Information-Theoretical Limits of Active Content Fingerprinting in Content-based Identification Systems

Farzad Farhadzadeh*, Frans M. J. Willems*
*Electrical Engineering Department
Eindhoven university of Technology
Eindhoven, The Netherlands
Email: {F.Farhadzadeh, F.M.J.Willems}@tue.nl

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

***Abstract*—Content fingerprinting and digital watermarking are techniques that are used for content protection and distribution monitoring and, more recently, for interaction with physical objects. Over the past few years, both techniques have been well studied and their shortcomings understood. In this paper, we introduce a new framework called active content fingerprinting, which takes the best from these two worlds, i.e., the world of content fingerprinting and that of digital watermarking, in order to overcome some of the fundamental restrictions of these techniques in terms of performance and complexity. The proposed framework extends the encoding process of conventional content fingerprinting such that it becomes possible to extract more robust fingerprints from the modified data. We consider different encoding strategies and examine the performance of the proposed schemes in terms of content identification rate in an information theoretical framework and compare them with those of conventional content fingerprinting and watermarking.**

I. Introduction

Content-based identification systems [1] are generally facing numerous requirements related to identification rate, complexity, privacy, security as well as memory storage. The trade-off between these requirements is a complex problem that remains still largely unsolved. To address this trade-off, *content fingerprints* are used [2], [3]. A content fingerprint is a short, robust and distinctive description of content.

In conventional content fingerprinting, the fingerprint is extracted directly from an original content and does not require any content modification that preserves the original content quality. In this sense, it can be considered as a *passive content fingerprinting* (PCFP). The extracted fingerprints resemble random codes, for which no efficient decoding algorithm is known in contrast to structured codes. Moreover, the performance of PCFP in terms of the identification rate is not satisfactory due to acquisition device imperfections.

Another approach for content identification is Digital Watermarking (DWM). DWM is already a well-studied domain these days, and its performance has been well investigated. Gelfand-Pinsker's channel with channel state information available at the transmitter [4] provided theoretical grounds for DWM. Later, Costa's famous result [5] was used in optimal watermarking schemes under Gaussian assumptions. DWM possesses two advantages over the PCFP: (a) each copy of content can be marked independently and (b) there is no need for complex search due to the usage of structured codes. However, as will be explained later on, PCFP can outperform the optimal watermarking scheme under Gaussian assumptions in terms of the identification rate.

Recently, *active content fingerprinting* (ACFP) was proposed in [6], [7], in which the basic idea was introduced and a feasibility study revealed higher performance with respect to PCFP and digital watermarking. ACFP by modifying digital contents takes the best from these two worlds, i.e. the world of content fingerprinting and that of digital watermarking, to overcome some of fundamental restrictions of these techniques in terms of performance and complexity. In the active content fingerprinting scheme, digital contents are modified in a way to enhance the identification rate and reduce the decoding complexity with respect to conventional content fingerprinting.

The main goal of this article is to analyze ACFP in an information theoretical framework and to investigate its fundamental limits in content-based identification systems. In this paper, we analyze ACFP from two different points of view. First, we briefly show the identification capacity based on ACFP in which a digital content can be arbitrarily modified under a certain distortion constraint. Then, we show the optimal encoding/modification scheme under Gaussian assumptions that can achieve the identification capacity based on ACFP. These primarily results were already investigated by the authors in [8].

Secondly, we extend our analysis by considering the content-based identification system based on ACFP using codes. Contrary to the previous scheme, in code-based ACFP modified contents can be chosen only from a predefined set of codes. While the procedure resembles vector quantization in Rate distortion theory the difference will be explained later on. The main goal of this analysis is to show the possibility of reducing the search complexity in content-based identification systems based on ACFP using structured codes (vector quantization). Finally, we consider the optimal encoding/quantization scheme under Gaussian assumptions.

The rest of this paper is organized as follows. Section II presents the identification system based on arbitrarily content modification, i.e., the *general model*. Section III presents the identification system based on vector quantization, the *coding-based model*. Concluding remarks follow in Section V.

