Active Content Fingerprinting Using Latent Data Representation, Extractor and Reconstructor

Dimche Kostadinov
Department of Computer Science
University of Geneva
Geneva, Switzerland
dimche.kostadinov@unige.ch

Sviatoslav Voloshynovskiy
Department of Computer Science
University of Geneva
Geneva, Switzerland
svolos@unige.ch

Sohrab Ferdowsi
Department of Computer Science
University of Geneva
Geneva, Switzerland
sohrab.ferdowsi@unige.ch

***Abstract*—This paper introduces a concept of Active Content Fingerprinting based on a Latent data Representation (aCFP-LR). The idea is to represent the data content by a constrained redundant description. The target is to estimate latent representation such that: (i) after applying a reconstructor function the result is close to the original data and (ii) after using an extraction function the resulting features are robust.**

A general problem formulation is proposed for aCFP-LR with an extractor-reconstructor pair of constraints. One particular case is considered under linear extractor (generator) and linear reconstructor (modulator) where a reduction is shown to a constrained projection problem.

Evaluation by numerical experiments is given using local image patches, extracted from publicly available data sets. Advantages and state-of-the-art performance is demonstrated under additive white Gaussian noise (AWGN), lossy JPEG compression and projective geometrical transform distortions.

***Index terms*—** active content fingerprint, latent representation, extractor, reconstructor, redundancy, robustness.

I. Introduction

Active Content Fingerprinting (aCFP) has emerged as a synergy between the digital watermarking (DWM) and passive content fingerprinting (pCFP) [1]. This alternative approach covers a range of applications in the case when content modulation is appropriate, prior to the content distribution/reproduction such as content authentication, identification and recognition.

It was also theoretically demonstrated that the identification capacity of aCFP [2] under the additive white Gaussian channel distortions and ℓ_2-norm embedding distortion is considerably higher to those of DWM and pCFP. Interestingly, the optimal modulation of aCFP produces the correlated modulation to the content in contrast to the optimal modulation of DWM where the watermark is independent to the host. Several scalar and vector modulation schemes for the aCFP were proposed [3], [4] and tested on synthetic signals and collections of images. Despite of the attractive theoretical properties of aCFP, the practical implementation of aCFP modulation with an acceptable complexity, capable to jointly withstand *signal processing* distortions such as additive white Gaussian noise (AWGN), lossy JPEG compression, histogram modifications, etc. and *geometrical distortions* (affine and projective transforms) remains an open and challenging problem.

On the other hand in the recent years, local, i.e., *patch-based*, compact, geometrically robust, binary descriptors such as SIFT [5], BRIEF [6], BRISK [7], ORB [8] and the family of LBP [9] have become a popular tool in image processing, computer vision and machine learning. These local descriptors are also considered as a form of local pCFP.

A. Prior Work

Up to our best knowledge, there is little prior work on the modulation of local descriptors in the scope of aCFP or DWM.

In [10] an aCFP was proposed with a linear modulation subject to convex constraint on the properties of the resulting local descriptors and the optimal solution was given when the feature map is invertible. The main open issues with the proposed optimal solution are related to the assumptions about the linear feature map.

The authors in [11] addressed the general case from two distinct perspectives. Firstly, they proposed a direct approximation of the linear feature map and secondly they presented a novel problem formulation for the linear modulation and the constraints on the properties of the resulting local descriptor. In the former case, the used linear map is predefined and analytic, therefore the open issues are related to the properties of the used linear map that are crucial for the achievable modulation distortion and the resulting feature descriptor.

To reduce the modulation distortion and explicitly regularize the features properties, [12] proposed the joint learning of a linear feature map and linear aCFP modulation. The authors provided a novel problem formulation and gave iterative alternating algorithm with convergence result.

B. Contributions

This paper proposes a concept of aCFP-LR to estimate a constrained representation that describes the content. The motivation is to use redundancy in the latent representation in order to be robust to noise. In general, this means that once the latent representation is perturbed by the noise (i) the reconstructor function should recover the original data, and (ii) the extractor function should provide the original features. The general scheme is shown in Figure 1.

