Identification In Desynchronization Channels

Koval, Oleksiy; Voloshynovskyy, Svyatoslav; Farhadzadeh, Farzad

How to cite

KOVAL, Oleksiy, VOLOSHYNOVSKYY, Svyatoslav, FARHADZADEH, Farzad. Identification In Desynchronization Channels. In: IEEE Information Theory Workshop (ITW). Paraty (Brazil). [s.l.] : [s.n.], 2011. p. 297–301. doi: 10.1109/ITW.2011.6089440

This publication URL: https://archive-ouverte.unige.ch//unige:47624
Publication DOI: 10.1109/ITW.2011.6089440

Identification In Desynchronization Channels

Oleksiy Koval
Computer Science Department
University of Geneva
Geneva, Switzerland
Email: Oleksiy.Koval@unige.ch

Svyatoslav Voloshynovskiy
Computer Science Department
University of Geneva
Geneva, Switzerland
Email: svolos@unige.ch

Farzad Farhadzadeh
Computer Science Department
University of Geneva
Geneva, Switzerland
Email: Farzad.Farhadzadeh@unige.ch

***Abstract*—In this paper we analyze the problem of object identification in channels with desynchronization. In our analysis we assume that the identification system is designed using a pilot-based re-synchronization mechanism that assists desynchronization compensation with a certain accuracy. We demonstrate how the accuracy of re-synchronization impacts the information-theoretic limits of identification system performance.**

I. Introduction

This paper deals with a problem of the content identification that arises in various multimedia management (broadcast monitoring, tracing and tracking, monetizing, copy detection) and security (biometric person identification, anti-counterfeiting, document authentication) applications. Due to an exponential growth of the analyzed data along with a significant progress in redistribution channels (radio, TV, Internet) observed in the last decades, accurate yet computationally efficient content identification tools are highly demanded. Besides the mentioned aspects of performance and complexity, the overall trade-off such means should satisfy includes as well protection of a privacy sensitive information that is critical in person identification and medical applications and storage memory.

Current solutions to this trade-off are usually designed based on digital fingerprints or robust perceptual hashes that can be considered as compact and robust representations of multimedia objects/contents designed for their distinctive, computationally efficient and privacy preserving management.

Digital fingerprinting that was yet recently considered as only an alternative to digital watermarking in the mentioned applications has become unavoidable in cases when modifications of the original content due to watermark embedding are undesirable (art objects), hardly possible without severe and unpredictable consequences (biometrics including DNA) or conflicting with the assumed protocol physical uniqueness and unclonability (physical unclonable functions (PUFs)).

Since mid of 90th, the domain of robust fingerprinting has performed an impressive evolution. The progress was mainly achieved along two directions: design of practical algorithms and theoretical analysis of attainable performance limits.

The advance attained along the former direction was mostly oriented on design of various robust feature extraction techniques and efficient matching strategies [1],[2]. The efforts along the latter one were mostly spent to analyze the achievable identification rates [3] in the formulation of infinite length codeword transmission over the discrete memoryless channel (DMC). Several groups of authors analyzed the identification problem within the information-detection framework for various channel and codeword length assumptions [4], [5], [6].

One of main underlying assumptions in the performance analysis accomplished in most of the referred papers is the perfect synchronization between the query and the content of the multimedia object database. Such an assumption could not be valid in multimedia object identification and might finally lead to an inaccurate performance analysis due to the potential mismatch between the identified query and the information enrolled in the database. The importance of the desynchronization consideration is additionally justified by omnipresence of such kind of degradations at both multimedia database enrollment and query identification stages even when the acquisition channel does not represent any aggressive attacking behavior. Specifically, it might be assumed that desynchronization in this case is introduced as a consequence of acquisition imperfections at these stages.

Identification in the presence of desynchronization was previously considered by the authors in [4], [5], [7], [8]. In the earlier works [4], [5], identification performance in terms of the average probability of error was analyzed under the generalized maximum likelihood hypothesis testing strategy.