Notations: Throughout this paper, we adopt the convention that a scalar random variable is denoted by a capital letter X, a specific value it may take is denoted by the lower case letter x, and its alphabet is designated by the script letter $\mathcal{X}$. As for vectors, a capital letter X^N with a corresponding superscript will denote an N-dimensional random vector $X^N = (X_1, \ldots, X_N)$. A lower case letter x^N will represent its particular realization $x^N = (x_1, \ldots, x_N)$.

II. General Model

A. Model Description

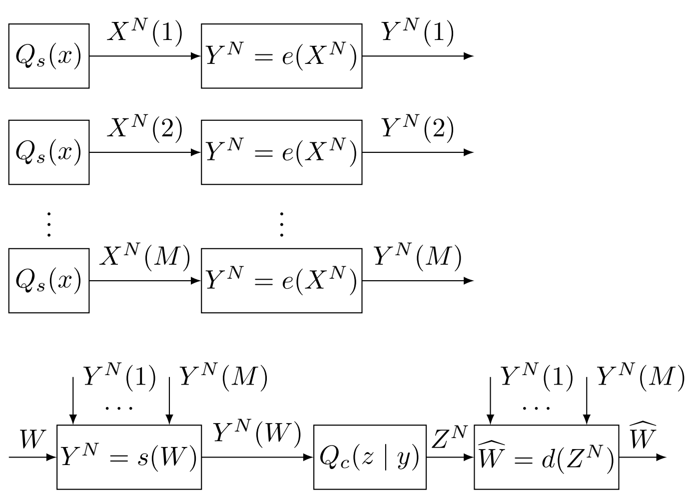


Fig. 1. Model of an identification system using modified content-sequences.

In an identification system, see Fig. 1, there are M items, which are indexed by$w \in \{1, 2, \cdots, M\}$, to be identified. A randomly generated content-sequence (vector) of length N corresponds to each item. This sequence has symbols $x_n, n = 1, 2, \cdots, N$ taking values in the discrete alphabet $\mathcal{X}$, and the probability that content-sequence $x^N = (x_1, x_2, \cdots, x_N)$ occurs for item w is

$$\Pr\{X^N(w) = x^N\} = \prod_{n=1}^{N} Q_s(x_n), \tag{1}$$

hence the components $X_1, X_2, \cdots, X_N$ are independent and identically distributed according to $\{Q_s(x), x \in \mathcal{X}\}$. Note that this probability does not depend on the index w.

An encoder $e(\cdot)$ transforms each content-sequence x^N into a modified content-sequence $y^N = (y_1, y_2, \cdots, y_N)$, where $y_n, n = 1, 2, \cdots, N$, is taking values in the discrete alphabet $\mathcal{Y}$. The distortion between the modified content-sequence and the original content-sequence cannot be too large. The modification distortion $\overline{D_{xy}}$ is defined as

$$\overline{D_{xy}} = \frac{1}{N} E\left[\sum_{n=1}^{N} D_{xy}(X_n, Y_n)\right], \tag{2}$$

where $\{D_{xy}(x, y), x \in \mathcal{X}, y \in \mathcal{Y}\}$ is the distortion matrix specifying the distortion per component. We assume that the distortion matrix only has finite non-negative entries. Moreover, we assume that all modified content-sequences are generated prior to the identification procedure. These modified content-sequences form the "database". This database $\mathcal{C}$ consists of the list of entries, hence

$$\mathcal{C} = \left(y^N(1), y^N(2), \cdots, y^N(M)\right). \tag{3}$$

In the identification process the probabilities for the items to be presented for identification are all equal, hence

$$\Pr\{W = w\} = 1/M \text{ for } w \in \{1, 2, \cdots, M\}. \tag{4}$$

When item w is presented for identification, its corresponding modified content-sequence $y^N(w)$ is "selected" from the database C and presented to the system, hence

