2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

Local active content fingerprint: solutions for general linear feature maps

Dimche Kostadinov, Slava Voloshynovskiy, Maurits Diephuis, Sohrab Ferdowsi and Taras Holotyak
Computer Science Department, University of Geneva, 7 Route de Drize, Carouge 1227 GE, Switzerland
e-mail: {Dimce.Kostadinov, svolos, Maurits.Diephuis, Sohrab.Ferdowsi and Taras.Holotyak}@unige.ch

***Abstract*—This paper presents solutions to the local patch based Active Content Fingerprint (aCFP) with linear modulation, general linear feature map and convex constraints on the properties of the local feature descriptor. A direct approximation of the linear feature map such that the image distortion is as small as possible and the approximate linear feature map is as close as possible to the original map is proposed. Then an explicit regularization of the trade-off between the modulation distortion and the robustness of the local feature is introduced trough a novel problem formulation.**

A computer simulation using local image patches, extracted from publicly available data set is provided, demonstrating the advantages under: additive white Gaussian noise (AWGN), lossy JPEG compression and projective geometrical transform distortions.

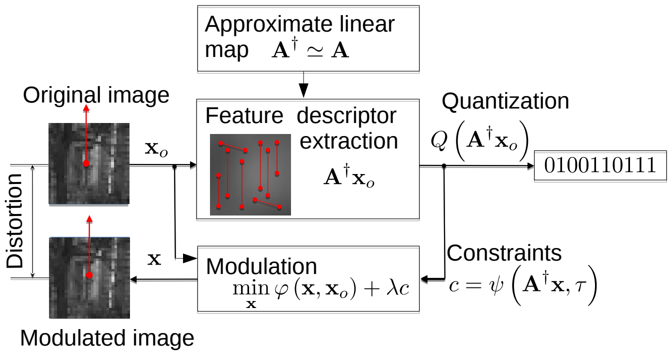


Fig. 1. Local aCFP framework

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. The advantages are related to a number of applications, including: content authentication, identification and recognition.

Recently, it was 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 have been proposed [3], [4] and have been 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] become a popular tool in image processing, computer vision and machine learning. These local descriptors are also considered as a form of local pCFP.

However, up to our best knowledge, there is small amount prior work on the modulation of local descriptors in the scope of aCFP or DWM.

In [10] an aCFP with a linear modulation subject to convex constraint on the properties of the resulting local descriptors was proposed, together with an optimal solution when the feature map is invertible.

The main open issues with the proposed optimal solution in [10] are related with the assumptions about the linear feature map and the general case with no constraints on the properties of the linear feature map is not addressed.

This paper approaches the general case from two distinct sides: firstly, by direct approximation of the linear feature map and secondly, by proposing novel problem formulation for the linear modulation and the constraints on the properties of the resulting local descriptor. The following contributions are presented:

- Approximation to the linear feature map, covering two cases: i) linear feature maps where the number of rows is bigger then the number of columns; ii) linear feature maps where the number of rows is smaller then the number of columns.
- Introduction of explicit regularization of the trade-off between the distortion and the robustness of the local feature by considering constraints on the pairwise encoding dependence's, on the distribution of the image modifications and on the distribution of the feature modifications with or without constraints on the range

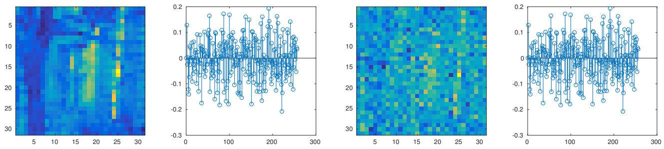


Fig. 2. **Column 1**: original image, **column 2**: features extracted from the original image, **column 3**: original image corrupted with AWGN noise, **column 4**: features extracted from the noisy original image;

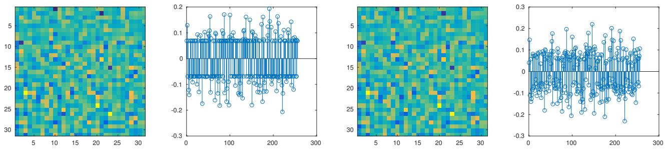


Fig. 3. **Column 1**: modulated image using the feature map **A**, **column 2**: features extracted from the modulated image, **column 3**: modulated image corrupted with AWGN noise, **column 4**: features extracted from noisy modulated image;

