolve these robustness-security requirements, we will use the hypothesis testing framework for the evaluation of object authenticity [8]. Therefore, we will compare the performance of unimodal authentication system with the multimodal one in terms of probability of false acceptance P_F and probability of correct acceptance (detection) P_D that form a receiver operation characteristic (ROC), establish the impact of fusion of noisy modalities on the ROC and evaluate the loss in performance due to the dimensionality reduction.

The paper has the following structure. The generic unimodal authentication problem is considered in Section 2 based on communication framework. The multimodal authentication in the direct domain is presented in Section 3 while in the random projections domain is provided in Section 4. The results of computer simulation are given in Section 5. Finally, Section 6 concludes the paper.

Notations We use capital letters to denote scalar random variables X and $\mathbf{X}$ to denote vector random variables, corresponding small letters x and $\mathbf{x}$ to denote the realizations of scalar and vector random variables, respectively. All vectors without sign tilde are assumed to be of the length N and with the sign tilde of length L with the corresponding subindexes. Calligraphic fonts $\mathcal{X}$ denote sets $X \in \mathcal{X}$ and $|\mathcal{X}|$ denotes the cardinality of set $\mathcal{X}$. We use $\mathbf{X} \sim p_{\mathbf{X}}(\mathbf{x})$ or simply $\mathbf{X} \sim p(\mathbf{x})$ to indicate that a random variable $\mathbf{X}$ is distributed according to $p_{\mathbf{X}}(\mathbf{x})$. The statistical hypothesis are denoted as H_i, $i = \{0, 1\}$ and the distributions under these hypothesis as $p(\mathbf{y}|H_i)$. $\mathcal{N}(\mu, \sigma_X^2)$ stands for Gaussian distribution with mean μ and variance σ_X^2. $||.||$ denotes Euclidean vector norm and $D_s(.,.)$ is Chernoff distance between distributions. The mathematical expectation of a random variable $X \sim p_X(x)$ is denoted by $E_{p_X}[X]$ or simply by $E[X]$.

2. MAIN DESIGNS OF UNIMODAL AUTHENTICATION SYSTEMS

2.1 Generic authentication problem

A generic authentication problem can be considered as a hypothesis testing [8] that requires the selection of authentication criteria and assumptions behind the statistics

of genuine and faked objects. In the most general case, the authentication problem can be considered as a decision making that the observed length-N codeword $\mathbf{v}$ representing the object under authentication is in some proximity to the genuine codeword $\mathbf{x}(m)$, $1 \leq m \leq |\mathcal{M}|$, for example specified in the Euclidean space as $\|\mathbf{x}(m) - \mathbf{v}\|^2 \leq e_m$, where e_m defines the acceptable distortions as well as the acceptance region $\mathcal{R}_1(m)$, while $\mathcal{R}_0(m)$ is considered to be the rejection region for the m^{th} codeword. This can be schematically shown as in Figure 3 for all realizations or codewords.

In the most general case, the above authentication problem can be considered as a composite hypothesis testing:

$$\begin{cases} H_0: & \mathbf{v} \sim p(\mathbf{v}|H_0), \\ H_1: & \mathbf{v} \sim p(\mathbf{v}|H_1), \end{cases} \tag{1}$$

where H_1 is a simple hypothesis with $p(\mathbf{v}|H_1) = p(\mathbf{v}|\mathbf{x}(m))$ and H_0 is a complex one for which:

$$p(\mathbf{v}|H_0) = \begin{cases} p(\mathbf{v}|\mathbf{x}(1)) & \text{with probability } p_1, \\ \dots & \dots \\ p(\mathbf{v}|\mathbf{x}(|\mathcal{M}|)) & \text{with probability } p_{|\mathcal{M}|}, \end{cases} \tag{2}$$

where $\sum_{n=1, m\neq m}^{|\mathcal{M}|} p_n = 1$ that includes all $\mathbf{x}(n)$ such that $n \neq m$, where $\mathbf{x}(m)$ is considered to be an authentic codeword corresponding to the above hypothesis H_1.

To design the decision rule for the binary hypothesis testing, we will use the worst case condition for the selection of the alternative hypothesis. We will assume that given the enrolled database of all object forensics and specified index m or equivalently $\mathbf{x}(m)$, one needs to ensure the desired ROC for the closest possible to $\mathbf{x}(m)$ codeword denoted as $\mathbf{x}(n)$, in Figure 3 among all $|\mathcal{M}|$ codewords. It should be noticed that one can also ensure the specified ROC taking into account all codewords and their corresponding probabilities of appearance according to the model (2) that will correspond to the Bayessian framework. [1] However, to avoid cumbersome integrations that reduces tractability, we will follow the worst case approach. Alternatively, one can also consider generalized maximum likelihood approach that is tractable but not always optimal [9].

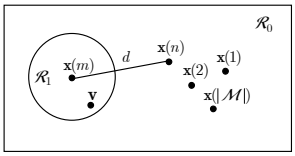


Figure 3: Authentication setup for the codeword space.

Therefore, we will reformulate the authentication problem (1)-(2) as the hypothesis testing with two simple hypothesis, i.e., H_1 corresponds to the case that the item is authentic and H_0 to the faked one, for the genuine codeword $\mathbf{x}(m)$ and its worst case counterpart $\mathbf{x}(n)$:

$$\begin{cases} H_0: & \mathbf{v} \sim p(\mathbf{v}|H_0) = p(\mathbf{v}|\mathbf{x}(n)), \\ H_1: & \mathbf{v} \sim p(\mathbf{v}|H_1) = p(\mathbf{v}|\mathbf{x}(m)). \end{cases} \tag{3}$$

[1] We assume that the attacker has access to the codebook and can choose the worst case counterpart to the codeword $\mathbf{x}(m)$ from the codebook of the enrolled codewords. We also assume that the designer of authentication system is aware of such an attacking strategy.

One can use the Neyman-Pearson decision rule that maximizes the probability of correct acceptance P_D subject to the constraint $P_F \leq \alpha$ that can be formulated as the likelihood ratio test:

$$\Lambda(\mathbf{v}) = \frac{p(\mathbf{v}|H_1)}{p(\mathbf{v}|H_0)} \lessgtr \eta, \tag{4}$$

with the threshold η chosen to satisfy $P_F = \int_{\Lambda(\mathbf{v})>\eta} p(\mathbf{v}|H_0) d\mathbf{v} = \alpha$.

Besides we will seek for a secure solution that does not leak any information to the counterfeiter. Therefore, the main challenge is to provide reliable and secure authentication based on the noisy observation $\mathbf{v}$.

2.2 Authentication architectures

The considered authentication setup is generic and presents the theoretical problem formulation for various biometric and physical item authentication systems. However, the need to store the entire codebook or database of all enrolled objects raises numerous complexity, memory storage, privacy and security concerns as discussed above. Therefore, modern authentication systems attempt to resolve these issues using recent achievements in coding theory. Before proceeding with the multimodal formulation and moving on to the core of this paper, it is important to provide a brief overview of state-of-the-art unimodal authentication systems and briefly describe how our contribution can be compared with the existing techniques. The existing unimodal authentication systems can be classified on two large groups, i.e., those based on *channel coding* and those based on *source coding* including distributed one.