Considering an extension of the concept of aCFP, this paper focuses only on one specific case with two constraints. The

ISBN 978-90-827970-1-5 © EURASIP 2018

first one is (ii) mentioned in the paragraph above under noise perturbations of the latent representation. The second is similar to (i), but, now the reconstructor function is applied only to the latent description and should provide the original data. The latter is also known as modulation.

Nonetheless, the advantage is that the redundancy adds one more element to compensate the trade-off between modulation distortion and feature robustness. In other words, we try to add redundancy to achieve small modulation distortion and high feature robustness.

The contributions of this paper are the following, we:

(i) introduce an extension of the core principle of active content fingerprinting by focusing on latent representation estimation with constraints imposed by the extractor and reconstructor functions pair,

(ii) propose a generalized problem formulation with explicit regularization of the trade-off between the distortion and the robustness of the local feature by considering constraints on the distribution of (i) the data modifications, (ii) feature modifications and (iii) on the actual latent representation of the content and

(iii) validate the proposed approach by a computer simulation using publicly available data under several image processing distortions, including AWGN, lossy JPEG compression, and projective geometrical transform.

C. Paper Organization

The organization of the paper is as follows. Section 2 introduces the approach, gives a short description of the aCFP and highlights the difference between aCFP and aCFP-LR. Section 3 presents the general problem formulation and under linear modulation and linear feature maps, shows a reduction to a low complexity constrained projection problem. Section 4 is devoted to computer simulation and Section 5 concludes the paper.

II. aCFP-LR vs aCFP

The aCFP framework consists of content modulation, prior to its reproduction and descriptor extraction that includes feature mapping and quantization. The core idea behind the aCFP modulation [3] and [10] is based on the observation that the magnitude of the feature coefficients before the quantization influences the probability of the bit error in the descriptor bits.

Descriptor bit flipping is more likely for low magnitude coefficients. Therefore, it is natural to modify the original content by an appropriate modulation and to increase these magnitudes subject to distortion constraints.

The main idea behind the proposed aCFP-RL is to produce a resilient to noise data representation such that after applying a feature generator function the resulting features are robust. At the same time, the reconstructor (modulation) function applied on the latent representation should give the original data.

Two operational modes are considered: (i) modulation and (ii) verification. The modulation estimates the latent representations that describe the content. During verification, the features from the noise perturbed latent representation are extracted and the fingerprint is computed.

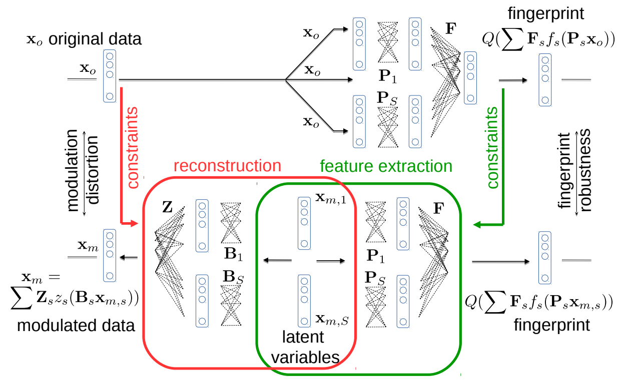

Fig. 1. A general scheme of aCFP-LR using a latent representation, extractor, and reconstructor functions.

A. pCFP, aCFP and aCFP-LR: Feature Generation

A shared component in pCFP, aCFP, and aFP-LR is the feature extraction.

Assume that from the original image is obtained a patch, denoted as $\mathbf{x}_o \in \Re^N$. This paper considers a generalized feature compositional case (with or without nonlinearity), where the extraction of the local features is defined as follows:

$$\mathbf{f}_o = \sum_s \mathbf{F}_s f_s(\mathbf{P}_s \mathbf{x}_o), \qquad (1)$$

$\mathbf{P}_s \in \Re^{M \times N}$ and $\mathbf{F} = [\mathbf{F}_1, ..., \mathbf{F}_S], \mathbf{F}_s \in \Re^{L \times M}$ are linear maps and $f_s : \Re^{M \times 1} \to \Re^M$ are functions describing element-wise nonlinearity, L, M and N are the lengths of the final, intermediate and input data representation, respectively, and $s \in \{1, .., S\}$. Note that the map can be either predefined, data independent and analytic or learned, data dependent and content adaptive.