$$y^N = s(w), \tag{5}$$

where $s(\cdot)$ is the selector. The system observes y^N via a discrete memoryless observation channel $\{Q_c(z|y), y \in \mathcal{Y}, z \in \mathcal{Z}\}$, with discrete alphabet $\mathcal{Z}$, and the resulting channel output sequence is $z^N = (z_1, z_2, \cdots, z_N)$, where $z_n \in \mathcal{Z}$ for $n = 1, 2, \cdots, N$. Now

$$\Pr\{Z^N = z^N | Y^N(w) = y^N\} = \Pi_{n=1}^{N} Q_c(z_n|y_n). \tag{6}$$

After observing Z^N, the decoder finds out to which modified content-sequence Z^N is related. If this is $Y^N(w)$, the decoder outputs $\widehat{W} = w$. The reliability of our identification system is measured by the error probability

$$P_{\mathcal{E}} = \Pr\{\widehat{W} \neq W\}. \tag{7}$$

B. Statement of Result

An identification rate-distortion pair (R, Δ) is said to be achievable if for all $\epsilon > 0$ there exist for all N large enough an encoder and a decoder such that

$$\begin{aligned} \overline{D_{xy}} &\leq \Delta + \epsilon, \\ \log_2 M &\geq N(R - \epsilon), \text{ and} \\ \Pr\{\widehat{W} \neq W\} &\leq \epsilon. \end{aligned} \tag{8}$$

Theorem 1. *The region of achievable rate-distortion pairs (R, Δ) for the identification system based on ACFP under an arbitrary modification scheme is given by*

$$\Big\{(R, \Delta) : R \leq I(Y; Z), \quad \Delta \geq \sum_{x,y} Q_s(x) P_t(y \mid x) D_{xy}(x, y), \quad \textit{for } P(x, y, z) = Q_s(x) P_t(y \mid x) Q_c(z \mid y)\Big\}. \tag{9}$$

Proof. It is given in [8]. □

The "capacity" of the content-based identification system based on ACFP, the maximum of all achievable identification rate, for a given distortion Δ is given by

$$C_{\text{ACFP}}(\Delta) = \max_{P_t(y|x):\sum_{x,y} Q_s(x)P_t(y|x)D_{xy}(x,y)\leq\Delta} I(Y;Z). \tag{10}$$

C. Gaussian Source

Let's assume the content-sequences are distributed i.i.d. according to a Gaussian distribution with variance V_X and mean zero. Moreover, the observation channel $Q_c(z \mid y)$ can be modelled as an Additive White Gaussian Noise (AWGN) channel where the noise has variance V_N.

Theorem 2. *Considering mean-squared error distortion, the capacity of the content-based identification system based on ACFP under arbitrary modification is given by*

$$C_{ACFP}(\Delta) = \frac{1}{2}\log_2\left(1 + \frac{(\sqrt{V_X}+\sqrt{\Delta})^2}{V_N}\right), \tag{11}$$

that can be achieved by scaling content by factor f, i.e. $Y^N = fX^N$, such that $(f-1)^2V_X = \Delta$.

Proof. It is given in [8]. □

III. Coding-based Model

A. Model Description

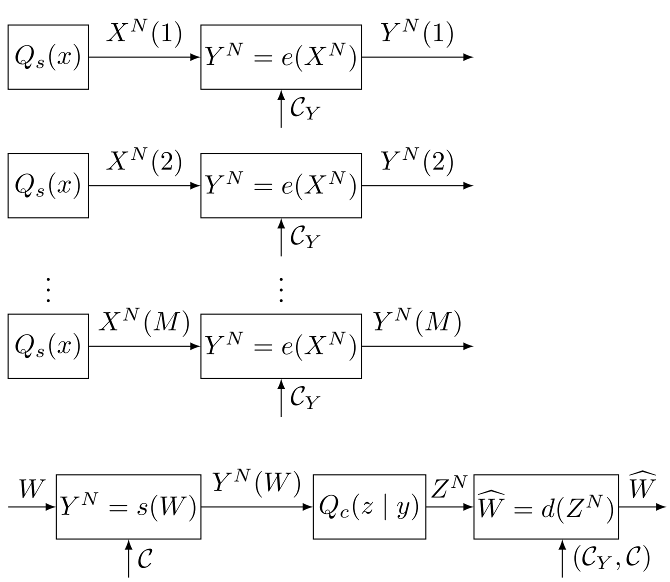


Fig. 2. Model of an identification system using predefined codewords.