The authentication system based on channel coding is schematically shown in Figure 4. Such an architecture also resembles the robust perceptual hashing and corresponding identification architecture [10, 11]. The genuine data $\mathbf{X}$ is supposed to be communicated via the channel, which consists of an active counterfeiter and a passive discrete memoryless channel (DMC) with the transition probability $p_{V|X}$. The counterfeiter is targeting to trick the authentication protocol by replacing $\mathbf{X}$ by $\mathbf{X}'$ keeping the same index m used for the genuine object to get the confirmative decision at the authentication stage according to the attacking strategies described in Section 2.1. The codeword $\mathbf{X}'$ can be considered in this sense as the worst case codeword to $\mathbf{X}$ given index m maximizing P_F. The index m is deduced at the encoder from $\mathbf{X}$ enrolled with the rate R_M such that any $\mathbf{X}$ can be represented by 2^{NR_M} sequences $\hat{\mathbf{X}}$. Obviously, to have lossless source representation, one needs to ensure the condition $R_M \geq H(X)$, where $H(X)$ is the entropy of X. However, at the same time the maximum number of uniquely distinguishable sequences that can be communicated via DMC $p_{V|X}$ is limited and determined by the identification capacity of this channel that is $C_{\text{id}} = I(X;V)$, where $I(X;V)$ is a mutual information between X and V [2] [12]. Therefore, to have the unique match of the decoded index $\hat{m}$ and the communicated index m, one should provide the conditions of reliable communications determined by $R_M \leq C_{\text{id}}$. This scheme is essentially based on the extension of classic cryptographic authentication techniques,

[2] Note, the difference with the communication capacity $C = \max_{p_X(x)} I(X;V)$, i.e., the maximization is performed with respect to the input distribution $p_X(x)$.

which use hashing. The only difference consists in the replacement of crypto-hash by a robust perceptual hash that is insensitive to the certain variations in the data. A fundamental shortcoming of this protocol is the unavoidable presence of collisions typical for hashing disregarding the chosen rate R_M due to the fact that in the most practically important cases $H(X) > C_{\text{id}}$. In any case, the decoder is uncapable to uniquely deduce the index $\hat{m}$ due to the overlapping of the decoding regions. Additional essential drawbacks of this architecture are relatively high rate R_M to be communicated, complexity of the decoder in the case of completely random codewords and the need for the database of enrolled signals and decision regions that might be either prohibitively large or represent the security leakage.

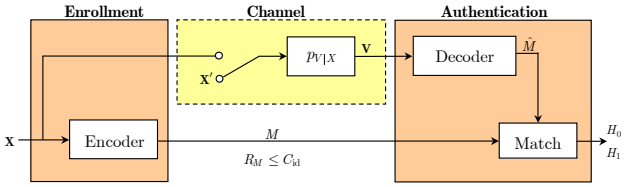


Figure 4: Authentication architecture based on channel coding (robust perceptual hashing).

The above drawbacks in part of collisions can be overcame by a scheme based on the source coding. In the scope of the source coding based authentication, the object index is deduced at the enrollment stage based on the observed data $\mathbf{x}$. We assume here that the *lossless coding* is used, where all sequences $\mathbf{x}$ of length N are generated from some distribution $p_{\mathbf{X}}(\mathbf{x})$. The encoder assigns the index m to each sequence and sends it to the decoder with the rate $R_X \geq H(X)$, where $H(X)$ is the entropy of X. At the authentication stage, one should make a decision about the item authenticity based on the observed vector $\mathbf{v}$ and the index m. For this purpose, the decoder retrieves the sequence $\hat{\mathbf{x}}(m)$ based on m, and the binary test produces the final decision by generating the hypothesis H_0, i.e., fake, or H_1, i.e., genuine. To reduce the rate for m, one can further apply *lossy source coding*. In this case, m can be considered as a hash obtained with the corresponding randomized codebook generation. However, this will cause well-known collisions. To avoid this undesirable effect and exploit the fact of $\mathbf{v}$ presence at the decoder that is correlated with the genuine $\mathbf{x}(m)$, one can use *distributed source coding* based on Slepian-Wolf framework [14]. This coding is based on binning, where m is considered as a bin index. In this case the rate can be reduced to $R_X^{SW} \geq H(X|V)$. Similar in spirit approaches were firstly introduced by Maurer [15] and Ahlswede and Csiszar [16], where the index m was considered as a helper data for *common randomness* extraction considered to be $\mathbf{x}$. Nevertheless, the above result applies to the discrete-value vectors. In the case of continuous-value vectors, one should apply first the quantization that will lead to the distortions at the reconstruction of $\hat{\mathbf{x}}$ and the corresponding collisions depending on the quantizer rate. More generally, the lossy distributed source coding can be considered based on Wyner-Ziv framework [17] using the binning similar to those used in the Slepian-Wolf coding with an auxiliary random vectors $\mathbf{U}$ constructed from $\mathbf{X}$ according to the mapping $p_{\mathbf{U}|\mathbf{X}}$. In this case, the rate-distortion function $R_X^{WZ}(D)$ is defined as:

$$R_X^{WZ}(D) = \min_{p_{U|X}: E[d^N(\mathbf{X},\hat{\mathbf{X}})] \leq D} I(X;U) - I(V;U), \qquad (5)$$

where $d^N(.,.)$ is the distortion measure between two vectors and D is the distortion. For the Gaussian setup considered in this paper, $\mathbf{X} \sim (\mathbf{0}, \sigma_X^2 \mathbf{I}_N)$ and $\mathbf{V} \sim (\mathbf{0}, \sigma_V^2 \mathbf{I}_N)$ the rate (5) turns out to be:

$$R_X^{WZ}(D) = \begin{cases} \frac{1}{2}\log_2\left(\frac{\sigma_X^2(1-\rho_{XV}^2)}{D}\right), & \text{for } D < \sigma_X^2(1-\rho_{XV}^2), \\ 0, & otherwise, \end{cases} \qquad (6)$$

where $\rho_{XV}^2 = \frac{E[XV]}{\sigma_X \sigma_V}$ is the correlation coefficient. The impact of distortions introduced by the lossy source coding can be considered as the equivalent noise in the reconstructed sequence $\hat{\mathbf{x}} = \mathbf{x} + \mathbf{z}$ according to the model of test channel from the rate-distortion theory [5] that can be taken into account in the corresponding hypothesis H_1. However, to avoid the collisions and to simplify the consideration, one can suppose that the rate $R_X^{WZ}(D)$ is chosen to be sufficiently high to guarantee the low distortion D that allows to assume that the distortions under both hypothesis H_0 and H_1 are the same.

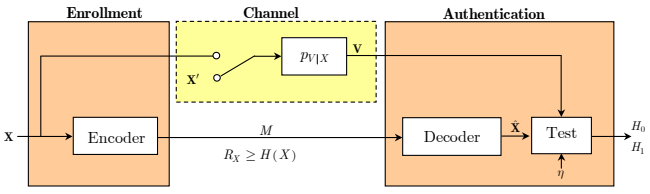
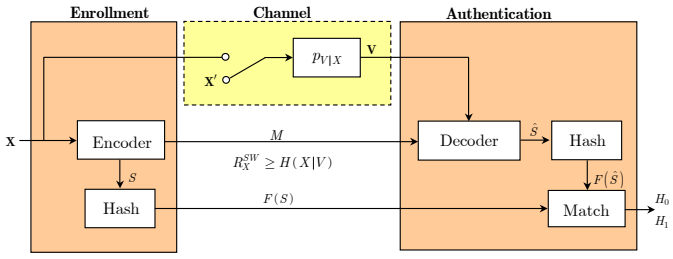


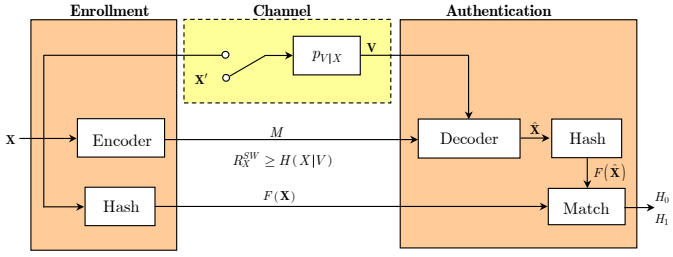
Figure 5: Authentication setup based on lossless coding.

The above framework was theoretically considered by Tyuls *et. al.* [12] and Ignatenko and Willems [18] and practically implemented by Martinian *et. al.* [19] and Lin *et. al.* [20] using low-density parity check codes (LDPC) for the distributed coding.