The feature extraction is followed by a quantization $Q(.)$ that results in a quantized local descriptor denoted as $\mathbf{b}_o = Q(\mathbf{f}_o) \in \{0, 1\}^L$, where

$$Q(f_o(i)) = \begin{cases} 1, & \text{if } f_o(l) \geq 0, \\ 0, & \text{if } f_o(l) < 0, \end{cases} \quad \forall l \in \{1, ..., L\}. \qquad (2)$$

Note that the differences between the existing classes of local descriptors are determined by the defined mapping (1) and the type of quantization $Q(.)$.

B. aCFP-LR: Reconstruction from Latent Variables

In aCFP, the concept is based around the modulated data $\mathbf{x}_m \in \Re^N$ that should be close to the original data representation $\mathbf{x}_o$. The aCFP-LR takes into account the reconstruction from S latent representations $\mathbf{x}_{m,s} \in \Re^L$, $s \in \{1, ..., S\}$ to $\mathbf{x}_o$. Assuming that the S latent variables $\mathbf{x}_{m,s}$ are given then a general reconstruction function is defined as follows:

$$\mathbf{x}_m = \sum_s \mathbf{Z}_s z_s(\mathbf{B}_s \mathbf{x}_{m,s}), \qquad (3)$$

where $\mathbf{B}_s \in \Re^{M \times L}$ and $\mathbf{Z} = [\mathbf{Z}_1, ..., \mathbf{Z}_S], \mathbf{Z}_s \in \Re^{N \times M}$ are linear maps and $z_s : \Re^{M \times 1} \to \Re^M$ are functions describing an element-wise nonlinearity.

ISBN 978-90-827970-1-5 © EURASIP 2018

III. aCFP-LR Problem Formulation

This paper presents a general problem formulation for estimation of the underlying data representation that describes the content. The corresponding optimization problem is the following:

$$\min_{\mathbf{h}_m} \varphi\left(v\left(\mathbf{h}_m\right), \mathbf{x}_o\right) + \lambda_1 \psi\left(g\left(\mathbf{h}_m\right), \tau\right) + \lambda_2 r\left(\mathbf{h}_m\right), \quad (4)$$

where $\mathbf{x}_o$ is the original data, $\mathbf{h}_m = [\mathbf{x}_{m,1}, .., \mathbf{x}_{m,S}]$ is the *latent data representation*, $v(\mathbf{h}_m)$ is the *reconstructor function*, $g(\mathbf{h}_m)$ is the *generator function* and the function $r(\mathbf{h}_m)$ imposes constraints on the properties of $\mathbf{h}_m$. The modulation level and the Lagrangian variables are denoted as τ, λ_1 and λ_2, respectively.

The first mapping function $v(\mathbf{h}_m)$ is the reconstructor function that is applied to the latent representation $\mathbf{h}_m$ in order to match $v(\mathbf{h}_m)$ to the original data representation $\mathbf{x}_o$, where $\varphi\left(v\left(\mathbf{h}_m\right), \mathbf{x}_o\right)$ is a function that penalizes the distortions in the original data domain.

The second function $g(\mathbf{h}_m)$ transforms $\mathbf{h}_m$ into features and tries to make $g\left(\mathbf{h}_m\right)$ robust, where $\psi\left(g\left(\mathbf{h}_m\right), \tau\right)$ is a function that penalizes non-robust feature components.

In general (4) represents one form of a min-max problem. Moreover, note that (4) is also one case of the formulation proposed in [10]. However, in this manuscript the target is not the actual data, but, rather the latent representation.

A. Linear Generator and Modulator

This paper addresses (4) by considering a setup of linear modulation under assumptions given as below.