More recent results [7], [8] justify content identification in channels with desynchronization for distortion-free and imperfect enrollment, respectively, from a twofold perspective. First, it is demonstrated that the maximum achievable rate in such systems asymptotically approaches identification capacity over the DMC [3] iff cardinality of a set of possible desynchronization distortions is subexponential in the length of the codewords that approaches infinity. Second, it is shown that desynchronization causes identification performance loss in terms of upper bounds on the probabilities of error of two kinds for finite length communication of binary fingerprints generated based on random projections and binaryzation for the case when the Bounded Distance Decoding (BDD) is applied. However, no optimization in terms of the decoder decision threshold is required versus the DMC identification channel case to attain the optimal performance.

The results in [7], [8] were obtained under an assumption that decoder is not computationally constrained and is able to apply an exhaustive decision making over the codebook with pre-distorted codewords in order to achieve asymptotic invariance to geometrical desynchronization. However,

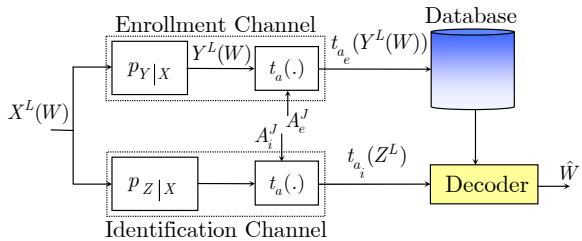

Fig. 1. Multimedia object identification as a communications problem in channels with desynchronization.

in various practical applications channel ambiguity is resolved using estimation/compensation approach rather then based on the exhaustive search [9], [10] in order to reduce the re-synchronization complexity. Therefore, the main goal of this paper is to extend the existing analysis results of the identification problem in channels with desynchronization distortions to the formulation with complexity constraints. The remaining part of the paper is organized as follows. The problem of identification capacity achieving geometrically resilient multimedia object identification with pilot-based resynchronization is formulated and analyzed in Section II. The impact of a database entry length on the system performance is considered in Section III. Section IV contains some experimental results.

Notations We use X to denote scalar random variables, X^N to denote vector random variables, x and x^N to denote realizations of X and X^N, respectively. Superscripts are used to designate length of vectors, i.e., $x^N = [x[1], x[2], ..., x[N]]$ with the ith element $x[i]$. We use $X \sim p(x)$ to indicate that X is distributed according to $p(x)$. Calligraphic fonts $\mathcal{X}$ denote sets $X \in \mathcal{X}$ with cardinalities $|\mathcal{X}|$.

II. PROBLEM FORMULATION AND ANALYSIS

The setup for multimedia object identification as a communication problem is presented in Fig. 1. We assume that there is a set of M multimedia objects represented by corresponding indexes $w \in \{1, 2, ..., M\}$. Every object is associated to a raw fingerprint $x^N = [x_1, x_2, ..., x_N]$ generated i.i.d. according to a certain distribution $p(x)$ from an alphabet $\mathcal{X}$, i.e.:

$$p(x^N) = \prod_{i=1}^{N} p(x[i]), \quad x^N \in \mathcal{X}^N. \tag{1}$$

Similarly to the case considered in [8], it is supposed that original $x^N(w), w \in \{1, 2, ..., M\}$, are not available and the database/codebook contains distorted versions of these fingerprints. The considered distortion model is compound and consists of a memoryless channel part $\{\mathcal{Y}, p(y|x), \mathcal{X}\}, p(y^N(w)|x^N(w)) = \prod_{i=1}^{N} p(y(w)[i]|x(w)[i])$, where $\mathcal{Y}$ is the enrollment channel output alphabet defined in appliance with a predefined feature extraction procedure (for example like in [1], [2]) and a desynchronization distortion part that is modeled as a parametric mapping:

$$t_a : \mathcal{A}^J \times \mathcal{Y}^L \rightarrow \mathcal{Y}^L. \tag{2}$$

According to this definition, it is assumed that the mapper is parametrized by a set of J parameters, $a^J = [a[1], a[2], ..., a[J]]$, taking their values in a set of finite cardinality $\mathcal{A}$. At the enrollment, this mapper is governed by a parameter vector a_e^J. A possible example of desynchronizations that can be modeled under (2) is a class of general affine transforms. In this case, $J = 6$ and vector parameters $[a[1], a[2], ..., a[6]]$ control rotation, scaling and translations in two dimensions.