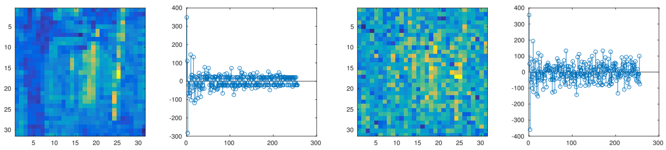


Fig. 4. **Column 1**: modulated image using the feature map $\mathbf{A}^{\dagger}$, **column 2**: features extracted from the modulated image, **column 3**: modulated image corrupted with AWGN noise, **column 4**: features extracted from noisy modulated image;

of values (compactness) of the solution.

- Validation by a computer simulation is presented using publicly available data set of images under: informed and non informed linear feature maps, original and approximate linear feature maps and several image processing distortions: AWGN, lossy JPEG compression and projective geometrical transform.

The paper is organized as follows. In Section 2 the problem is introduced, a short description of the local pCFP is given and the local aCFP modulation is presented. In Section 3 the main result is stated. Section 4 is devoted to computer simulation and Section 5 concludes the paper.

II. Local (patch based) linear modulation

The proposed aCFP framework consists of content modulation, prior to it's reproduction and descriptor extraction that includes feature mapping and quantization.

The scheme of the local aCFP framework is shown in Figure I. Fist an image block is presented to the system, then the local features are extracted. Based on the properties of the extracted features an constrain on the modulation is added and the modulated image block is estimated. The linear map approximation is independent of the actual modulation.

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 to the bit error in the descriptor bits: the 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 some distortion constraint. Obviously, the modulation faces a trade-off between two conflicting requirements of feature coefficient magnitude increase for the probability of bit error reduction and the modulation distortion. Fortunately, the low magnitude coefficient are concentrated near zero and are easily affected by a low distortion modulation.

Note that the local aCFP scheme is applicable also in the context of global image description.

A. Local descriptor extraction (pCFP): no patch modulation

Given an original image, around a local key point, local image patch $\mathbf{x}_o \in \Re^{N\times 1}$ is extracted. Usually the patch extraction is performed according to the patch orientation defined for example by a patch gradient (shown in red in Figure I).

Given a patch $\mathbf{x}_o$ in the most general case, the local features are extracted using a mapping function $f_2 : \Re^{N\times 1} \rightarrow \Re^{L\times 1}$, where L is the length of the descriptor. Consider a linear function $f_2(\mathbf{x}_o) = \mathbf{A}\mathbf{x}_o$, then $\mathbf{A} \in \Re^{L\times N}$ is a map (note that the map is either predefined, data independent and analytic or learned, data dependent and adaptive). The mapping, followed by a quantization $Q(.)$ results in the local descriptor $\mathbf{b}_x = Q(\mathbf{A}\mathbf{x}_o)$. The differences between the existing classes of local descriptors are determined by the defined mapping $f_2(.)$ and the type of the quantization $Q(.)$.

B. Local aCFP: patch modulation

The analysis here is focused on the solutions under linear maps with general properties and scalar quantizers used in such descriptors as ORB and LBP.

Linear modulation. We consider aCFP modulation to local image patches. The aCFP modulation function $f_1 : \Re^{N\times 1} \rightarrow \Re^{N\times 1}$ that is considered is linear:

$$f_1(\mathbf{x_o}) = \mathbf{Z}\mathbf{x}_o, \mathbf{Z} \in \Re^{N\times N} \tag{1}$$

Linear feature extraction. The considered feature extraction is linear: $f_2(\mathbf{x_o}) = \mathbf{A}\mathbf{x}_o$ defined as $\mathbf{A} = \mathbf{CT}$. The matrix $\mathbf{T} \in \Re^{M\times N}$ represents a linear transform, examples include low pass filter, DCT, FFT, WDT, random projections and others that typically are used by most of the known local descriptors for decorrelation and "robustification" of the features. The matrix $\mathbf{C} \in \{-1, 0, +1\}^{L\times M}$ represents the m-wise (pairwise, triplewise, etc.) constraints that describe the geometrical configuration of the considered pixel interactions.