1) Reconstructor and Generator Distortions: The function $\varphi\left(v\left(\mathbf{h}_m\right), \mathbf{x}_o\right)$ is defined as $\varphi\left(v\left(\mathbf{h}_m\right), \mathbf{x}_o\right) = \|\mathbf{x}_o - v\left(\mathbf{h}_m\right)\|_2^2$. The function $\psi\left(g\left(\mathbf{h}_m\right), \tau\right)$ is replaced by an explicit inequality constraint $|\mathbf{A}_F \mathbf{h}_m| \geq_e \tau \mathbf{1}$ where $\mathbf{A}_F \in \Re^{L \times SN}$ and $\geq_e$ represents an element-wise inequality.

It is assumed that there is no element-wise nonlinearity, *i.e.*, $z_s(\mathbf{x}_{m,s}) = \mathbf{x}_{m,s}$ and no specific function $r(\mathbf{h}_m)$ is defined other then the explicit inequality constraint $|\mathbf{A}_F \mathbf{h}_m| \geq_e \tau \mathbf{1}$.

2) Linear Generator (Feature Extraction): The feature extraction function is defined as the linear version of (1). Assuming the modulated data $\mathbf{x}_{m,s}$ are given, the features are extracted as follows:

$$\mathbf{f}_m = g(\mathbf{h}_m) = \sum_s \mathbf{F}_s \mathbf{P}_s \mathbf{x}_{m,s}, \quad (5)$$

where the linear maps $\mathbf{P}_s$ are generated at random. The linear maps $\mathbf{F}_s$ are defined using a constraint matrix, data samples (or their clusters) and random sampling. We describe the construction of the linear maps $\mathbf{F}_s$ in the following.

Let $\mathbf{C} \in \{-1, 0, +1\}^{L \times M}$ be the matrix that encodes the m-wise (pair-wise, triple-wise, etc.) constraints that describe the geometrical configuration of the considered data (pixel) interactions. Given a data set of patches, assume that M centroids $[\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_S]$ are estimated using a k-means algorithm. Denote a transformation matrix as $\mathbf{T}_s \in \Re^{M \times N}$, where $\mathbf{T}_s = \mathbf{R}\left[\mathbf{a}_s \mathbf{a}_s^T\right]^{-1} \in \Re^{N \times N}$ and $\mathbf{R} \in \Re^{N \times N}$ is a randomly constructed matrix. First the matrix $\mathbf{a}_s \mathbf{a}_s^T$ is quantized to J levels, where for every quantization level $q \in \{1, 2, 3, ..., J\}$, there exists a set $\mathcal{L}_q$ of indexes to the elements in $\mathbf{a}_i \mathbf{a}_s^T$, all of the $\mathcal{L}_q$ are with the same cardinality. Then for every index set $\mathcal{L}_q$ the corresponding elements of $\mathbf{R}$ are generated from a uniform distribution with support $[0, 1]$. The main idea is to try to have an equal contribution of the elements of $\mathbf{a}_s \mathbf{a}_s^T$ in the linear feature map $\mathbf{C}\mathbf{T}_s$.

Finally, the linear maps $\mathbf{F}_s$ are estimated as follows:

$$\mathbf{F}_s = \left(\mathbf{U}\begin{bmatrix} \mathbf{\Sigma}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{V}^T\right)^T, \quad (6)$$

where $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ is the SVD of $(\mathbf{C}\mathbf{T}_s)^T$ and $\mathbf{\Sigma}_p$ is a diagonal matrix having p non-zero diagonal elements equal to the largest p singular values from $\mathbf{\Sigma}$. The value of p determines the rank of $\mathbf{F}_s$.

3) Linear Reconstruction (Modulator): Let $\mathbf{B}_s = \mathbf{I}$ and $\mathbf{Z}_s = \mathbf{I}$, then the simplest modulation on multiple latent variables is defined as:

$$\mathbf{x}_m = v(\mathbf{h}_m) = \sum_s z_s(\mathbf{x}_{m,s}). \quad (7)$$

Note that by using (7) and $\varphi\left(v\left(\mathbf{h}_m\right), \mathbf{x}_o\right)$ the modulation and the reconstruction are seen as equivalent.

The following presents the problem formulation under the previous assumptions $1, 2$ and 3.