Similar to the identification scheme defined in Section II-A, in coding-based identification, see Fig. 2, there are M items indexed $w \in \{1, 2, \cdots, M\}$ to be identified. Randomly generated content-sequences $x^N(w)$ corresponding to each item w is generated according to (1).

A predefined set of codes $\mathcal{C}_Y = \left(y^N(1), y^N(2), \cdots, y^N(M_C)\right), y^N \in \mathcal{Y}^N$ of cardinality M_C is revealed to the encoder at the enrollment phases. At the enrollment phase, the encoder $e(\cdot)$ maps each content-sequence x^N into a code sequence $y^N \in \mathcal{C}_Y$ and the assigned codeword is indexed by the content-sequence index. The indexed codewords form the database $\mathcal{C} = \left(y^N(1), y^N(2), \cdots, y^N(M)\right)$ and is shared with the decoder (Fig. 2). The encoding/modification distortion is defined similarly to (2).

The identification process follows similar to the model defined in Section II-A.

B. Statement of Result

Theorem 3. *The region of achievable rate-distortion pair (R, Δ), which is defined in (8), for the identification system based on ACFP using the coding scheme is given by*

$$\Big\{(R,\Delta): R \leq I(Y;Z),\quad I(X;Y) \leq I(Y;Z),\quad \Delta \geq \sum_{x,y} Q_s(x)P_t(y \mid x)D_{xy}(x,y),\quad \textit{for } P(x,y,z) = Q_s(x)P_t(y\mid x)Q_c(z\mid y)\Big\}. \tag{12}$$

C. Proof

The proof consists of the achievability part and a converse. We start with the converse.

1) Converse Part: First, we define the random variable I that takes values in $\{1, 2, \cdots, N\}$ with probability $1/N$. Then the random triple (X, Y, Z) is defined as $(X, Y, Z) \triangleq (X_i, Y_i, Z_i)$ if $I = i$. Hence, the joint distribution of (X, Y, Z) is given by

$$\begin{aligned} P(x,y,z) &= \frac{1}{N}\sum_{i=1}^{N} \Pr\{X_i = x, Y_i = y, Z_i = z\} \\ &= \frac{1}{N}\sum_{i=1}^{N} Q_s(x_i)P_t(y_i \mid x_i)Q_c(z_i \mid y_i) \\ &= Q_s(x)P_t(y \mid x)Q_c(z \mid y). \end{aligned} \tag{13}$$

We have

$$\begin{aligned} H(Y^N) &\geq H(Y^N) - H(Y^N|X^N) \\ &\overset{(a)}{\geq} \sum_{i=1}^{N} H(X_i) - H(X_i \mid Y_i) \\ &= NH(X \mid I) - NH(X \mid Y, I) \\ &\overset{(b)}{\geq} NI(X;Y), \end{aligned} \tag{14}$$

where (a) and (b) follow from the facts that $X_1, X_2, \ldots, X_N$ are distributed i.i.d., and conditioning reduces entropy.

Using Fano's inequality $H(W|Z^N) \leq H(W|\widehat{W}) \leq F$ where $F = 1 + \Pr\{\widehat{W} \neq W\}\log_2 M$, we have

$$F \geq H(W|Z^N)$$

$$
\begin{aligned}
&\overset{(a)}{=} H(W, Y^N|Z^N) \\
&\geq H(Y^N|Z^N), \qquad (15)
\end{aligned}
$$

where (a) follows from the fact that $H(Y^N|Z^N, W) = 0$ since $H(Y^N|Z^N, W) \leq H(Y^N|W) = 0$. Hence,

$$
\begin{aligned}
H(Y^N) &\leq H(Y^N) - H(Y^N|Z^N) + F \\
&\overset{(a)}{\leq} \sum_{i=1}^{N} H(Z_i) - H(Z_i|Y_i) + F \\
&= NH(Z|I) - NH(Z|Y, I) + F \\
&\overset{(b)}{\leq} NH(Z) - NH(Z|Y) + F \\
&= NI(Z;Y) + F \qquad (16)
\end{aligned}
$$

where (a) and (b) follows by definition of a discrete memoryless channel, Z_i depends only on Y_i and is conditionally independent of everything else.