Binary quantization. Let $\mathbf{t}_o = \mathbf{A}\mathbf{x}_o$, then the quantization is defined as:

$$\forall i, Q(t_o(i)) = \begin{cases} 1 & \text{if } t_o(i) \geq 0 \\ 0 & \text{if } t_o(i) < 0. \end{cases} \tag{2}$$

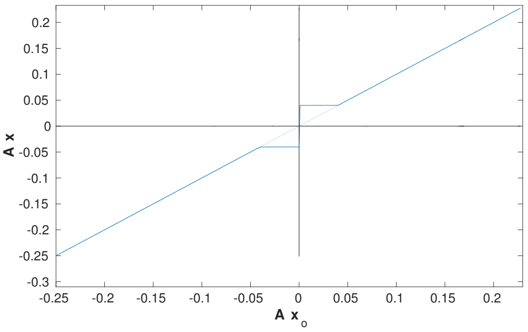


Fig. 5. Original image $\mathbf{x}_o$, modulated image $\mathbf{x}$ using (5), robustified feature $\mathbf{Ax}$

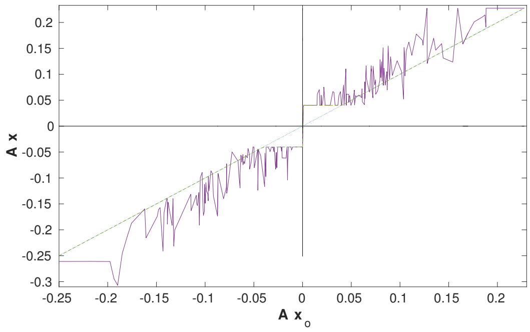


Fig. 6. Original image $\mathbf{x}_o$, modulated image $\mathbf{x}$ using (13), robustified feature $\mathbf{Ax}$

C. Generalization and reduction to a constrained projection problem

The generalized aCFP is a solution to a problem of functions estimation [10]:

$$\min_{f_1,f_2} \varphi\left(f_1\left(\mathbf{x}_o\right),\mathbf{x}_o\right) + \lambda_1 \psi\left(f_2\left(f_1\left(\mathbf{x}_o\right)\right),\tau\right), \tag{3}$$

where the first mapping function f_1 is the aCFP modulation that modifies the local data in the original domain and $\varphi(.)$ is a function that penalizes the modulation distortions in the original data domain. The second mapping function f_2 transforms the modified local data $f_1\left(\mathbf{x}_o\right)$ and $\psi(.)$ is a function that penalizes unwanted properties in the feature domain. The variable τ is the given modulation threshold, λ_1 is Lagrangian dual variable.

The aCFP with linear modulation, linear feature map and convex constraints on the properties of the features is a constrained projection problem [10]:

$$\begin{aligned} \hat{\mathbf{x}} = arg\min_{\mathbf{x}} \frac{1}{2}\left(\mathbf{x}_o - \mathbf{x}\right)^T\left(\mathbf{x}_o - \mathbf{x}\right) \\ \textbf{subject to} \\ |\mathbf{CTx}| \geq_e \tau\mathbf{1}, \end{aligned} \tag{4}$$

where $\geq_e$ represents element-wise inequality.

It was shown in [10] that if $\mathbf{A}$ is invertible the global optimal solution of (4) is:

$$\mathbf{x} = \mathbf{x}_o + \mathbf{A}^{-1}(sign\left(\mathbf{Ax}_o\right) \odot \max\{\tau\mathbf{1} - |\mathbf{Ax}_o|, \mathbf{0}\}), \tag{5}$$

where $\odot$ represents Hadamard (element-wise) product (the proof is given in [10]). Further, if $\mathbf{A}$ is a square orthogonal matrix $\mathbf{AA}^T = \mathbf{I}$, then $\mathbf{A}^{-1}$ in (5) is replaced with $\mathbf{A}^T$.

III. Trades between modulation distortion and feature robustness

A. Giving up distortion

It is possible to use (4) in the general case: with no constraints on the properties of the linear map $\mathbf{A}$. It that case the robustification of the feature descriptor will be achieved, however the level of modulation distortion is not regularized and it is possible for the nose level to become unacceptably high.

B. Giving up exact feature descriptor properties

On the other hand, instead of using the exact linear map $\mathbf{A}$ it is possible to use an approximate one such that the image distortions are as small as possible and the approximate map is as close as possible to the true map. There are two main cases.