The main idea is to avoid soft hypothesis testing by replacing it by the direct matching of hashes extracted from the reconstructed data based on noisy measurements using common randomness extraction framework. Two possible schemes were considered in [12, 19, 20]. The first scheme uses the index s, $s = \{1, 2, \cdots, 2^{NI(X;V)}\}$, where $I(X;V)$ denotes the mutual information between X and V, of the sequence $\mathbf{x}$ in the bin m for the hashing (Figure 2.2,a) and the second one is based on the hash computed from the original $\mathbf{x}$ sequence (Figure 2.2,b). The necessity to introduce extra side information is dictated by the need to distinguish the sequences within the same bin m in the case when the informed attacker might produce a fake that will be jointly typical with $\mathbf{v}$. This will lead to the false acceptance. Thus, the hashed values of s or $\mathbf{x}$ aim protecting against such kind of attack. Obviously, the rate of the hash in the second case is higher. A common drawback of these schemes is unavoidable presence of collisions due to the hashing. Therefore, to avoid it as well as to enable the usage of biometrics of the same person in various applications, Ignatenko and Willems [18] suggested the scheme based on XORring of index s with the secret key k and its validation at the authentication stage based on the decoded version $\hat{s}$ (Figure 7).



(a)



(b)

Figure 6: Authentication setup based on distributed coding and hashing of: (a) sequence index s within the bin m and (b) original sequence x.

At the same time, the practical implementation of the above schemes was envisioned by the conversion of the continuous vectors to the discrete representations, e.g., extraction of minutia features [19], quantization of randomly projected data from non-overlapping blocks of size 16×16 [20], binarization of speckle images based on Gabor transform subbands thresholding [18] and similar transformations predicted for optical and coated physical unclonable functions [12]. All these transformations are not invertible and obviously lead to the information loss due to data processing inequality [5]. The quantitative estimation of such a loss was not reported besides the results of computer modeling. Moreover, the definition of security has also different notion for the considered biometrics authentication systems [12, 18, 19] and physical item protection. In the biometric context, it is assumed that both $\mathbf{x}$ and $\mathbf{v}$ can be only obtained from the physical person and thus are unknown for the attacker. Therefore, the security leakage sources were considered with respect to the indexes m and s and the corresponding efforts have been dedicated to protect the scheme from the direct disclose of biometric data $\mathbf{x}$ that can be exploited by the attacker for impersonation. In the item authentication application, the data $\mathbf{x}$ is inherently present for the counterfeiter and the security relies on the physical impossibility to reproduce the duplicate or clone that altogether with the index m and any assisting data might be accepted as the authentic item. Nevertheless, there are a number of crypto-based attacks that can benefit from the disclosure of the coding part of the authentication scheme to present a fake $\mathbf{x}'$ with the index m that can be falsely accepted.

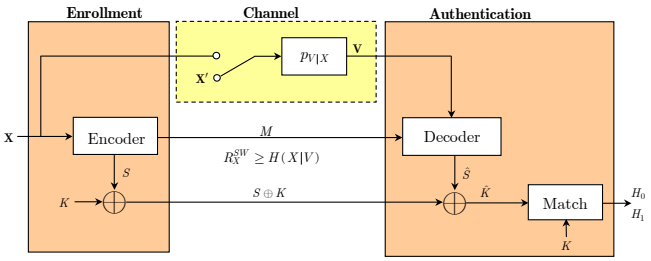


Figure 7: Authentication setup based on distributed coding and sequence index s XORing, which identifies the sequence x within the bin m.

Therefore, as the first step on the way toward the theoretical quantification of the loss due to the above feature extraction and countermeasures related to the protection of the codebook against impersonation attack, we will consider the performance of a simple lossless source coding based authentication presented in Figure 5 accompanied by a generic random projection operator $\mathbf{\Phi}$ shown in Figure 8. Such kind of projection into a secure key-defined domain besides the security insures the dimensionality reduction, complexity as well as memory storage. The transform $\mathbf{\Phi}$ produces vectors $\tilde{\mathbf{x}}(m)$ and $\tilde{\mathbf{v}}$ of dimensionality L, where $L \leq N$. Additionally, the transform can be chosen in such a way to guarantee a certain robustness to the legitimate distortions. In the rest, this protocol is similar to one from Figure 5.

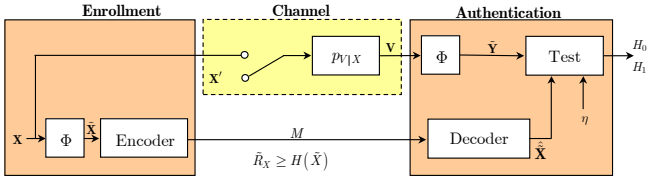


Figure 8: Authentication setup with random projections: $\mathbf{\Phi}$ is the key-based random projection operator.

3. MULTIMODAL AUTHENTICATION AS BINARY HYPOTHESIS TESTING

We will formulate the multimodal authentication problem shown in Figure 9 as the hypothesis testing with two hypothesis H_1 and H_0 considered in Section 2.1, where H_1 corresponds to the case that the object is authentic and represented by a pair of multimodal vectors $(\mathbf{x}(m), \mathbf{y}(m))$ of dimensions N_X and N_Y, respectively, and H_0 corresponds to the case of non-authentic object represented by the worst case counterpart pair $(\mathbf{x}(n), \mathbf{y}(n))$:

$$\begin{cases} H_0: & \mathbf{v} = \mathbf{x}(n) + \mathbf{z}_X, \ \mathbf{w} = \mathbf{y}(n) + \mathbf{z}_Y, \\ H_1: & \mathbf{v} = \mathbf{x}(m) + \mathbf{z}_X, \ \mathbf{w} = \mathbf{y}(m) + \mathbf{z}_Y, \end{cases} \tag{7}$$

where $\mathbf{z}_X$ and $\mathbf{z}_Y$ are the noise components in each modality that can be more generally represented as:

$$\begin{cases} H_0: & p(\mathbf{v}, \mathbf{w}|H_0) = p(\mathbf{v}|H_0)p(\mathbf{w}|H_0), \\ H_1: & p(\mathbf{v}, \mathbf{w}|H_1) = p(\mathbf{v}|H_1)p(\mathbf{w}|H_1), \end{cases} \tag{8}$$

with the assumption of independence between modalities $\mathbf{X}$ and $\mathbf{Y}$ reflected by the product of the corresponding pdfs.

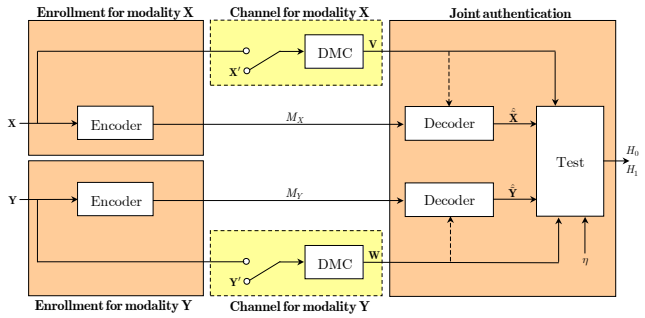


Figure 9: Multimodal authentication architecture based on source coding.

We will use the Neyman-Pearson decision rule that maximizes P_D subject to the constraint $P_F \leq \alpha$ according to the unimodal analog (4):

$$\Lambda(\mathbf{v},\mathbf{w}) = \frac{p(\mathbf{v},\mathbf{w}|H_1)}{p(\mathbf{v},\mathbf{w}|H_0)} = \frac{p(\mathbf{v}|H_1)p(\mathbf{w}|H_1)}{p(\mathbf{v}|H_0)p(\mathbf{w}|H_0)} \lessgtr \eta, \quad (9)$$

with η such that $P_F = \int_{\Lambda(\mathbf{v},\mathbf{w})>\eta} p(\mathbf{v},\mathbf{w}|H_0)d\mathbf{v}d\mathbf{w} = \alpha$.