Moreover,

$$
\begin{aligned}
log_2 M &= H(W) \\
&\leq H(W) - H(W|Z^N) + F \\
&\overset{(a)}{\leq} I(Y^N; Z^N) + F \\
&\overset{(b)}{\leq} NI(Y;Z) + F. \qquad (17)
\end{aligned}
$$

where (a) follows from the Markov chain $W \leftrightarrow Y^N \leftrightarrow Z^N$, and (b) follows similar to (16).

Now for the distortion part we have

$$
\begin{aligned}
\overline{D_{xy}} &= \frac{1}{N} E\left[\sum_{n=1}^{N} D_{x,y}(X_n, Y_n)\right] \\
&= \frac{1}{N} \sum_{n=1}^{N} \sum_{x_n, y_n} Q_s(x_n) P_t(y_n \mid x_n) D_{x,y}(x_n, y_n) \\
&= \sum_{x,y} Q_s(x) P_t(y \mid x) D_{x,y}(x, y) \\
&= D_{xy}(X, Y). \qquad (18)
\end{aligned}
$$

Now, assume that (R, Δ) is achievable. Then $F \leq 1 + \epsilon \log_2(M)$ and $\Delta \geq \overline{D_{xy}} - \epsilon$. For all blocklengths N and small enough $\epsilon > 0$, we obtain from (14), (16) and (17) that

$$
NI(X;Y) \leq NI(Y;Z) + 1 + \epsilon \log_2 M \qquad (19a)
$$

$$
NR \leq \frac{1}{1-\epsilon}(NI(Y;Z) + 1) \qquad (19b)
$$

for some $P(x, y, z) = Q_s(x) P_t(y \mid x) Q_c(z \mid y)$. If we now let $\epsilon \downarrow 0$ and $N \to \infty$, then we obtain the converse of Theorem 3 from (18) and (19).

2) Achievability: We give an outline of the achievability proof here. We use weak typicality here. Fix an $\epsilon > 0$ and the conditional distribution $P_t(y \mid x)$ that satisfies $\overline{D_{xy}} \leq \Delta$.

We analyze the error events occur during enrollment and identification phases as follows.

Enrollment. Calculate $P(y) = \sum_x Q_s(x) P_t(y \mid x)$. Randomly generate a code $\mathcal{C}_Y$ consisting of $M_C = 2^{NR_C}, R_C = I(Y;Z)$ sequences Y^N drawn i.i.d. with probability $\prod_{i=1}^{N} P(y_i)$. Reveal this code to the enrollment and identification parts.

For a given content-sequence, the encoder first constructs a list of codewords $\mathcal{L}(x^N(w))$ such that $(x^N(w), y^N(w')) \in \mathcal{A}_\epsilon^{(N)}(XY)$ for all $w = 1 \cdots M$, i.e., $\mathcal{L}(x^N(w)) = \{y^N(w') \in \mathcal{C}_Y : (x^N(w), y^N(w')) \in \mathcal{A}_\epsilon^{(N)}(XY), w' = 1 \cdots M_C\}$. From each set $\mathcal{L}(x^N(w))$ a codeword $y^N(w') \in \mathcal{L}(x^N(w))$ will be chosen uniformly at random and assigned to the content-sequence $x^N(w)$. The distortion is as expected $D_{xy}(X, Y)$ because of the law of large numbers.