Furthermore, we specifically suppose in this paper that such a mapping impacts only the coordinates of the elements of their input and does not modify this input cardinality. We suppose as well that the defined desynchronization preserves sample independence of its input.

In the identification stage, a raw fingerprint representing unknown multimedia object z^N is received at the output of the memoryless identification channel $\{\mathcal{Z}, p(z|x), \mathcal{X}\}$, where $\mathcal{Z}$ is the identification channel output alphabet:

$$p(z^N|x^N) = \prod_{i=1}^{N} p(z[i]|x[i]) \tag{3}$$

that is further distorted according to (2) with $a_i^J \neq a_e^J$.

Evidently, since $a_i^J \neq a_e^J$ in a general situation, no reliable communications is possible due to desynchronization of the input of the enrollment and the query. In order to compensate this type of distortions in the identification capacity achieving regime [7], [8], it was suggested to use an optimal decoding over the entire set $|\mathcal{A}|^J$. The main motivation was that in the case the query is distorted by $a^J = a_i^J - a_e^J$, the influence of this part of the channel is compensated and further analysis can be accomplished under the DMC part only. It was demonstrated that in price of a certain decoder complexity increase that asymptotically vanishes with N, the identification capacity [3] over the DMC is achievable.

In the scope of this paper, we are advocating a less conservative decoding approach that was successfully applied in several applications [9], [10]. This decoding strategy is targeting estimation and compensation of the channel matrix A^J based on a pilot signal.

Estimation of desynchronization parameters. Assume that channel desynchronization distortions introduced during enrollment and identification are due to acquisition imperfections only and can be modeled within the class of generalized affine transforms. Then, possible ways of an accurate desynchronization estimation are to use a pilot signal that might be specifically designed and transmitted prior to the main communication session [10] or based on invariants, i.e., specific features deduced from the original data, that are robust with respect to a predefined class of desynchronization transforms. A possible example of such affine invariant features are SIFT image descriptors [11] that are the state of the art in matching and indexing. It is easy to realize that a strong link exists between both approaches since the extracted invariants can be considered as a pilot that is present within the data without any their additional alteration. This feature is extremely important for digital fingerprinting-based content identification, where no alteration of the data is tolerated in general.

Given the specified type of desynchronization, one has the

following desynchronization input/output relationship:

$$\begin{bmatrix} \acute{\theta}[1] \\ \acute{\theta}[2] \\ 1 \end{bmatrix} = \begin{bmatrix} a[1] & a[2] & 0 \\ a[3] & a[4] & 0 \\ a[5] & a[6] & 1 \end{bmatrix} \begin{bmatrix} \theta[1] \\ \theta[2] \\ 1 \end{bmatrix} + \begin{bmatrix} D[1] \\ D[2] \\ 1 \end{bmatrix}, \quad (4)$$

where the invariant with spatial Cartesian coordinates $[\theta[1]\ \theta[2]]$ is transformed to a new one with the coordinates $\acute{\theta}[1]\ \acute{\theta}[2]$ via a linear map (4). Here, $a[i], i \in \{1, 2, ..., 6\}$ are realizations of random parameters that define rotation, scaling and translation affine components. Finally, $D[1], D[2]$ are i.i.d. zero-mean Gaussian random variables with variance σ_D^2 that model data noise introduced due to acquisition and content manufacturing imperfections [12].

Rewriting (4) in a matrix form, one obtains for K invariants:

$$\acute{\Theta}^{3K} = A^{3K\times 3K} \cdot \Theta^{3K} + D^{3K}, \quad (5)$$

where $A^{3K\times 3K}$ is an $3K \times 3K$ block diagonal matrix with $\begin{bmatrix} a[1] & a[2] & 0 \\ a[3] & a[4] & 0 \\ a[5] & a[6] & 1 \end{bmatrix}$ on the main diagonal.