Under the Gaussian assumption about noise $\mathbf{Z}_X \sim \mathcal{N}(\mathbf{0}, \sigma^2_{Z_X}\mathbf{I}_{N_X})$ and $\mathbf{Z}_Y \sim \mathcal{N}(\mathbf{0}, \sigma^2_{Z_Y}\mathbf{I}_{N_Y})$ [3] and known signal pairs $(\mathbf{x}(m), \mathbf{y}(m))$ and $(\mathbf{x}(n), \mathbf{y}(n))$, we have:

$$\begin{cases} H_0: & p(\mathbf{v},\mathbf{w}|H_0) = \mathcal{N}(\mathbf{x}(n), \sigma^2_{Z_X}\mathbf{I}_{N_X})\mathcal{N}(\mathbf{y}(n), \sigma^2_{Z_Y}\mathbf{I}_{N_Y}), \\ H_1: & p(\mathbf{v},\mathbf{w}|H_1) = \mathcal{N}(\mathbf{x}(m), \sigma^2_{Z_X}\mathbf{I}_{N_X})\mathcal{N}(\mathbf{y}(m), \sigma^2_{Z_Y}\mathbf{I}_{N_Y}). \end{cases} \quad (10)$$

The test (9) can be reformulated by taking the logarithm as:

$$\begin{aligned} &[\log p(\mathbf{v}|H_1) - \log p(\mathbf{v}|H_0)] \\ &+ [\log p(\mathbf{w}|H_1) - \log p(\mathbf{w}|H_0)] \lessgtr \log\eta, \end{aligned} \quad (11)$$

that can be reduced to the sufficient statistic t:

$$\begin{aligned} t(\mathbf{v},\mathbf{w}) &:= \frac{1}{\sigma^2_{Z_X}}\left[\mathbf{v}^T(\mathbf{x}(m)-\mathbf{x}(n)) - \frac{1}{2}(\epsilon_X(m)-\epsilon_X(n))\right] \\ &+ \frac{1}{\sigma^2_{Z_Y}}\left[\mathbf{w}^T(\mathbf{y}(m)-\mathbf{y}(n)) - \frac{1}{2}(\epsilon_Y(m)-\epsilon_Y(n))\right] \lessgtr \gamma, \end{aligned} \quad (12)$$

where $\gamma = \log\eta$ and $\epsilon_X(m) = \mathbf{x}^T(m)\mathbf{x}(m) = \|\mathbf{x}(m)\|^2$, $\epsilon_X(n) = \|\mathbf{x}(n)\|^2$, $\epsilon_Y(m) = \|\mathbf{y}(m)\|^2$, $\epsilon_Y(n) = \|\mathbf{y}(n)\|^2$ are the energies of signals $\mathbf{x}(m)$, $\mathbf{x}(n)$, $\mathbf{y}(m)$ and $\mathbf{y}(n)$, respectively, and the test t is characterized by:

$$\begin{cases} H_0: & T \sim \mathcal{N}\left(-\frac{1}{2}\left(\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}\right), \frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}\right), \\ H_1: & T \sim \mathcal{N}\left(+\frac{1}{2}\left(\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}\right), \frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}\right), \end{cases} \quad (13)$$

where $d^2_X = \|\mathbf{x}(m)-\mathbf{x}(n)\|^2$ and $d^2_Y = \|\mathbf{y}(m)-\mathbf{y}(n)\|^2$.

The probabilities of false acceptance P_F and correct detection P_D can be now found as:

$$\begin{cases} P_F = \Pr[T>\gamma|H_0] = Q\left(\dfrac{\gamma + \frac{1}{2}\left(\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}\right)}{\sqrt{\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}}}\right), \\ P_D = \Pr[T>\gamma|H_1] = Q\left(\dfrac{\gamma - \frac{1}{2}\left(\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}\right)}{\sqrt{\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}}}\right). \end{cases} \quad (14)$$

To determine a threshold γ, we set $P_F = \alpha$, which yields:

$$\gamma = \sqrt{\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}}\,Q^{-1}(\alpha) - \frac{1}{2}\left(\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}\right), \quad (15)$$

where $Q(.)$ is the Q-function, that results in:

$$P_D(\alpha) = Q\left(Q^{-1}(\alpha) - \sqrt{\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}}\right). \quad (16)$$

Let $\bar{\epsilon}_X = \frac{1}{2}(\epsilon_X(m) + \epsilon_X(n))$ and $\bar{\epsilon}_Y = \frac{1}{2}(\epsilon_Y(m) + \epsilon_Y(n))$, which assume equals prior probabilities. Then:

$$\begin{aligned} d^2_X &= \|\mathbf{x}(m) - \mathbf{x}(n)\|^2 = 2\bar{\epsilon}_X(1-\kappa_X), \\ d^2_Y &= \|\mathbf{y}(m) - \mathbf{y}(n)\|^2 = 2\bar{\epsilon}_Y(1-\kappa_Y), \end{aligned} \quad (17)$$

where κ_X and κ_Y are the coefficients such that $|\kappa_X| \leq 1$ and $|\kappa_Y| \leq 1$ and defined as:

$$\begin{aligned} \kappa_X &= \frac{\mathbf{x}^T(m)\mathbf{x}(n)}{\frac{1}{2}(\mathbf{x}^T(m)\mathbf{x}(m) + \mathbf{x}^T(n)\mathbf{x}(n))}, \\ \kappa_Y &= \frac{\mathbf{y}^T(m)\mathbf{y}(n)}{\frac{1}{2}(\mathbf{y}^T(m)\mathbf{y}(m) + \mathbf{y}^T(n)\mathbf{y}(n))}. \end{aligned} \quad (18)$$

If $\kappa_X = 0$, $\mathbf{x}^T(m)\mathbf{x}(n) = 0$ and $\kappa_Y = 0$, $\mathbf{y}^T(m)\mathbf{y}(n) = 0$ and the corresponding pairs of vectors are orthogonal.

If we also assume that both pairs of signals have the same energy, i.e., $\epsilon_X(m) = \epsilon_X(n) = \epsilon_X$ and $\epsilon_Y(m) = \epsilon_Y(n) = \epsilon_Y$ and $\kappa_X = \kappa_Y = 0$, then:

$$P_D(\alpha) = Q\left(Q^{-1}(\alpha) - \sqrt{2(\xi_X + \xi_Y)}\right), \quad (19)$$

where $\xi_X = \frac{\epsilon_X}{\sigma^2_{Z_X}}$ and $\xi_Y = \frac{\epsilon_Y}{\sigma^2_{Z_Y}}$ (we also define signal-to-noise ratios $SNR_X = 10\log_{10}\xi_X$ and $SNR_Y = 10\log_{10}\xi_Y$).

The average probability of error is:

$$\begin{aligned} P_e &= \tfrac{1}{2}P_F + \tfrac{1}{2}(1-P_D) \\ &= Q\left(\frac{1}{2}\sqrt{\frac{d^2_X}{\sigma^2_{Z_X}} + \frac{d^2_Y}{\sigma^2_{Z_Y}}}\right) = Q\left(\sqrt{\tfrac{1}{2}(\xi_X + \xi_Y)}\right). \end{aligned} \quad (20)$$

Therefore, the fusion of independent modalities is beneficial in terms of increase of overall SNR.

We conclude our analysis by considering the bounds on error probabilities for the generic distributions and particular Gaussian assumptions. For this reason, it is important to establish the impact of modality fusion on the distribution distances such as Chernoff distance.

We will define the Chernoff distance between two distributions as:

$$\begin{aligned} &D_s(p(\mathbf{v},\mathbf{w}|H_1), p(\mathbf{v},\mathbf{w}|H_0)) = \\ &-\log\int_{\mathcal{V}^{N_X}}\int_{\mathcal{W}^{N_Y}} p(\mathbf{v},\mathbf{w}|H_1)\left(\frac{p(\mathbf{v},\mathbf{w}|H_1)}{p(\mathbf{v},\mathbf{w}|H_0)}\right)^s d\mathbf{v}d\mathbf{w}, \end{aligned} \quad (21)$$

and its first derivative with respect to s as $\dot{D}_s(.,.)$. For the simplicity of notations, we redenote $\mu(s) = -D_s(p(\mathbf{v},\mathbf{w}|H_1), p(\mathbf{v},\mathbf{w}|H_0))$ and $\dot{\mu}(s) = -\dot{D}_s(p(\mathbf{v},\mathbf{w}|H_1), p(\mathbf{v},\mathbf{w}|H_0))$.

The Chernoff distance provides an upper bound on both the probability of false acceptance P_F and probability of

[3] The selection of the Gaussian noise is explained by the largest differential entropy (worst case conditions for the authentication) among all distributions with the bounded variance.

miss $P_M = 1 - P_D$ [21] [4]:

$$\begin{cases} P_F \leq e^{\mu(s) - s\dot{\mu}(s)}, \\ P_M \leq e^{\mu(s) + (1-s)\dot{\mu}(s)}, \end{cases} \tag{22}$$

for $0 \leq s \leq 1$ and with $\eta = \dot{\mu}(s)$. The fastest convergence rate of the exponential terms is achieved for the optimal selection of s.