Proposition 1: *The closest orthogonal matrix* $\mathbf{Z}$ *given the matrix* $\mathbf{A} \in \Re^{L\times N}, L \leq N$ *in a Gaussian sense is the solution to the problem*:

$$\begin{aligned} \mathbf{Z} = arg\min_{\mathbf{Z}} \|\mathbf{A} - \mathbf{Z}\|_F \\ \textbf{subject to} \\ \mathbf{ZZ}^T = \mathbf{I}, \end{aligned} \tag{6}$$

the optimal solution to (6) is $\mathbf{Z} = \mathbf{UI}_{L\times N}\mathbf{V}^T$ where $\mathbf{U\Sigma V}^T$ is a singular value decomposition (SVD) of $\mathbf{A}$. Then (5) is:

$$\begin{aligned} \mathbf{x} = \mathbf{x}_o + \mathbf{VI}_{N\times L}\mathbf{U}^T(sign\left(\mathbf{UI}_{L\times N}\mathbf{V}^T\mathbf{x}_o\right) \odot \\ \max\{\tau\mathbf{1} - |\mathbf{UI}_{L\times N}\mathbf{V}^T\mathbf{x}_o|, \mathbf{0}\}). \end{aligned} \tag{7}$$

Proposition 2: *The closest incoherent matrix* $\mathbf{Z}$ *given the matrix* $\mathbf{A} \in \Re^{L\times N}, L > N$, *in Gaussian sense is the solution to the problem* :

$$\begin{aligned} \mathbf{Z} = arg\min_{\mathbf{Z}} \|\mathbf{A} - \mathbf{Z}\|_F \\ \textbf{subject to} \\ \mu\left(\mathbf{Z}\right) \leq \epsilon_\mu, \end{aligned} \tag{8}$$

where $\mu\left(\mathbf{Z}\right) = \max_{\substack{i\neq j \\ i,j\in\{1,2,3,\ldots,L\}}} \frac{|\mathbf{z}_i\mathbf{z}_j^T|}{\|\mathbf{z}_i\|_2^2\|\mathbf{z}_j\|_2^2}$ and $\mathbf{z}_i$ is the ith column of $\mathbf{Z}$.

Given any incoherent $\mathbf{B} \in \{\mathbf{B} \in \Re^{L\times N}, \mu\left(\mathbf{B}\right) \leq \epsilon_\mu\}$, the solution of (8) is equivalent to a product of a rotation $\mathbf{R}$ matrix and the incoherent $\mathbf{B}$ matrix. This decomposition is not unique. Nevertheless, the rotation $\mathbf{R}$ matrix is a solution to the following problem [11]:

$$\begin{aligned} \mathbf{R} = arg\min_{\mathbf{R}} &\|\mathbf{RB}-\mathbf{A}\|_F \\ &\textbf{subject to} \\ &\mathbf{RR}^T = \mathbf{I}, \end{aligned} \tag{9}$$

the optimal solution is $\mathbf{R}=\mathbf{UV}^T$ where $\mathbf{U}\Sigma\mathbf{V}^T$ is a SVD of $\mathbf{AB}^T$. Therefore $\mathbf{Z}=\mathbf{UV}^T\mathbf{B}$ and the final solution is:

$$\begin{aligned} \mathbf{x} = \mathbf{x}_o + &\mathbf{B}^{-1}\mathbf{VU}^T(sign\left(\mathbf{UV}^T\mathbf{Bx}_o\right) \odot \\ &\max\{\tau\mathbf{1} - |\mathbf{UV}^T\mathbf{Bx}_o|, \mathbf{0}\}). \end{aligned} \tag{10}$$

C. Explicit regularization of the trade-off between modulation distortion and feature robustness

An explicit regularization to (4) is introduced in order to improve it's solution, where several constraints are considered:

1) the dependencies for a subsets of pairwise constraints represented by the matrix $\mathbf{C}$
2) the distribution of the modulated modifications
3) the distribution of the feature descriptor modifications
4) constraints on the range of values (compactness) of the optimal solution.

Define $\mathbf{t}_o = \mathbf{CTx}_o = \mathbf{Ax}_o$, 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 $\mathbf{A}_s$ be rows reordered $\mathbf{A}$ such that $\mathbf{A}_s\mathbf{x}_o = \mathbf{s}_o$.