Then, the problem of the desynchronization parameters estimation can be formulated using a Maximum Likelihood (ML) approach:

$$\hat{A}^{3K\times 3K} = \arg\max_{a^6 \in \mathcal{A}^6} p(\acute{\Theta}^{3K} | \hat{A}^{3K\times 3K}, \Theta^{3K}) \quad (6)$$

$$= \arg\min_{a^6 \in \mathcal{A}^6} \left\| \acute{\Theta}^{3K} - A^{3K\times 3K} \cdot \Theta^{3K} \right\|^2. \quad (7)$$

The solution to the former minimization problem is given by:

$$\hat{A}_{ML}^{3K\times 3K} = A^{3K\times 3K} + \Xi^{3K\times 3K}, \quad (8)$$

where $\Xi^{3K\times 3K} = (D^{3K})^T \Theta^{3K} (\Theta^{3K} (\Theta^{3K})^T))^{-1}$, ^{T} defines matrix transposition operation and $\Xi^{3K\times 3K}$ is the estimation error matrix that for a sufficiently large K will have zero-mean Gaussian distributed rows with a covariance matrix defined as $\sigma_D^2 (\Theta^{3K} (\Theta^{3K})^T)^{-1}$, supposing that the inverse exists. Assuming orthogonality of Θ^{3K} according to the SIFT invariance design, one has that the elements of the estimation error matrix will be i.i.d. Gaussian with zero mean and variance $\sigma_D^2 \left\| \Theta^{3K} \right\|^2$.

Given the desynchronization matrix estimate, one can perform its approximate inverse at both enrollment and identification under an assumption that a common synchronizing pilot can be extracted from the data to be identified. Then, the output of the inverse at the identification stage is stored at the database and made available at identity verification.

Provided approximately desynchronization compensated enrolled database/codebook is observed at the identification stage together with a symmetrically transformed query, the decoder should decide which out of M enrolled objects is observed by the system or to reject authenticity if the system input is originated from the data that is irrelevant to the database/codebook content:

$$g : \mathcal{Z}^N \to \{\delta, 1, 2, ..., M\}, \quad (9)$$

where δ denotes the decoder output for the case when the data irrelevant to the analyzed at enrollment are observed by the system. Since the obtained estimate in not perfectly accurate, this deterministic mapping should be applied at all elements of the set of residual desynchronization parameters $a_r^J = a^J - \hat{a}_{ML}^J \in \mathcal{A}_r^J$.

It is known that the decoder (9) will produce an incorrect result with the following average probability [13]:

$$P_e^G = \frac{1}{M|\mathcal{A}_r^J|} \sum_{w=1}^{M} \sum_{a_i \in \mathcal{A}_r^J} \Pr[g(t_{a^{-1}}(t_{a_i}(Z^L))) \quad (10)$$

$$\neq w \mid t_{A^{-1}}(t_{A_e}(Y^L(w)) = t_{a^{-1}}(t_{a_e}(y^L(w))] \quad (11)$$

taking into account the cardinality of $\mathcal{A}_r^J$, defined in our case for a fixed rate of identification:

$$R_{id} = \frac{1}{N} \log_2 M. \quad (12)$$

Finally, capacity of the identification system C_{id} is defined as a supremum of the identification rates R_{id} such that $P_e \to 0$ for a sufficiently large N [13].

Theorem. The capacity C_{id} of an identification system operating over DMCs that are followed by respective desynchronization (2) that models acquisition distortions between the original fingerprint and (a) the database entry and (b) the query, respectively, is given by $I(Y; Z)$ in the case $N \to \infty$, where $p(y, z) = \sum_{x \in \mathcal{X}} p(x)p(y|x)p(z|x)$ for all $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$ and $I(Y; Z)$ stays for a mutual information between two random variables Y and Z.

The proof of this theorem follows a random coding-based strategy and jointly typical decoding applied at all points within $\mathcal{A}_r^J$ in the achievability part and is using Fano inequality [13], [14] in the converse part and is omitted in this paper.