For the average probability of error [21]:

$$P_e \leq \frac{1}{2} e^{\mu(s_m)}, \tag{23}$$

where s_m is the value for which $\dot{\mu}(s) = 0$.

For the independent modalities, the Chernoff distance satisfies the additivity property and reduces to:

$$\begin{array}{l} D_s(p(\mathbf{v},\mathbf{w}|H_1), p(\mathbf{v},\mathbf{w}|H_0)) = \\ D_s(p(\mathbf{v}|H_1), p(\mathbf{v}|H_0)) + D_s(p(\mathbf{w}|H_1), p(\mathbf{w}|H_0)). \end{array} \tag{24}$$

Therefore, similarly to the exact result for the Gaussian distributions the presence of the second modality increases the convergence rate to zero. One can also extend these bounds to the considered Gaussian case (10) for which:

$$\begin{array}{l} \mu(s) = \frac{s(s-1)}{2} \left(\frac{d_X^2}{\sigma_{Z_X}^2} + \frac{d_Y^2}{\sigma_{Z_Y}^2} \right), \\ \dot{\mu}(s) = \frac{(2s-1)}{2} \left(\frac{d_X^2}{\sigma_{Z_X}^2} + \frac{d_Y^2}{\sigma_{Z_Y}^2} \right), \end{array} \tag{25}$$

which is maximized for $s = s_m = 0.5$ and that results in:

$$\begin{cases} P_F \leq e^{-\frac{1}{8}\left(\frac{d_X^2}{\sigma_{Z_X}^2} + \frac{d_Y^2}{\sigma_{Z_Y}^2} \right)} = e^{-\frac{1}{4}(\xi_X + \xi_Y)}, \\ P_M \leq e^{-\frac{1}{8}\left(\frac{d_X^2}{\sigma_{Z_X}^2} + \frac{d_Y^2}{\sigma_{Z_Y}^2} \right)} = e^{-\frac{1}{4}(\xi_X + \xi_Y)}, \end{cases} \tag{26}$$

and the average probability of error:

$$P_e \leq \frac{1}{2} e^{-\frac{1}{8}\left(\frac{d_X^2}{\sigma_{Z_X}^2} + \frac{d_Y^2}{\sigma_{Z_Y}^2} \right)} = \frac{1}{2} e^{-\frac{1}{4}(\xi_X + \xi_Y)}. \tag{27}$$

To show the link with the previous exact results, we will demonstrate on the case of P_e (20) the link with the above bounds. For this purpose, we will use the approximation for large argument α of $Q(\alpha) \approx \frac{1}{\sqrt{2\pi\alpha^2}} e^{-\frac{\alpha^2}{2}}$, which yields for the exact P_e:

$$P_e \approx \frac{1}{\sqrt{\frac{\pi}{2}\left(\frac{d_X^2}{\sigma_{Z_X}^2} + \frac{d_Y^2}{\sigma_{Z_Y}^2} \right)}} e^{-\frac{1}{8}\left(\frac{d_X^2}{\sigma_{Z_X}^2} + \frac{d_Y^2}{\sigma_{Z_Y}^2} \right)}, \tag{28}$$

that coincides up to the exponential term with (27).

4. AUTHENTICATION BASED ON RANDOM PROJECTIONS

In this section, we will consider the impact of the dimensionality reduction on the performance according to the setup shown in Figure 10.

[4]These bounds are general and applied to any distributions and vector lengths. While this gives a precise exponential rate for the convergence of these probabilities to zero, the result can be improved using asymptotic integral expansion technique for large N and independent components.

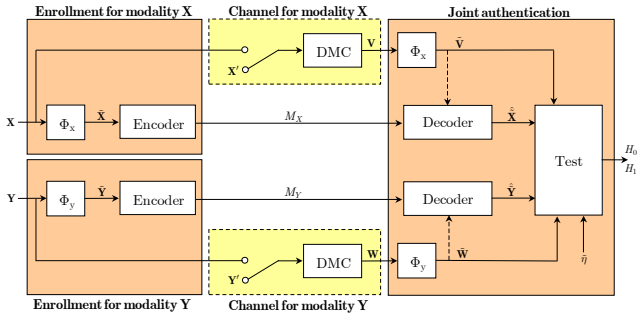


Figure 10: Multimodal authentication architecture based on source coding and random projections.

The main difference with respect to the previously considered setup consists in the usage of key-defined projection operators $\mathbf{\Phi_x}$ and $\mathbf{\Phi_y}$ defined as:

$$\begin{array}{l} \tilde{\mathbf{x}} = \mathbf{\Phi_x}\mathbf{x}, \\ \tilde{\mathbf{y}} = \mathbf{\Phi_y}\mathbf{y}, \end{array} \tag{29}$$

where $\mathbf{x} \in \mathbb{R}^{N_X}$, $\mathbf{y} \in \mathbb{R}^{N_Y}$, $\tilde{\mathbf{x}} \in \mathbb{R}^{L_X}$, $\tilde{\mathbf{y}} \in \mathbb{R}^{L_Y}$, $\mathbf{\Phi_x} \in \mathbb{R}^{L_X \times N_X}$, $\mathbf{\Phi_y} \in \mathbb{R}^{L_Y \times N_Y}$, $L_X \leq N_X$ and $L_Y \leq N_Y$. We assume that the matrices $\mathbf{\Phi_x}$ and $\mathbf{\Phi_y}$ are orthoprojectors for which $\mathbf{\Phi_x}\mathbf{\Phi_x}^T = \mathbf{I}_{L_X}$ and $\mathbf{\Phi_y}\mathbf{\Phi_y}^T = \mathbf{I}_{L_Y}$.

The corresponding hypotheses (7) can be reformulated as:

$$\begin{cases} \tilde{H}_0: & \tilde{\mathbf{v}} = \mathbf{\Phi_x}(\mathbf{x}(n) + \mathbf{z}_X) = \tilde{\mathbf{x}}(n) + \tilde{\mathbf{z}}_X, \\ & \tilde{\mathbf{w}} = \mathbf{\Phi_y}(\mathbf{y}(n) + \mathbf{z}_Y) = \tilde{\mathbf{y}}(n) + \tilde{\mathbf{z}}_Y, \\ \tilde{H}_1: & \tilde{\mathbf{v}} = \mathbf{\Phi_x}(\mathbf{x}(m) + \mathbf{z}_X) = \tilde{\mathbf{x}}(m) + \tilde{\mathbf{z}}_X, \\ & \tilde{\mathbf{w}} = \mathbf{\Phi_y}(\mathbf{y}(m) + \mathbf{z}_Y) = \tilde{\mathbf{y}}(m) + \tilde{\mathbf{z}}_Y, \end{cases} \tag{30}$$

that leads to the test:

$$\Lambda(\tilde{\mathbf{v}}, \tilde{\mathbf{w}}) = \frac{p(\tilde{\mathbf{v}}, \tilde{\mathbf{w}}|\tilde{H}_1)}{p(\tilde{\mathbf{v}}, \tilde{\mathbf{w}}|\tilde{H}_0)} = \frac{p(\tilde{\mathbf{v}}|\tilde{H}_1) p(\tilde{\mathbf{w}}|\tilde{H}_1)}{p(\tilde{\mathbf{v}}|\tilde{H}_0) p(\tilde{\mathbf{w}}|\tilde{H}_0)} \lessgtr \tilde{\eta}, \tag{31}$$

with the distributions under hypothesis:

$$\begin{array}{l} p(\tilde{\mathbf{v}}, \tilde{\mathbf{w}}|\tilde{H}_0) = \mathcal{N}(\tilde{\mathbf{x}}(n), \sigma_{Z_X}^2 \mathbf{C_x}) \mathcal{N}(\tilde{\mathbf{y}}(n), \sigma_{Z_Y}^2 \mathbf{C_y}), \\ p(\tilde{\mathbf{v}}, \tilde{\mathbf{w}}|\tilde{H}_1) = \mathcal{N}(\tilde{\mathbf{x}}(m), \sigma_{Z_X}^2 \mathbf{C_x}) \mathcal{N}(\tilde{\mathbf{y}}(m), \sigma_{Z_Y}^2 \mathbf{C_y}), \end{array} \tag{32}$$

where $\mathbf{C_x} = \mathbf{\Phi_x}\mathbf{\Phi_x}^T$ and $\mathbf{C_y} = \mathbf{\Phi_y}\mathbf{\Phi_y}^T$.