Definition 1: The strength of dependence for a set of connected points:

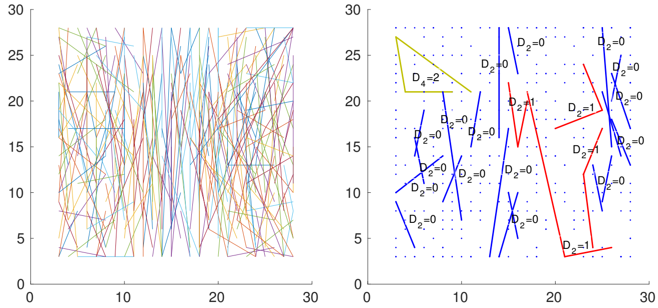
$$\begin{aligned} \mathcal{S}_{\{i_1,i_2,...,i_Q\}} = &\{\{x(i_1), x(i_2), ..., x(i_Q)\}, \\ &\{x(i_j), x(i_{j+1})\}, i_j \neq i_{j+1}, \\ &j, j+1 \in \{1,2,3,...,Q-1\}, \\ &i_j \in \{1,2,3,...,N\}\}, \end{aligned} \tag{11}$$

is defined as $D_{|\mathcal{S}_{\{i_1,i_2,...,i_Q\}}|} = |\mathcal{S}_{\{i_1,i_2,...,i_Q\}}| - 2$.

Lemma 1: $\forall \mathbf{C} \in \{-1,1\}^{L\times N}$ *if there exists at least one subset* $\mathcal{S}_{i_1,i_2,...,i_Q}$ *from* $\mathbf{C}$ *such that* $D_{|\mathcal{S}_{\{i_1,i_2,...,i_Q\}}|} \geq 1$ *then there exists at least one pair of elements* $t_o(i), t_o(j), i \neq j, i, j \in \{1,2,3,...L\}$ *from* $\mathbf{t}_o \in \Re^{L\times 1}$ *that are linearly dependent.*

Proposition 3: *An aCFP with explicitly regularized trade-off between modulation distortion and feature robustness, linear modulation, linear feature map and convex constraints on the properties of the features is a solution to the constrained projection problem*:

$$\begin{aligned} \hat{\mathbf{x}} = arg \min_{\mathbf{x},\mathbf{s}_l,\mathbf{s}_h} &(d_0\left(\mathbf{x}_o, \mathbf{x}\right) + \lambda_1 d_1\left(\mathbf{s}_l\right) + \\ &\lambda_2 d_2\left(\mathbf{s}_h\right) + \lambda_3\left(l+u\right)) \\ &\textbf{subject to} \\ &\mathbf{A}_s\mathbf{x} =_e \left[\begin{smallmatrix}\mathbf{s}_l\\ \mathbf{s}_g\end{smallmatrix}\right] \\ &d_3\left(\left[\begin{smallmatrix}\mathbf{s}_l\\ \mathbf{s}_g\end{smallmatrix}\right]\right) \geq_e \tau \begin{bmatrix}\mathbf{1}_l\\ \mathbf{1}_g\end{bmatrix} \\ &\mathbf{x} \geq_e l\mathbf{1} \\ &l \geq 0 \\ &\mathbf{x} \leq_e u\mathbf{1}, \end{aligned} \tag{12}$$



a) b)

Fig. 7. a) All of the pairwise constraints encoded by the matrix $\mathbf{C}$. b) Cases of subset $\mathcal{S}_{i_1,i_2,...,i_Q}$ from $\mathbf{C}$ with different strength of linear dependence

where λ_1, λ_2 and λ_3 are Lagrangian multipliers, $d_0\left(.,.\right)$, $d_1\left(.,.\right)$, $d_2\left(.,.\right)$ and $d_3\left(.,.\right)$ are penalty function, l and u are bounds on the range of values.

In the decomposition $\left[\begin{smallmatrix}\mathbf{s}_l\\ \mathbf{s}_g\end{smallmatrix}\right]$, $\mathbf{s}_l$ are the components of the modified feature descriptor related to the components of the original descriptor $\mathbf{A}_s\mathbf{x}$ that are less than the modulation threshold τ and $\mathbf{s}_g$ are the components of the modified feature descriptor related to the components of the original descriptor $\mathbf{A}_s\mathbf{x}$ that are greater than the modulation threshold τ.