III. Analysis of the database entry length impact on identification system design and performance

The analysis of the achievable rates of object identification in channels with desynchronization was accomplished using a capacity achieving argument that explicitly assumes infinite length of database entries. However, practical identity verification systems are aiming at providing optimal probabilities of miss-classification that should be attained for a codebook/database of a fixed and not necessary relevant to identification capacity cardinality [15] and finite codeword length ($N < \infty$). In order to justify how codeword length impacts these system performance measures, we formulate content identification problem as follows (Fig. 2).

Similarly to the previous Section, it is assumed that outputs of both enrollment $Y^N(w), w \in \{1, 2, ...M\}$, and identification Z^N are distorted and desynchronized versions of original raw fingerprints $X^N(w), w \in \{1, 2, ...M\}$ for the enrollment related case, and of X'^N in the opposite case.

Estimation and compensation of the introduced desynchronization is accomplished using the strategy presented earlier. However, instead of considering the entire $\mathcal{A}_r^6$ set (ML strategy), re-synchronization is performed in one trial based on $\hat{a}_{ii} = \hat{a}_{ii}^{3\times 3} = \frac{1}{K} \sum_{k=1}^{K} \hat{A}_{ML}^{3K\times 3K}[3k-2 : 3k, 3k-2 : 3k]$, where $ii \in \{e, i\}$ and e, i denotes enrollment and identification, respectively, that is usually referred to as mismatched metric [10].

It is assumed that to cope with mentioned privacy, complexity and storage memory requirements identification is performed based on binary data. Binary fingerprints are generated using the following two-stage procedure. First, given $t_{a_e}(y^N(w)), w \in \{1, 2, ..., M\}$, their reduced dimensionality $\tilde{y}^L(w)$ representations are generated using random projections:

$$\tilde{y}^L(w) = \mathbf{\Psi} t_{\hat{a}_e^{-1}}(t_{a_e}(y^N(w))) \tag{13}$$

at the enrollment stage. We assume that $\mathbf{\Psi} \in \mathbb{R}^{L \times N}$, where $L \leq N$ and $\mathbf{\Psi} = (\psi^N[1], \psi^N[2], \cdots, \psi^N[L])^T$, denotes a dimensionality reducing operator. It is supposed that the elements of $\mathbf{\Psi}$, $\psi_{i,j}$, are independently and identically generated from a Gaussian distribution, i.e., $\Psi[i,j] \sim \mathcal{N}(0, \frac{1}{N}), i \in \{1, 2, ..., L\}, j \in \{1, 2, ..., N\}$. The parameters of the generating distribution are adjusted to guarantee that $\mathbf{\Psi}$ is an approximate orthoprojector, i.e., $\mathbf{\Psi}\mathbf{\Psi}^T \approx \mathbf{I}_L$. The selection of the Gaussian distribution for $\mathbf{\Psi}$ is justified by a desired property of such a projection output to have the Gaussian statistics too. It should be admitted that such a property will be attributed to other statistical designs of $\mathbf{\Psi}$ for a sufficiently long input vectors due to the Central Limit Theorem.

Then, random projection outputs are converted to a length L binary fingerprint by extracting signs of $\tilde{y}^L(w)[i]$:

$$b_{\mathbf{y}_a}[i] = sign(\psi^N[i] t_{\hat{a}_e^{-1}}(t_{a_e}(y^N(w)))), i \in \{1, 2, ..., L\}, \tag{14}$$

where $\psi^N[i]$ stays for the ith row of $\mathbf{\Psi}$, and stored in the database.

The identification stage is organized as follows (Fig. 2). First, the output of the identification channel $t_{a_i}(Z^N)$ passes desysnchronization compensation based on $\hat{a}_i^{3\times 3}$. Then, to preserve symmetry with the enrollment stage, one applies:

$$\tilde{z}^L = \mathbf{\Psi} t_{\hat{a}_i^{-1}}(t_{a_i}(z^N)), \tag{15}$$

$$b_{\mathbf{z}_a}[i] = sign(\psi_i^T t_{\hat{a}_i^{-1}}(t_{a_i}(z^N))), i \in \{1, 2, ..., L\}. \tag{16}$$

It is possible to demonstrate that under rather mild conditions on X^N, binary fingerprints generated at enrollment and identification will have a Binomial distribution with 0.5 probabilities of both binary events due to particularities of the fingerprint extraction [16].