Similarly, one can deduce the sufficient statistic $\tilde{t}$:

$$\begin{array}{l} \tilde{t}(\tilde{\mathbf{v}}, \tilde{\mathbf{w}}) := \frac{1}{\sigma_{Z_X}^2} \tilde{\mathbf{v}}^T \mathbf{C_x}^{-1} \left(\tilde{\mathbf{x}}(m) - \tilde{\mathbf{x}}(n) \right) \\ -\frac{1}{2\sigma_{Z_X}^2} (\tilde{\mathbf{x}}^T(m) \mathbf{C_x}^{-1} \tilde{\mathbf{x}}(m) - \tilde{\mathbf{x}}^T(n) \mathbf{C_x}^{-1} \tilde{\mathbf{x}}(n)) \\ +\frac{1}{\sigma_{Z_Y}^2} \tilde{\mathbf{w}}^T \left(\tilde{\mathbf{y}}(m) - \tilde{\mathbf{y}}(n) \right) \\ -\frac{1}{2\sigma_{Z_Y}^2} (\tilde{\mathbf{y}}^T(m) \mathbf{C_y}^{-1} \tilde{\mathbf{y}}(m) - \tilde{\mathbf{y}}^T(n) \mathbf{C_y}^{-1} \tilde{\mathbf{y}}(n)) \lessgtr \tilde{\gamma}, \end{array} \tag{33}$$

where $\tilde{\gamma} = \log \tilde{\eta}$, which is characterized by:

$$\begin{cases} \tilde{H}_0: & \tilde{T} \sim \mathcal{N}\left(-\frac{1}{2}\left(\frac{\tilde{d}_X^2}{\sigma_{Z_X}^2} + \frac{\tilde{d}_Y^2}{\sigma_{Z_Y}^2} \right), \frac{\tilde{d}_X^2}{\sigma_{Z_X}^2} + \frac{\tilde{d}_Y^2}{\sigma_{Z_Y}^2} \right), \\ \tilde{H}_1: & \tilde{T} \sim \mathcal{N}\left(+\frac{1}{2}\left(\frac{\tilde{d}_X^2}{\sigma_{Z_X}^2} + \frac{\tilde{d}_Y^2}{\sigma_{Z_Y}^2} \right), \frac{\tilde{d}_X^2}{\sigma_{Z_X}^2} + \frac{\tilde{d}_Y^2}{\sigma_{Z_Y}^2} \right), \end{cases} \tag{34}$$

where $\tilde{d}_X^2 = (\mathbf{x}(m) - \mathbf{x}(n))^T \mathbf{\Phi_x}^T \mathbf{C_x}^{-1} \mathbf{\Phi_x} (\mathbf{x}(m) - \mathbf{x}(n))$ and $\tilde{d}_Y^2 = (\mathbf{y}(m) - \mathbf{y}(n))^T \mathbf{\Phi_y}^T \mathbf{C_y}^{-1} \mathbf{\Phi_y} (\mathbf{y}(m) - \mathbf{y}(n))$.

The probabilities of false acceptance P_F and correct de-

tection P_D can be now found as:

$$\begin{cases} \tilde{P}_F = \Pr[\tilde{T} > \tilde{\gamma}|\tilde{H}_0] = Q\left(\dfrac{\tilde{\gamma}+\frac{1}{2}\left(\frac{\tilde{d}_X^2}{\sigma_{Z_X}^2}+\frac{\tilde{d}_Y^2}{\sigma_{Z_Y}^2}\right)}{\sqrt{\frac{d_X^2}{\sigma_{Z_X}^2}+\frac{d_Y^2}{\sigma_{Z_Y}^2}}}\right), \\ \tilde{P}_D = \Pr[\tilde{T} > \tilde{\gamma}|\tilde{H}_1] = Q\left(\dfrac{\tilde{\gamma}-\frac{1}{2}\left(\frac{\tilde{d}_X^2}{\sigma_{Z_X}^2}+\frac{\tilde{d}_Y^2}{\sigma_{Z_Y}^2}\right)}{\sqrt{\frac{\tilde{d}_X^2}{\sigma_{Z_X}^2}+\frac{\tilde{d}_Y^2}{\sigma_{Z_Y}^2}}}\right). \end{cases} \tag{35}$$

Assuming $\tilde{P}_F = \alpha$, one obtains:

$$\tilde{P}_D(\alpha) = Q\left(Q^{-1}(\alpha) - \sqrt{\frac{\tilde{d}_X^2}{\sigma_{Z_X}^2}+\frac{\tilde{d}_Y^2}{\sigma_{Z_Y}^2}}\right). \tag{36}$$

Taking into account the condition of orthoprojection for $\mathbf{\Phi_x}$ and $\mathbf{\Phi_y}$ one obtains $\mathbf{C_x} = \mathbf{\Phi_x}\mathbf{\Phi_x}^T = \mathbf{I}_{L_X}$, $\mathbf{C_y} = \mathbf{\Phi_y}\mathbf{\Phi_y}^T = \mathbf{I}_{L_Y}$ and for the distances:

$$\begin{aligned} \tilde{d}_X^2 &= (\mathbf{x}(m)-\mathbf{x}(n))^T \mathbf{\Phi}_\mathbf{x}^T\mathbf{\Phi_x}(\mathbf{x}(m)-\mathbf{x}(n)), \\ &= \|\mathbf{\Phi_x}(\mathbf{x}(m)-\mathbf{x}(n))\|^2, \end{aligned} \tag{37}$$

$$\begin{aligned} \tilde{d}_Y^2 &= (\mathbf{y}(m)-\mathbf{y}(n))^T \mathbf{\Phi}_\mathbf{y}^T\mathbf{\Phi_y}(\mathbf{y}(m)-\mathbf{y}(n)), \\ &= \|\mathbf{\Phi_y}(\mathbf{y}(m)-\mathbf{y}(n))\|^2. \end{aligned} \tag{38}$$

To introduce the bounds on the distance $\tilde{d}^2$ we will use the results of Johnson-Lindenstrauss lemma [22], which states that with high probability the geometry of a point cloud is not disturbed by certain Lipschitz mappings onto a space of dimension logarithmic in the number of points.

According to Jonhnson-Lindenstrauss result [22]:

$$\begin{aligned} (1-\zeta)\sqrt{\tfrac{L_X}{N_X}} \le \tfrac{||\mathbf{\Phi_x x}||}{||\mathbf{x}||} \le (1+\zeta)\sqrt{\tfrac{L_X}{N_X}}, \\ (1-\zeta)\sqrt{\tfrac{L_Y}{N_Y}} \le \tfrac{||\mathbf{\Phi_y y}||}{||\mathbf{y}||} \le (1+\zeta)\sqrt{\tfrac{L_Y}{N_Y}}, \end{aligned} \tag{39}$$

where $0 < \zeta < 1$ with high probability.

This allows to use the approximation for the random orthoprojectors $\mathbf{\Phi_x}$ and $\mathbf{\Phi_y}$ as:

$$\begin{aligned} (1-\zeta)\sqrt{\tfrac{L_X}{N_X}}||\mathbf{x}|| \le ||\mathbf{\Phi_x x}|| \le (1+\zeta)\sqrt{\tfrac{L_X}{N_X}}||\mathbf{x}||, \\ (1-\zeta)\sqrt{\tfrac{L_Y}{N_Y}}||\mathbf{y}|| \le ||\mathbf{\Phi_y y}|| \le (1+\zeta)\sqrt{\tfrac{L_Y}{N_Y}}||\mathbf{y}||. \end{aligned} \tag{40}$$

Thus, with high probability one can approximate (36) as follows:

$$\tilde{P}_D(\alpha) \approx Q\left(Q^{-1}(\alpha) - \sqrt{\frac{L_X}{N_X}\frac{d_X^2}{\sigma_{Z_X}^2}+\frac{L_Y}{N_Y}\frac{d_Y^2}{\sigma_{Z_Y}^2}}\right), \tag{41}$$

that makes possible to estimate the corresponding loss with respect to the equation (16). The random projections introduce the loss in the distance between codewords proportional to $\sqrt{\frac{L_X}{N_X}}$ and $\sqrt{\frac{L_Y}{N_Y}}$ for each modality, respectively.