Proposition 1: *The aCFP-LR under constraints* $1, 2$ *and* 3 *with linear modulation, linear feature map and convex constraints on the properties of the latent representations is a constrained projection problem*:

$$\mathbf{h}_m = \arg\min_{\mathbf{h}_m} \frac{1}{2}\|\mathbf{x}_o - [\mathbf{I}_{N \times N}, ..., \mathbf{I}_{N \times N}]\mathbf{h}_m\|_2^2 \quad (8)$$
$$\textbf{subject to } |\mathbf{A}_F \mathbf{h}_m| \geq_e \tau \mathbf{1},$$

where $\mathbf{I}_{N \times N}$ is an $N \times N$ identity matrix, $\mathbf{A}_F = [\mathbf{A}_1, ..., \mathbf{A}_{S-1}, \mathbf{0}_{L \times N}]$, $\mathbf{0}_{L \times N}$ is a zero matrix with dimensions $L \times N$, $\mathbf{h}_m = \begin{bmatrix} \mathbf{x}_{m,1} \\ \vdots \\ \mathbf{x}_{m,S} \end{bmatrix}$, $\mathbf{x}_{m,s} \in \Re^N$ are the latent representations, $\mathbf{A}_s = \mathbf{F}_s \mathbf{P}_s$ are the linear feature maps and $s \in \{1, ..., S-1\}$.

The goal is to allow an arbitrary large distortion between $\mathbf{x}_o$ and $\mathbf{x}_{m,i}$, but, very small distortion between $\mathbf{x}_o$ and $\sum_{s=1}^{S} \mathbf{x}_{m,s}$ and robust features $\mathbf{f}_m = [\mathbf{A}_1, ..., \mathbf{A}_{S=1}]\begin{bmatrix} \mathbf{x}_{m,1} \\ \vdots \\ \mathbf{x}_{m,S-1} \end{bmatrix}$ that would satisfy the constraints in (8). The key is to estimate all $\mathbf{x}_{m,s}$ and use every modulated component $\mathbf{x}_{m,s}$ independently, rather then the original data $\mathbf{x}_o$. In this way the additional element that is added to the concept of aCFP is the redundancy.

In this case the trade-off is between modulation distortion, feature robustness and the amount of redundancy.

B. Verification

Only the first $S - 1$ latent representations $\mathbf{x}_{m,s}$ from $\mathbf{h}_m$ are involved in the constraint in (8). The inequality $|\left[\mathbf{A}_1, ..., \mathbf{A}_{S-1}, \mathbf{0}_{L \times N}\right]\mathbf{h}_m| \geq_e \tau \mathbf{1}$ enforces these first $S - 1$

latent representations $\mathbf{x}_{m,s}$ to be sparse. There is no constraint on the variable $\mathbf{x}_{m,S}$ that appears in the cost function $\|\mathbf{x}_o - \sum_s \mathbf{x}_{m,s}\|_2^2$ of (8). Therefore, $\mathbf{x}_{m,S}$ is not sparse and it will always have a larger ℓ_2-norm than the ℓ_2-norm of the rest $\mathbf{x}_{m,s}, s \in \{1, ..., S-1\}$.

This is important at the verification stage. Since even if we have all $\mathbf{x}_{m,s}$ but we do not know the component $\mathbf{x}_{m,S}$, if we have the prior knowledge that the ℓ_2-norm of $\mathbf{x}_{m,S}$ is greater then any of the ℓ_2-norms of the rest of $\mathbf{x}_{m,s}$ then the component $\mathbf{x}_{m,\hat{s}}$ can be simply estimated by computing the following:

$$\hat{s} = \arg \max_{1 \leq s \leq S} \|\mathbf{x}_{m,s}\|_2. \tag{9}$$

At the verification stage, it is assumed that noise is independently added to every modulated component $\mathbf{x}_{m,s}$. Afterwards, the noisy modulated component is $\mathbf{y}_{m,s} = \mathbf{x}_{m,s} + \mathbf{x}_{\mathcal{N},s}$ and the fingerprint is estimated as:

$$\mathbf{b}_y = Q\left(\sum_{s=1}^{S-1} \mathbf{A}_s \mathbf{y}_{m,s}\right). \tag{10}$$

IV. Computer Simulations

This section validates the proposed approach by numerical experiments and demonstrates the advantages of the aCFP-LR. The performance is evaluated under several signal processing distortions, including AWGN, lossy JPEG compression and the projective geometrical transform. The results are compared with those from the pCFP, aCFP and aCFPL schemes.