We include in the setup analysis of this Section the cases of system irrelevant inputs X'^N since such situations are not rare in identification setups. These inputs were not analyzed in details in the previous Sections due to capability of the jointly typical decoder to reliably eliminate them from consideration with high probability for $N \to \infty$ [13].

Therefore, one can formulate the multimedia object identification problem in channels with desynchronization as a multiple hypothesis testing problem with corresponding prior probabilities defined by an equivalent Binary Symmetric Channel (BSC) with a crossover probability P_{b_e} [8]:

$$\begin{cases} H_0: & B_{\mathbf{y}}^L \sim \frac{1}{2^L}, \\ H_w: & B_{\mathbf{y}}^L \sim P_{b_e}^{d_H(b_{\mathbf{z}_a}^L, b_{\mathbf{y}_a}^L(w))}(1-P_{b_e})^{L-d_H(b_{\mathbf{z}_a}^L, b_{\mathbf{y}_a}^L(w))}, \end{cases} \tag{17}$$

where $w \in \{1, 2, ..., M\}$ and $d_H(.,.)$ is the Hamming distance.

Having access to the database content and $b_{\mathbf{z}_a}^L$, the decoder should decide which one out of $M+1$ alternatives is present at the input of the identification system. We assume that it operates according to the Bounded Distance Decoding (BDD):

$$d_H(b_{\mathbf{z}_a}^L, b_{\mathbf{y}_a}^L(w)) \leq L\gamma, \tag{18}$$

defined for a certain threshold γ. The optimal selection of γ for the BSC distortion model case was considered in [17]. Using the equal error rate probability of error analysis, it was demonstrated that the optimal value of the threshold is a function of the channel:

$$\gamma_{\text{opt}} = \frac{1 - R_{id} + \log_2(1-P_{b_e})}{\log_2\left(\frac{1-P_{b_e}}{P_{b_e}}\right)}. \tag{19}$$

According to [17], the threshold value that determines the performance of the identification system in terms of probabilities of error coincides with the classical BDD.

Similarly to [8], we analyze probabilities of error of two kinds. For the probability of false acceptance P_f one has:

$$\begin{aligned} P_f &= \Pr[\bigcup_{w=1}^{M} d_H(b_{\mathbf{y}_a}^L(w), b_{\mathbf{z}_a}^L) \leq \gamma L | H_0] \\ &\leq 2^{-L(1-H_2(\gamma)-R_{id})}, \end{aligned} \tag{20}$$

for a fixed P_{b_e} and due to the union bound and the Chernoff bound application on the tail of binomial distribution $\mathcal{B}(L, 0.5)$ [17]. Similarly, one can bound the probability of incorrect identification P_{ic} in the following way:

$$\begin{aligned} P_{ic} &= \Pr[d_H(b_{\mathbf{y}_a}^L(w), b_{\mathbf{z}_a}^L) > \gamma L \\ &\cup \bigcup_{s \neq w}^{M} d_H(b_{\mathbf{y}_a}^L(s), b_{\mathbf{z}_a}^L) \leq \gamma L | H_w] \\ &\leq 2^{-L(D(\gamma || P_{b_e}))} + 2^{-L(1-H_2(\gamma)-R_{id})}. \end{aligned} \tag{21}$$

where $D(\gamma || P_{b_e}) = \gamma \log_2 \frac{\gamma}{P_{b_e}} + (1-\gamma)\log_2 \frac{1-\gamma}{1-P_{b_e}}$ and the result is justified by the same arguments as above.

Finally, it is easy to show by minimizing average probability of error based on (20) and (21) that γ_{opt} coincides with (19) and in the capacity achieving regime case is equal to P_{b_e}.

However, the desynchronization matrix estimation error $\Xi^{3K\times 3K}$ (8) breaks the assumption of the constant P_{b_e} used in the analysis of P_f, P_{ic} and γ_{opt} and converts the considered BSC model with a fixed P_{b_e} to the BSC with a random crossover probability that is distributed according to a certain $p(P_{b_e} = p)$ that depends on the accuracy of the desynchronization estimation/compensation.