Equivalently to (19) for the equiprobable orthogonal signals with the same energy, one can rewrite the above approximation as:

$$\tilde{P}_D(\alpha) = Q\left(Q^{-1}(\alpha) - \sqrt{2\left(\frac{L_X}{N_X}\xi_X+\frac{L_Y}{N_Y}\xi_Y\right)}\right). \tag{42}$$

Finally, the average probability of error computed for the direct domain according to (20) can be found for the random projections domain as:

$$\tilde{P}_e = Q\left(\sqrt{\frac{1}{2}\left(\frac{\tilde{\epsilon}_X}{\sigma_{Z_X}^2}+\frac{\tilde{\epsilon}_Y}{\sigma_{Z_Y}^2}\right)}\right), \tag{43}$$

where $\tilde{\epsilon}_X = \|\Phi_\mathbf{x}\mathbf{x}(m)\|^2 = \|\Phi_\mathbf{x}\mathbf{x}(n)\|^2$ and $\tilde{\epsilon}_Y = \|\Phi_\mathbf{y}\mathbf{y}(m)\|^2 = \|\Phi_\mathbf{y}\mathbf{y}(n)\|^2$ with the approximation:

$$\tilde{P}_e \approx Q\left(\sqrt{\frac{1}{2}\left(\frac{L_X}{N_X}\xi_X+\frac{L_Y}{N_Y}\xi_Y\right)}\right). \tag{44}$$

It should be also pointed out that the random projections reduce the information distances in the corresponding error exponents due to the data processing inequality [5]:

$$D_s(p(\tilde{\mathbf{v}},\tilde{\mathbf{w}}|\tilde{H}_1), p(\tilde{\mathbf{v}},\tilde{\mathbf{w}}|\tilde{H}_0)) \le D_s(p(\mathbf{v},\mathbf{w}|H_1), p(\mathbf{v},\mathbf{w}|H_0)). \tag{45}$$

In particular, for the considered independent modalities:

$$\begin{aligned} &D_s(p(\tilde{\mathbf{v}},\tilde{\mathbf{w}}|\tilde{H}_1), p(\tilde{\mathbf{v}},\tilde{\mathbf{w}}|\tilde{H}_0)) = \\ &D_s(p(\tilde{\mathbf{v}}|\tilde{H}_1), p(\tilde{\mathbf{v}}|\tilde{H}_0)) + D_s(p(\tilde{\mathbf{w}}|\tilde{H}_1), p(\tilde{\mathbf{w}}|\tilde{H}_0)), \end{aligned} \tag{46}$$

and for the Gaussian case (32) with the orthoprojectors $\mathbf{\Phi_x}$ and $\mathbf{\Phi_y}$, one can rewrite (25) as:

$$D_s(p(\tilde{\mathbf{v}},\tilde{\mathbf{w}}|\tilde{H}_1), p(\tilde{\mathbf{v}},\tilde{\mathbf{w}}|\tilde{H}_0)) = -\frac{s(s-1)}{2}\left(\frac{\tilde{d}_X^2}{\sigma_{Z_X}^2}+\frac{\tilde{d}_Y^2}{\sigma_{Z_Y}^2}\right). \tag{47}$$

Using the result of Jonhnson-Lindenstrauss (40):

$$\begin{aligned} &D_s(p(\tilde{\mathbf{v}},\tilde{\mathbf{w}}|\tilde{H}_1), p(\tilde{\mathbf{v}},\tilde{\mathbf{w}}|\tilde{H}_0)) \approx \\ &-\tfrac{s(s-1)}{2}\left(\tfrac{L_X}{N_X}\tfrac{d_X^2}{\sigma_{Z_X}^2}+\tfrac{L_Y}{N_Y}\tfrac{d_Y^2}{\sigma_{Z_Y}^2}\right). \end{aligned} \tag{48}$$

This confirms the data processing inequality result (45) and demonstrates the decease in the Chernoff distance due to the random projections.

5. RESULTS OF COMPUTER SIMULATION

In this section we will demonstrate the main results: the impact of additional possibly noisy modality on the performance enhancement of authentication system in terms of ROC; the impact of dimensionality reduction based on the random projection and approximation accuracy of authentication performance based on the random projection using the Jonhnson-Lindenstrauss lemma.

The impact of the second modality $\mathbf{Y}$ on the performance of authentication system based on the modality $\mathbf{X}$ is demonstrated in Figures 11 for various P_F without dimensionality reduction. In all considered cases, the presence of noisy modality $\mathbf{Y}$ increases P_D. Contrarily, the impact of the second modality reduces to zero as its $SNR_Y \to -\infty$. The overall performance as the function of SNR_X and SNR_Y for the fixed $P_F = 10^{-5}$ is shown in Figure 12.

The impact of dimensionality reduction based on the random projection and approximation accuracy was investigated using both analytical formulas and Monte Carlo simulation for Gaussian data of lengths $N_X = N_Y = 3500$. The orthoprojectors $\mathbf{\Phi_x}$ and $\mathbf{\Phi_y}$ are generated from the independent realizations of Gaussian random variables $\Phi_{\mathbf{x}_{i,j}} \sim \mathcal{N}(0, \frac{1}{N_X})$ and $\Phi_{\mathbf{y}_{i,j}} \sim \mathcal{N}(0, \frac{1}{N_Y})$. The results of simulation for the probability P_D are shown in Figures 13 and 14. The dimensionality reduction of modality $\mathbf{X}$ for the fixed length of modality $\mathbf{Y}$ in the indicated ranges revealed the loss in performance about 5 dB for both SNR_Ys. This loss is relative small price for the reduced complexity of processing and memory storage as well as security/privacy enhancement. Moreover, the random projection approximation based on the Jonhnson-Lindenstrauss lemma demonstrates quite accurate results and can be used for the estimation of performance. Finally, as in the previous case, the presence of

the second modality with the positive SNR increased the accuracy of authentication.

6. CONCLUSIONS

In this paper, we have investigated the impact of additional modality presence on the authentication performance in terms of ROC and average probability of error. To avoid the information loss, the authentication setup was based on the raw data fusion. Even in the case of statistically independent modalities belonging to the same object, we observed the reduction of the corresponding error probabilities computed according to the exact formulas for the Gaussian assumptions and generic error exponent bounds for any distributions. In the same line, we also investigated the impact of dimensionality reduction performed using random orthoprojectors and proposed the corresponding approximations using the Jonhnson-Lindenstrauss lemma. It was established that the orthoprojectors reduce the vector and distribution distances proportionally to the ratio of the vector lengths after and before projection. These findings might be of interest for the design of practical authentication systems. In future, we plan to consider the security of the presented setup and study possible attacks.

7. ACKNOWLEDGMENTS

This paper was partially supported by SNF projects 114613, 200021–111643, 200021–1119770 and IM2 project.

8. REFERENCES

[1] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: A tool for information security," *IEEE Trans. on Information Forensics and Security*, vol. 1, no. 2, pp. 125–143, June 2006.

[2] L. Hong, A. Jain, and S. Pankanti, "Can multibiometrics improve performance," Summit, NJ, pp. 59–64, Oct 1999.

[3] B. Ulery, W. Fellner, A. H. P. Hallinan and, and C. Watson, "Evaluation of selected biometric fusion techniques," NIST report 7346, 2006, Tech. Rep.

[4] R. Brunelli and D. Falavigna, "Person identification using multiple cues," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 10, pp. 955–966, 1995.