The UCID [13] image database was used to extract local image patches. The ORB detector [8] was run on all images, and $\sqrt{N} \times \sqrt{N}$ pixel patches, $\sqrt{N} = 31$ were extracted around each detected feature point. The features were sorted by scale-space, 30 patches were used from each individual image.

A. Lenear Feature Maps

Given the extracted image patches, pCFP, aCFP, aCFPL and aCFP-LR used the following linear maps for feature extraction:

1) pCFP and aCFP: Both use $\mathbf{A}_1$. It is defined as $\mathbf{A}_1 = \left(\mathbf{U}\mathbf{I}_{L\times M}\mathbf{V}^T\right)^T$ where $\mathbf{U}, \mathbf{V}$ are obtained by singular value decomposition (SVD) of $(\mathbf{C}\mathbf{T}_1)^T$. The matrix $\mathbf{T}_1 = \mathbf{R}\left[\mathbf{x}_o\mathbf{x}_o^T\right]^{-1} \in \Re^{N\times N}$ and $\mathbf{R} \in \Re^{N\times N}$ is a random matrix, generated from a uniform distribution with the support $[0, 1]$. The matrix $\mathbf{A}_1^T$ is the closest orthogonal to $(\mathbf{C}\mathbf{T})^T$, satisfying $\mathbf{A}_1^T\mathbf{A}_1 = \mathbf{I}$ and easily invertible.

2) aCFPL: It uses the matrix $\mathbf{A}_L$ that is learned from the half of the total available patches. The remaining half is used for evaluation.

3) aCFP-LR: It uses the matrx $\mathbf{A} = [\mathbf{A}_1, ..., \mathbf{A}_s] = [\mathbf{F}_1\mathbf{P}_1, ..., \mathbf{F}_S\mathbf{P}_S]$, where $\mathbf{F}_s$ and $\mathbf{P}_s$ are defined as in Section III-A, $s \in \{1, .., S-1\}$ and the number of redundant representations is set to $S = 12$. The solution of (8) is found using the CVX [14] publicly available solver.

B. Noise Distortions

The following distortions are simulated:

1) AWGN: The results from a single patch were obtained as the average of 100 AWGN realizations. Four different noise levels were used, defined in PSNR= $10\log_{10}\frac{255^2}{\sigma^2}$ are 0dB, 5dB, 10dB and 20dB. The used modulation level (mL) is 60 for pCFP, aCFP and aCFPL, and $mL = 100$ for aCFP-LR. The results are shown in Table I.

2) Lossy JPEG Compression: Three JPEG quality factor (QF) levels 0, 5 and 10 were used. The used modulation level (mL) is 30 for pCFP, aCFP and aCFPL, and $mL = 100$ for aCFP-LR. The results are shown in Table II.

3) Projective Transform with Lossy JPEG Compression: A projective transformation $\mathbf{P} \in \Re^{3\times 3}$ was used, where:

$$\mathbf{P} = \begin{bmatrix} 1.0763 & 0.0325 & 0 \\ 0.0119 & 1.09 & 0 \\ -24.32 & -70.37 & 1 \end{bmatrix},$$

followed by a lossy JPEG compression with QF=5. The used modulation level is 60 for pCFP, aCFP and aCFPL, and $mL = 100$ for aCFP-LR. The results are shown in Table III.