Justifying the impact of $\Xi^{3K\times 3K}$ on the BSC crossover probability by the following mapping:

$$\phi : \mathcal{A}_r^6 \to [0, 0.5], \tag{22}$$

the bounds on the probabilities of error become:

$$P_f \leq 2^{-L(1-H_2(\gamma'_{opt})-R_{id})}, \tag{23}$$

$$P_{ic} \leq 2^{-L(D(\gamma'_{opt}||P_{b_e}))} + 2^{-L(1-H_2(\gamma'_{opt})-R_{id})}, \tag{24}$$

for $\gamma'_{opt} = \max_{[0,0.5]} P_{b_e} = 0.5$ and no reliable communications will be possible in the worst case. However, if ϕ maps to an interval with the upper limit $P_{b_e}^{max} < 0.5$,

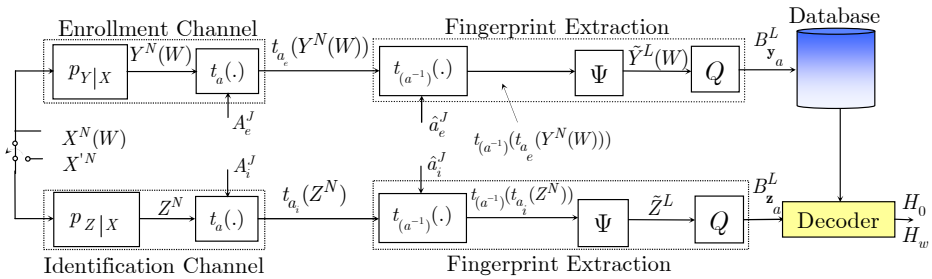

Fig. 2. Multimedia object identification as a communications problem: finite length codeword case.

reliable communication will be still possible but with a rate $R_{id} < 1 - H_2(\gamma'_{opt}) = 1 - H_2(P_{b_e}^{max})$.

IV. Experimental validation

Our experiments were dedicated to the analysis of the desynchronization impact on the crossover probability P_{b_e} of the equivalent identification BSC. Here, it is assumed that the original data of 32×32 cardinality are generated i.i.d. from a standard Gaussian distribution and from a separable 2D AR(1) process with $\theta = 0.95$. It was assumed that prior to enrollment and identification these data were corrupted by the AWGN channel with the variance $10^{-2.5}$ and affine desynchronized with the residual rotation and scale distributed on the intervals $\alpha \in [-0.5, 0.5]$ degree with a step 0.1 degree and $\rho \in [0.9, 1.1]$ with a step 0.02, respectively. 10000 realizations of the AWGN were used for every rotation angle/scale pair. Binary fingerprints of length 32 were extracted from the imperfectly re-synchronized data using random projections and binarization according to Section III. The obtained experimental results allow to conclude that uncompensated affine desynchronization significantly increase maximal P_{b_e} in both i.i.d. (from 0.018 to 0.42) and AR(1) (0.0027 to 0.16) cases while keeping the equivalent crossover probability below 0.5.

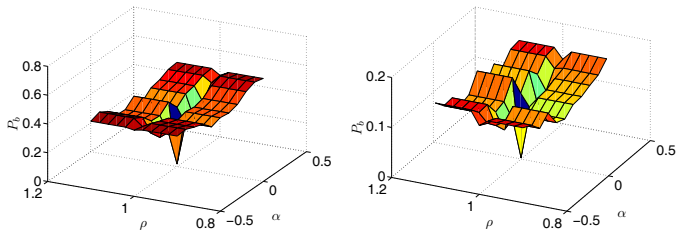

Fig. 3. P_{b_e} as a function of re-synchronization accuracy: (a) Gaussian i.i.d. case and (b) 2D AR(1) case.

V. Conclusions

In this paper we considered the problem of decoding complexity constraint identification in channels with desynchronization. We have demonstrated that application of the channel estimation/compensation strategy allows reduce decoding complexity in the identification capacity achieving regime. The analysis of practical finite-length identification based on missmatched metric revealed a strong dependence of the attained probabilities of error on the resynchronization accuracy that might even lead in the worst case to $R_{id} = 0$. Experimental results confirm our theoretical findings.