Without knowing the priors on $d_0(.,.)$, $d_1(.)$ and $d_2(.)$ the principle idea is to consider the worst case where the residual elements $\left(x(i) - x_o(i)\right), \forall i \in \{1,2,3,..,N\}$ are Gaussian distributed.

Define:

- $d_0(\mathbf{x},\mathbf{x}_0) = \frac{1}{2}\|\mathbf{x}-\mathbf{x}_0\|_2^2$ and $d_1(.) = d_2(.) = 0$
- a weak constraints on the feature descriptor $\mathbf{Ax}$ such that it's elements are only greater/less than the required modulation level τ
- a weak constraints on u.

then (12) is:

$$\begin{aligned} \hat{\mathbf{x}} = arg \min_{\mathbf{x},u} &\frac{1}{2}\left(\mathbf{x}_o - \mathbf{x}\right)^T\left(\mathbf{x}_o - \mathbf{x}\right) + \lambda_3 u \\ &\textbf{subject to} \\ &\left(\mathbf{A}_s\mathbf{x}\right)^- \leq_e -\tau\mathbf{1} \\ &\left(\mathbf{A}_s\mathbf{x}\right)^+ \geq_e \tau\mathbf{1} \\ &\mathbf{x} \geq_e 0 \\ &\mathbf{x} \leq_e u\mathbf{1}, \end{aligned} \tag{13}$$

where $\left(\mathbf{Ax}\right)^-$ and $\left(\mathbf{Ax}\right)^+$ are the negative and positive components respectivly.

Note that if the encoding matrix $\mathbf{C}$ does not contain dependent pairs than (13) will not bring improvement and the solution of (5) is equivalent to the solution of (13). However, having at least two dependent pairs than a trade between the modulation error and constrained feature descriptor modification (greater/less than the required modulation level τ) is possible.

IV. COMPUTER SIMULATIONS

A computer simulation is performed to demonstrate the advantages of the local aCFP scheme over pCFP under several signal processing distortions, including AWGN, lossy JPEG compression and projective geometrical transform.

The UCID [12] 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, with $\sqrt{N} = 31$ were extracted around each detected feature point. The features were sorted by scale-space, 30 patches were extracted from individual image.

Let the matrix $\mathbf{X}_o = [\mathbf{x}_{o,1}, \mathbf{x}_{o,2}, \mathbf{x}_{o,3}, ..., \mathbf{x}_{o,1000}]$ represent all the original available local image patches and define the matrix $\mathbf{T} \in \Re^{N \times N}$ to represents low pass filter with 11×11 window.

Two cases are simulated an informed and non-informed one. Consider that an local image block is transmitted trough a noisy channel.

Informed case: At the receiver end, in the informed case an amount of available information is presented about the original image. Only one matrix $\mathbf{A}_I$ is used in the aCFP scenario. $\mathbf{A}_I = \left(\mathbf{U}\mathbf{I}_{L \times M}\mathbf{V}^T\right)^T$ where $\mathbf{U}, \mathbf{V}$ are obtained by SVD of $(\mathbf{C}\mathbf{T}_I)^T$.

The matrix $\mathbf{T}_I = \mathbf{R}\left[\mathbf{x}_o\mathbf{x}_o^T\right]^{-1} \in \Re^{N \times N}$, where $\mathbf{R} \in \Re^{N \times N}$ is a random matrix is generated as follows. First the matrix $\mathbf{x}_o\mathbf{x}_o^T$ is quantized in 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 $\mathbf{x}_o\mathbf{x}_o^T$, 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 uniform distribution with the support $[0, 1]$. The main idea is to try to make the contribution of the elements of $\mathbf{x}_o\mathbf{x}_o^T$ in the linear feature map $\mathbf{C}\mathbf{T}_I$ equi likely.

The matrix $\mathbf{A}_I^T$ is the closest orthogonal to $(\mathbf{C}\mathbf{T})^T$, satisfying $\mathbf{A}_I\mathbf{A}_I^T = \mathbf{I}$.

Non-informed case: In the non-informed case, an information about the original image is not presented. Three different matrices $\mathbf{A}$, $\mathbf{A}^\dagger \simeq \mathbf{A}$ with $\mathbf{A}^\dagger(\mathbf{A}^\dagger)^T = \mathbf{I}$ and $\mathbf{A}^r$, $\forall i, j, A^r(i, j) \sim \mathcal{N}(0, 1)$ with $\mathbf{A}^r(\mathbf{A}^r)^T = \mathbf{I}$ are used in the pCFP and the aCFP scenario.