C. Measures

Three measured quantities are used in the evaluation:

1) Modulation Level: Define $\mathbf{t}_o = \mathbf{A}_T\mathbf{x}_o = \left(\sum_{s=1}^{S-1}\mathbf{A}_s\right)\mathbf{x}_o$, where $\mathbf{A}_T = \left(\sum_{s=1}^{S-1}\mathbf{A}_s\right)$ then let $\mathbf{s}_o, |s(i)| \leq |s(j)|, \forall i \leq j, i, j \in \{1, 2, 3, ..., L\}$ be a sorted $|\mathbf{t}_o|$ vector and let $\mathbf{A}_{T1}$ be rows reordered $\mathbf{A}_T$ such that $\mathbf{A}_{T1}\mathbf{x}_o = \mathbf{s}_o$. The modulation level mL is defined in percentage $mL = \frac{K}{L}100, 1 \leq K \leq L$ and it represents the fraction of coefficients $\mathbf{s}_o$ that are modified. At a single modulation level, the modulation threshold τ for the aCFP-LR method is defined as $\tau = 100\max_{1\leq i\leq K}|s_o(i)|$, for the aCFP and aCFPL $\tau = \max_{1\leq i\leq K}|s_o(i)|$.

2) Modulation Distortion: The modulation distortion for aCFP-LR is defined as $DWR = 10\log_{10}\left(\frac{255^2}{\Delta^2}\right)$, $\Delta = \frac{1}{N}\|\mathbf{x}_o - \sum_s \mathbf{x}_{m,s}\|_2$ for aCFP and aCFPL is defined as $DWR = 10\log_{10}\left(\frac{255^2}{\Delta^2}\right)$, $\Delta = \frac{1}{N}\|\mathbf{x}_o - \mathbf{x}_m\|_2$.

3) Probability of Bit Error: The probability of bit error defined by the probability of correct bit $p_e = 1 - p_c$, $p_c = \frac{1}{L}\sum_{i=1}^{L} I\{b_x(i), b_y(i)\}$ with $L = 256$ bits, where $\mathbf{b}_x = Q((\sum_s \mathbf{A}_s)\mathbf{x}_o)$, $\mathbf{b}_y = Q\left(\sum_s \mathbf{A}_s(\mathbf{x}_{m,s} + \mathbf{x}_{\mathcal{N},s})\right)$, $\mathbf{x}_{\mathcal{N},s}$ is the introduced distortion and I is an indicator function $I\{a, b\} = 1$ if $a = b$ and $I\{a, b\} = 0$ otherwise.

D. Results

The provided evaluation takes into account an average results for a total of 500 image patches.

The results in Tables I, II and III show that the proposed approach has a superior performance in terms of probability of bit error and modulation distortion compared to pCFP, aCFP and aCFPL.

In Table IV are presented the results considering extreme cases of noise. Even under severe, very high noise levels the aCFP-LR achieves a small probability of error and small modulation distortion.

ISBN 978-90-827970-1-5 © EURASIP 2018

In summary, this evaluation shows that using the proposed approach it is possible to achieve a small modulation error and very high robustness on the features, even under severe noise levels. However, this performance is at the cost of adding redundancy in the underlying representation and the ability to differentiate $\mathbf{x}_{m,S}$ from the rest $\mathbf{x}_{m,s}, s \in \{1, ..., S-1\}$ under noise perturbations.

V. Conclusion

This papers introduced the concept of aCFP-LR and proposed a novel general problem formulation described by a latent representation, extractor and reconstructor functions. A linear modulation was addressed on a latent data representation with constraints on the modulation distortion and the resulting feature properties.

A computer simulation was provided using local image patches. Superior performance was demonstrated under the distortions AWGN, lossy JPEG compression and projective geometrical transform compared to the pCFP, aCFP and aCFPL schemes.

References

[1] S.Voloshynovskiy, F. Farhadzadeh, O. Koval, and T. Holotyak, "Active content fingerprinting: a marriage of digital watermarking and content fingerprinting," in *IEEE WIFS*, Tenerife, Spain, 2012.

[2] F. Farhadzadeh, F. Willems, and S.Voloshynovskiy, "Information-theoretical limits of active content fingerprinting in content-based identification systems," in *Proceedings of IEEE International Workshop on Information Forensics and Security*, Rome, Italy, November 16–19 2015.

[3] F. Farhadzadeh and S. Voloshynovskiy, "Active content fingerpriting," *Information Forensics and Security, IEEE Transactions on*, June 2014.