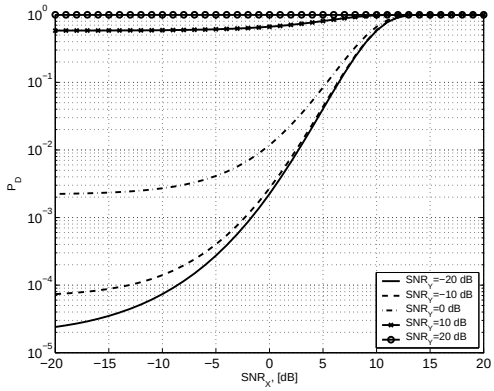
[5] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley and Sons, New York, 1991.

[6] A. Ross, R. G. Hong, A. K. Jain, and S. Pankanti, "Feature level fusion using hand and face biometrics," in *SPIE Conf. Biometric Technology for Human Identification II*, 2005, pp. 196–204.

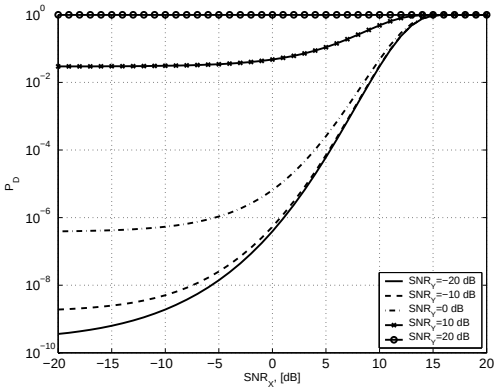
[7] N. Ratha, J. Connell, and R. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM Systems Journal*, vol. 40, no. 3, pp. 614–634, 2001.

[8] U. Maurer, "A unified and generalized treatment of authentication theory," in *Proc. 13th Symp. on Theoretical Aspects of Computer Science (STACS'96)*, ser. Lecture Notes in Computer Science, vol. 1046. Springer-Verlag, Feb 1996, pp. 387–398.

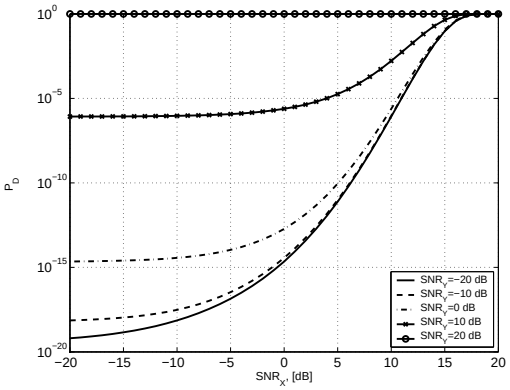
[9] O. Zeitouni, J. Ziv, and N. Merhav, "When is the generalized likelihood ratio test optimal?" *IEEE Trans Information Theory*, vol. 38, no. 5, pp. 1597–1602, 1992.



(a)



(b)



(c)

Figure 11: Probability of correct detection P_D for P_F equals (a) 10^{-5}, (b) 10^{-10} and (c) 10^{-20} for $\frac{L_X}{N_X} = \frac{L_Y}{N_Y} = 1$.

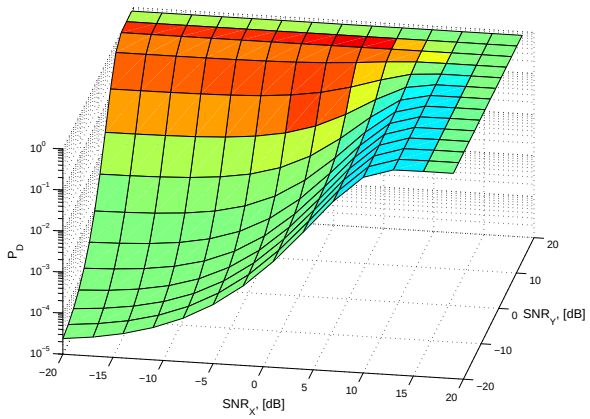
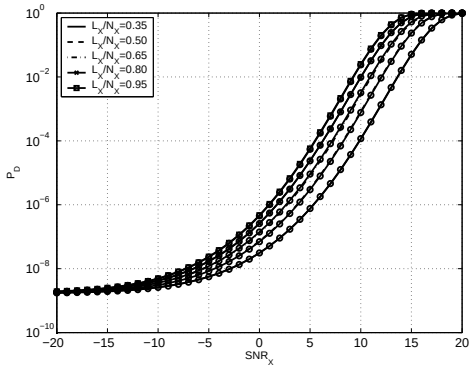
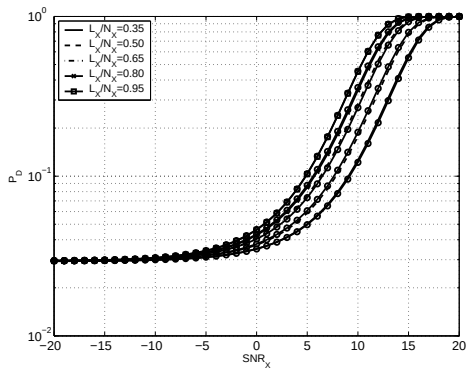


Figure 12: Probability of correct detection P_D: $\frac{L_X}{N_X} = \frac{L_Y}{N_Y} = 1$, $P_F = 10^{-5}$.



(a)



(b)

Figure 13: Probability of correct detection P_D for the fixed $P_F = 10^{-10}$ and various dimensionality X reduction ratios with $\frac{L_Y}{N_Y} = 1$ for: (a) $SNR_Y = -10$ dB and (b) $SNR_Y = +10$ dB. The lines with circles show the corresponding approximations.

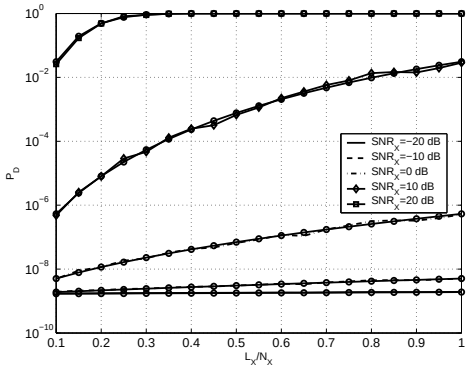


Figure 14: Probability of correct detection P_D for the fixed $P_F = 10^{-10}$ and $SNR_Y = -10$ dB and $\frac{L_Y}{N_Y} = 1$. The lines with circles show the corresponding approximations.

[10] J. Fridrich, "Robust bit extraction from images," in *Proceedings ICMCS'99*, vol. 2, Florence, Italy, June 1999, pp. 536–540.

[11] S. S. Kozat, R. Venkatesan, and M. K. Mihcak, "Robust hashing via matrix invariances," in *ICIP 2004*, vol. 3, Singapore, October 2004, pp. 677–680.

[12] P. Tuyls, B. Skoric, and T. K. (Eds.), *Security with Noisy Data: On Private Biometrics, Secure Key Storage and Anti-Counterfeiting*. Springer, 2007.

[13] D. Slepian and J. Wolf, "Noiseless encoding of correlated information sources," *IEEE Trans. Information Theory*, vol. 19, pp. 471–480, July 1973.

[14] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inform. Theory*, vol. 39, pp. 733–742, 1993.

[15] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography - Part I: secret sharing," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1121–1132, 1993.

[16] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Information Theory*, vol. 22, no. 1, pp. 1–10, January 1976.

[17] T. Ignatenko and F. Willems, "On the security of xor-method in biometric authentication systems," in *The twenty-seventh symposium on Information Theory in the Benelux*, Noordwijk, The Netherlands, June 8-9 2006, pp. 197–204.

[18] E. Martinian, S. Yekhanin, and J. Yedidia, "Secure biometrics via syndromes," in *43rd Annual Allerton Conference on Communications, Control, and Computing*, Monticello, IL, USA, October 2005.

[19] Y.-C. Lin, D. Varodayan, and B. Girod, "Image authentication based on distributed source coding," in *IEEE International Conference on Image Processing (ICIP2007)*, San Antonio, USA, September 2007.

[20] H. L. V. Trees, *Detection, Estimation, and Modulation Theory*. New York: John Wiley and Sons, 1968.

[21] W. B. Johnson and J. Lindenstrauss, "Extensions of Lipschitz mappings into Hilbert space," *Contemporary Mathematics*, no. 26, pp. 189–206, 1984.