Measures: Three measured quantities are used in the evaluation:

- 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 is defined as $\tau = \max_{1 \leq i \leq K} |s_o(i)|$.
- The *modulation distortion* DWR is defined as $DWR = 10\log_{10}\left(\frac{255^2}{\Delta^2}\right) [dB]$, $\Delta = \frac{1}{N}\|\mathbf{x} - \mathbf{x}_o\|_2$.
- The *probability of bit error* $p_e$is defined by the probability of correct bit $p_e = 1 - p_c$, $p_c = \frac{1}{L}\sum_{i=1}^{L}\mathbf{1}\{b_x(i), b_y(i)\}$ with $L = 256$ bits, where $\mathbf{b}_x = Q(\mathbf{A}\mathbf{x}_o)$, $\mathbf{b}_y = Q(\mathbf{A}\mathbf{x}_o + \mathbf{A}\mathbf{x}_e + \mathbf{A}\mathbf{x}_n)$, $\mathbf{x}_n$ is the introduced distortion and $\mathbf{1}\{\}$ is an indicator function such that $\mathbf{1}\{a, b\} = 1$, if $a = b$ and $\mathbf{1}\{a, b\} = 0$, otherwise.

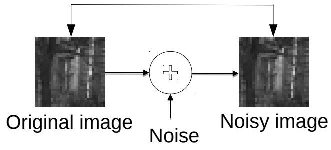


Fig. 8. Informed case

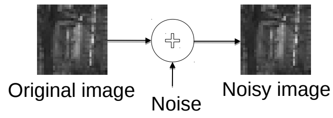


Fig. 9. Non informed case

A. *AWGN*

The results from a single patch were obtained as an 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. Two modulation levels (mL) were used 10 and 60.

B. *Lossy JPEG compression*

Three strong levels of JPEG quality factors (QF) 0, 5 and 10 were used. The modulation levels (mL) that were used are 10 and 30.

C. *Projective transform with lossy JPEG compression*

A projective transformation $\mathbf{P} \in \Re^{3 \times 3}$ was used, where:

$$\mathbf{P} = \begin{bmatrix} 1.0763 & 0.0325 & 0 \\ 1.0763 & 0.0325 & 0 \\ -24.32 & -70.37 & 1 \end{bmatrix}, \tag{14}$$

followed by a lossy JPEG compression with QF=5. The modulation levels (mL) that were used are 10 and 60.

In Tables I, II and III are provided the average results for a total of 1000 image patches.

The results show that the pair of highest DWR and lowest p_e is achieved in the informed case for the aCFP scenario under all types of noise.

The non-informed aCFP consistently outperforms the non-informed pCFP in terms of p_e for all types of noise, however at cost of introducing a modulation distortion.

It is important to highlight that the greatest reduction in p_e is .15, achieved at AWGN noisy level of 0dB and modulation level $mL = 60$ using the proposal (13). The results by (13) produce even smaller p_e than the results in the informed case when using $\mathbf{A}_I$, however the modulation noise DWR is smaller for the former case.

On the other hand the proposal consisting of approximate liner map provides small improvement in terms of p_e, however it achieves high DWR, that is small modulation distortion.

V. CONCLUSION

This paper presented solutions to the local patch based aCFP with linear modulation, general linear feature map and convex constraints on the properties of the local feature descriptor. A linear feature map approximation was proposed. An explicit regularization of the trade-off between the modulation distortion and the robustness of the local feature was introduced trough a novel problem formulation.