[4] F. Farhadzadeh, S. Voloshynovskiy, T. Holotyak, and F. Beekhof, "Active content fingerprinting: Shrinkage and lattice based modulations," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, pp. 3073–3077.

[5] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[6] M. Calonder, V. Lepetit, M. Ozuysal, T. Trzcinski, C. Strecha, and P. Fua, "BRIEF: Computing a Local Binary Descriptor Very Fast," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1281–1298, 2012.

[7] S Leutenegger, M Chli, and RY Siegwart, "Brisk: Binary robust invariant scalable keypoints," 2011, pp. 2548–2555.

[8] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski, "Orb: an efficient alternative to sift or surf," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2564–2571.

[9] Matti Pietikainen, Abdenour Hadid, Guoying Zhao, and Timo Ahonen, *Computer Vision Using Local Binary Patterns*, Computational imaging and vision. Springer Verlag, London, 2011.

[10] Dimche Kostadinov, Sviatoslav Voloshynovskiy, Maurits Diephuis, and Taras Holotyak, "Local active content fingerprinting: optimal solution under linear modulation," in *ICIP*, Phoenix, USA, September, 25-28 2016.

[11] Dimche Kostadinov, Sviatoslav Voloshynovskiy, Maurits Diephuis, Sohrab Ferdowsi, and Taras Holotyak, "On local active content fingerprinting: solutions for general linear feature maps," in *ICPR*, Cancun, Mexico, December, 2-4 2016.

[12] Dimche Kostadinov, Sviatoslav Volshinovsky, and Sohrab Ferdowsi, "Joint learning of local fingerprint and content modulation," in *EU-SIPCO, Kos, Grece*, October, 1-5 2017.

[13] Gerald Schaefer and Michal Stich, "Ucid: an uncompressed color image database," in *Electronic Imaging 2004*. International Society for Optics and Photonics, 2003, pp. 472–480.

[14] Michael Grant and Stephen Boyd, "Cvx: Matlab software for disciplined convex programming, version 2.1," March 2014.

		p_e			
		pCF	aCFP	aCFPL	aCFP-LR
AWGN	0dB	.149	.064	.063	.009
	5dB	.12.1	.022	.022	.008
	10dB	.092	.005	.004	.007
	20dB	.028	0	0	0

	pCF	aCFP	aCFPL	aCFP-LR
mL	0	60	60	100
DWR	∞	4.7	6.9	291

TABLE I
THE DWR AND THE p_e OF UNDER DIFFERENT ADDITIVE WHITE GAUSSIAN NOISE (AWGN) LEVEL.

		p_e			
		pCF	aCFP	aCFPL	aCFP-LR
QF	0	.082	.025	.024	.012
	5	.082	.015	.012	.010
	10	.028	.012	.011	.007

	pCF	aCFP	aCFPL	aCFP-LR
mL	0	30	30	100
DWR	∞	4.7	6.9	291

TABLE II
THE DWR AND THE p_e UNDER DIFFERENT JPEQ QUALITY FACTOR (QF).

	p_e			
	pCF	aCFP	aCFPL	aCFP-LR
Projec. QF=5	.058	.048	.047	.04

	pCF	aCFP	aCFPL	aCFP-LR
mL	0	60	60	100
DWR	∞	4.7	6.9	291

TABLE III
THE DWR AND THE p_e UNDER JPEQ QUALITY FACTOR QF=5 AND PROJECTIVE TRANSFORMATION.

		p_e
AWGN	-10dB	.015
	-20dB	.019
	-30dB	.020

		p_e
AWGN	-40dB	.071
	-50dB	.301
	-60dB	.321

		DWR
mL	100	291

		DWR
mL	100	291

		p_e
Proj. QF	3	.042
	2	.054
	1	.061

		p_e
Proj. AWGN	-10dB	.061
	-20dB	.074
	-30dB	.123

		DWR
mL	100	291

		DWR
mL	100	291

TABLE IV
THE DWR AND THE p_e FOR aCFP-LR UNDER PROJECTIVE TRANSFORM AND EXTREMELY LOW QF AND HIGH AWGN LEVELS.

ISBN 978-90-827970-1-5 © EURASIP 2018