An error can occur if the encoder can not construct such list for each content-sequence. This error can be made $\leq 2\epsilon$ letting $N \to \infty$, if $R_C > I(X;Y)$.

Identification. Let $x^N(1)$ be encoded to $y^N(1)$ at the enrollment phase and, as a result, z^N be observed at the identification phase.

The following errors can occur during the decoding part as follows:

(1) an error occurs if $(Y^N(1), Z^N) \notin \mathcal{A}_\epsilon^{(N)}(YZ)$, the probability of this event is $\leq \epsilon$ for sufficiently large N.

(2) an error occurs if there exists $y^N(w'), w' \neq 1$ such that $(Y^N(w'), Z^N) \in \mathcal{A}_\epsilon^{(N)}(YZ)$. This error can be $\leq \epsilon$ for $R_C \leq I(Y;Z)$ and sufficiently large N.

(3) an error occurs if there exists $x^N(w), w \neq 1$ such that $x^N(w)$ is encoded to $y^N(1)$. In other words, this error occurs if the same $y^N(1)$ is assigned to two or more different content-sequences. It can be shown that the probability of occurrence of this event is $\approx 2^{NR} 2^{-NI(X;Y)} 2^{-N(R_C - I(X;Y))} = 2^{-N(R_C - R)}$. The second term $2^{-NI(X;Y)}$ is the probability that $(Y^N(1), X^N(w)) \in \mathcal{A}_\epsilon^{(N)}(XY)$ and as a result $Y^N(1) \in \mathcal{L}(x^N(w))$. The third term $2^{-N(R_C - I(X;Y))}$ is the probability that $y^N(1) \in \mathcal{L}(x^N(w))$ is chosen uniformly at random from list $\mathcal{L}(x^N(w))$ since $|\mathcal{L}(x^N(w))| \approx 2^{N(R_C - I(X;Y))}$. Consequently, this error can be $\leq \epsilon$ for $R \leq R_C$ and sufficiently large N.

From above analysis we have $I(X;Y) \leq R_C \leq I(Y;Z)$ and $R \leq R_C$, this implies (12) and concludes the achievability proof.

Remark 1. *Although the encoding procedure in the coding-based ACFP resembles the Rate distortion theory [9], the main difference is that in the Rate distortion theory several content sequences can be mapped to a single codeword, but in coding-based ACFP each content sequence should be mapped to a unique codeword. In other words, in the coding-based ACFP there is an one-to-one relationship between content sequences and codewords, while in the Rate distortion theory there can be a many-to-one relationship.*

D. Gaussian Source

Similar to Section II-C, we consider a Gaussian source with variance V_X and mean zero, and an AWGN channel with variance V_N.

Theorem 4. *Considering distortion as the mean-squared error, the maximum of achievable identification rate based on ACFP using the coding scheme is given by*

$$R_{\mathrm{VQ}}(\Delta) = \frac{1}{2}\log_2\left(\frac{1}{1-\rho^2}\right), \tag{20}$$

where $(1-\Delta/V_X) \le \rho^2 < 1, 0 < \Delta \le V_X$ *is a solution of*

$$2\rho^2 + 2\rho\sqrt{\rho^2 - \left(1-\frac{\Delta}{V_X}\right)} - \left(1-\frac{\Delta}{V_X}\right) = \frac{V_N}{V_X}\left(\frac{\rho^2}{1-\rho^2}\right).$$

Proof. Consider a joint density $p(x,y)$ where the marginal density $p(x)$ is Gaussian with variance V_X and let ρ be the normalized correlation between X and Y. Consider also a Gaussian density $g(x,y)$ with the same covariance matrix as $p(x,y)$. We have

$$\begin{aligned} 0 &\le \int p(x,y)\log\frac{p(x\mid y)}{g(x\mid y)}\mathrm{d}x\mathrm{d}y \\ &= -h(X\mid Y) - \int p(x,y)\log g(x\mid y)\mathrm{d}x\mathrm{d}y \\ &\overset{(a)}{=} -h(X\mid Y) - \int g(x,y)\log g(x\mid y)\mathrm{d}x\mathrm{d}y \\ &= -h(X\mid Y) + \frac{1}{2}\log 2\pi e V_X(1-\rho^2), \end{aligned} \tag{21}$$