Non-informed and informed case: pCFP results

		p_e			
		$\mathbf{A}$	$\mathbf{A}^\dagger$	$\mathbf{A}^r$	$\mathbf{A}_I$
AWGN	0dB	.224	.422	.3485	.15
	5dB	.150	.373	.2661	.12
	10dB	.095	.310	.1795	.09
	20dB	.034	.160	.0640	.03
QF	0	.082	.244	.072	.03
	5	.051	.190	.056	.02
	10	.028	.144	.044	.01
Proj., QF=5		.058	.233	.0769	.05

TABLE I
THE DWR AND THE p_e UNDER PCFP USING VARYING AWGN NOISE, JPEQ QUALITY FACTOR AND PROJECTIVE TRANSFORMATION WITH QF=5 FOR THE FEATURE MAPS $\mathbf{A}$, $\mathbf{A}^\dagger$, $\mathbf{A}^r$ AND $\mathbf{A}_I$

Informed case: aCFP results using the proposal (5)

		p_e	
		$\mathbf{A}_I$	
mL		10	60
DWR		51	28
AWGN	0dB	.15	.11
	5dB	.11	.05
	10dB	.08	.01
	20dB	.02	0
QF	0	.02	0
	5	.01	0
	10	0	0
Projective, QF=05		.05	.03

Non-informed case: aCFP results using the proposal (5)

		p_e	
		$\mathbf{A}$	
mL		10	60
DWR		20.0	-6.9
AWGN	0dB	.220	.121
	5dB	.145	.045
	10dB	.086	.010
	20dB	.019	0
QF	0	.082	.253
	5	.049	.217
	10	.022	.204
Proj., QF=5		.053	.263

TABLE II
THE DWR AND THE p_e USING VARYING ACFP MODULATION UNDER VARYING AWGN NOISE, JPEQ QUALITY FACTOR AND PROJECTIVE TRANSFORMATION WITH QF=5 FOR THE FEATURE MAPS $\mathbf{A}$ AND $\mathbf{A}_I$

Non-informed case: aCFP results using row orthogonal linear feature maps and the proposal (5)

		p_e			
		$\mathbf{A}^\dagger$		$\mathbf{A}^r$	
mL		10	60	10	60
DWR		52.4	27.9	43.9	19.6
AWGN	0dB	.421	.400	.348	.297
	5dB	.374	.337	.264	.184
	10dB	.310	.248	.176	.072
	20dB	.157	.047	.053	.001
QF	0	.243	.218	.070	.050
	5	.190	.154	.054	.027
	10	.143	.088	.041	.009
Proj., QF=5		.232	.210	.075	.049

Non-informed case: aCFP results using the proposal (13)

		p_e					
		$\mathbf{A}$		$\mathbf{A}^\dagger$		$\mathbf{A}^r$	
mL		10	60	10	60	10	60
DWR		33.8	4.7	52.4	27.9	43.9	19.6
AWGN	0dB	.217	.064	.421	.401	.348	.297
	5dB	.142	.022	.374	.338	.264	.184
	10dB	.084	.005	.310	.249	.177	.073
	20dB	.018	0	.157	.047	.053	.001
QF	0	.074	.025	.242	.219	.070	.050
	5	.040	.015	.190	.153	.054	.026
	10	.015	.012	.143	.088	.041	.009
Proj., QF=5		.049	.048	.232	.209	.075	.046

TABLE III
THE DWR AND THE p_e USING VARYING ACFP MODULATION UNDER VARYING AWGN NOISE, JPEQ QUALITY FACTOR AND PROJECTIVE TRANSFORMATION WITH QF=5 FOR THE FEATURE MAPS $\mathbf{A}$, $\mathbf{A}^\dagger$ AND $\mathbf{A}^r$

The computer simulation using local image patches, extracted from publicly available data set was provided and the advantages under the distortions AWGN, lossy JPEG compression and projective geometrical transform were demonstrated.

The results produced by the proposed linear modulation show that small p_e is achievable under different and severe signal processing distortions, however at cost of introducing modulation distortion.

REFERENCES

[1] S.Voloshynovskiy, F. Farhadzadeh, O. Koval, and T. Holotyak, "Active content fingerprinting: a marriage of digital watermarking and content fingerprinting," in *Proceedings of IEEE International Workshop on Information Forensics and Security*, Tenerife, Spain, December 2–5 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*, vol. 9, no. 6, pp. 905–920, 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, Slava Voloshynovskiy, Maurits Diephuis, and Taras Holotyak, "Local active content fingerprint: optimal solution under linear modulation," *submitted to the International Conference of Image Processing*, 2016.

[11] Daniele Barchiesi and Mark Plumbley, "Learning incoherent dictionaries for sparse approximation using iterative projections and rotations," *IEEE Transactions on signal processing*, vol. 61, no. 8, pp. 2055–2065, 2013.

[12] 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.