Our future work will be dedicated to the development of practical desynchronization resilient identification schemes.

Acknowledgment

This paper was partially supported by SNF project 200020-134595 and CRADA project.

References

[1] J. Haitsma, T. Kalker, and J. Oostveen, "Robust audio hashing for content identification," in *Proc. Int. Workshop on Content-Based Multimedia Indexing*, pp. 117–125, (Brescia, Italy), September 2001.

[2] F. Lefebvre and B. Macq, "Rash : RAdon Soft Hash algorithm," in *Proceedings of EUSIPCO 2002*, (Toulouse, France), 2002.

[3] F. Willems, T. Kalker, J. Goseling, and J.-P. Linnartz, "On the capacity of a biometrical identification system," in *Proc. 2003 IEEE Int. Symp. Inform. Theory*, p. 82, (Yokohama, Japan), June 29 - July 4 2003.

[4] O. Koval, S. Voloshynovskiy, F. Beekhof, and T. Pun, "Decision-theoretic consideration of robust perceptual hashing: link to practical algorithms," in *Proc. WaCha2007*, (Saint Malo, France), 2007.

[5] S. Voloshynovskiy, O. Koval, F. Beekhof, and T. Pun, "Robust perceptual hashing as classification problem: decision-theoretic and practical considerations," in *Proc. of the IEEE 2007 Int. Workshop on Multimedia Sig. Proc.*, (Chania, Crete, Greece), October 1–3 2007.

[6] A. L. Varna, A. Swaminathan, and M. Wu, "A decision theoretic framework for analyzing hash-based content identification systems," in *ACM Digital Rights Management Workshop*, pp. 67–76, Oct. 2008.

[7] O. Koval, S. Voloshynovskiy, F. Farhadzadeh, T. Holotyak, and F. Beekhof, "Information-theoretic analysis of desynchronization invariant object identification," in *Proc. ICASSP 2011*, May 22–27 2011.

[8] O. Koval, S. Voloshynovskiy, F. Farhadzadeh, T. Holotyak, and F. Beekhof, "Geometrically robust perceptual fingerprinting: an asymmetric case," in *Proc. SPIE 2011*, January 23–27 2011.

[9] S. Pereira and T. Pun, "Robust template matching for affine resistant image watermarks," *IEEE Trans. on Im. Proc.* **9**, pp. 1123–1129, 2000.

[10] G. Taricco and E. Biglieri, "Space-time decoding with imperfect channel estimation," *IEEE Trans. on Wir. Comm.* **4**, pp. 1874–1888, 2005.

[11] D. Lowe, "Object recognition from local scale-invariant features," in *Proc. ICCR 1999*, pp. 1150–1157, 1999.

[12] V. M. Govindu and M. Werman, "On using priors in affine matching," *Image and Vision Computing* **22**(14), pp. 1157–1164, 2005.

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

[14] T. Ignatenko and F.M.J. Willems, "Secret-key and identification rates for biometric identification systems with protected templates," in *The thirty-first symposium on Information Theory in the Benelux*, Rotterdam, The Netherlands, May 11-12 2006, pp. 121–128.

[15] F. Beekhof, S. Voloshynovskiy, O. Koval, and T. Holotyak, "Fast identification algorithms for forensic applications," in *Proc. IEEE IFS 1999*, (London, UK), 2009.

[16] F. Farhadzadeh, S. Voloshynovskiy, O. Koval, T. Holotyak, and F. Beekhof, "Statistical analysis of digital fingerprinting based on random projections," in *Proc. IEEE ICIP 2011*, September 11–14 2011.

[17] S. Voloshynovskiy, O. Koval, F. Beekhof, F. Farhadzadeh, and T. Holotyak, "Information-theoretical analysis of private content identification," in *Proc. IEEE Inf. Theory Workshop (ITW 2010)*, (Dublin, Ireland), August,30 - September, 3 2010.