where (a) follows from the fact that $\log g(x\mid y)$ is a quadratic function of x and y, and $p(x,y)$ and $g(x,y)$ have the same covariance matrix. Hence,

$$h(X\mid Y) \le \frac{1}{2}\log 2\pi e V_X(1-\rho^2). \tag{22}$$

Consequently, if we replace $p(x,y)$ by the Gaussian $g(x,y)$, the conditional entropy $h(X\mid Y)$ can only increase. Thus, since $h(X)$ is fixed, the mutual information $I(X;Y) = h(X) - h(X\mid Y)$ can only decrease. Moreover, by replacing $p(x,y)$ with $g(x,y)$, $I(Y;Z)$ is maximized, since the observation channel is AWGN and $g(y)$, i.e., the distribution of the channel input, is Gaussian now. Therefore, any R with $R \le I(Y;Z)$ subject to $I(X;Y) \le I(Y;Z)$ is also achievable with a Gaussian assignment. Furthermore, since the mean-squared error distortion only depends on the second moment, replacing $p(x,y)$ with $g(x,y)$ does not violate the distortion constraint.

As such, we need only look at Gaussian assignments $g(x,y)$ in which $I(Y;Z) = 1/2\log_2(1+V_Y/V_N)$, $I(X;Y) = \frac{1}{2}\log_2(1/(1-\rho^2))$, $I(Y;Z) = 1/2\log_2(1+V_Y/V_N)$, and $E[(X-Y)^2] = V_X+V_Y-2\rho\sqrt{V_XV_Y} \le \Delta$. Hence, to evaluate the maximum identification rate in the coding-based ACFP, we should solve the following optimization problem:

$$\begin{aligned} \max_{\rho,V_Y} \quad & \frac{1}{2}\log_2\left(1+\frac{V_Y}{V_N}\right) \\ \text{subject to} \quad & \frac{1}{2}\log_2\left(\frac{1}{1-\rho^2}\right) \le \frac{1}{2}\log_2\left(1+\frac{V_Y}{V_N}\right) \\ & V_X + V_Y - 2\rho\sqrt{V_XV_Y} \le \Delta. \end{aligned} \tag{23}$$

It can be shown that the maximum occurs for such a ρ and V_Y that satisfies the constraints with equality. □

Remark 2. *From Theorem 2 one can conclude that for* $\Delta = 0$ *the identification rate in the general ACFP converges to the identification rate in PCFP. However, in the coding-based ACFP this convergence, for very small* Δ*, only happens if the variance of the observation channel* V_N *is very small, since in the coding-based ACFP the modified content sequence can be only chosen from a predefined set of codewords.*

Remark 3. *Similar to the Rate distortion theory under Gaussian assumptions, in the coding-based ACFP* (20) *the distortion* Δ *can not be bigger than the variance of the content-sequence* V_X*, i.e.,* $\Delta \le V_X$*. Moreover, for* $\Delta = V_X$ *the maximum identification rate for the general ACFP is* $1/2\log_2(1+4V_X/V_N)$*, but in the coding-based ACFP this is equal to* $1/2\log_2(4V_X/V_N)$*. These identification rates for* $\Delta = V_X$ *are getting close to each other as* V_X/V_N *is getting larger.*

IV. Comparison with PCFP and DWM

In this section, we compare the performance of the content-based identification based on PCFP, DWM, ACFP and coding based ACFP in terms of the identification rate. We analyze their performance under Gaussian assumptions. The content-sequences are drawn i.i.d. from a zero mean Gaussian distribution with variance V_X. And, the observation channel $Q_c(z\mid y)$ is an AWGN channel with noise variance V_N.

In the content-based identification setup based on PCFP, the encoder is an identity function, i.e., $Y^N = X^N$. Hence, the capacity of identification based on PCFP is given by [1]

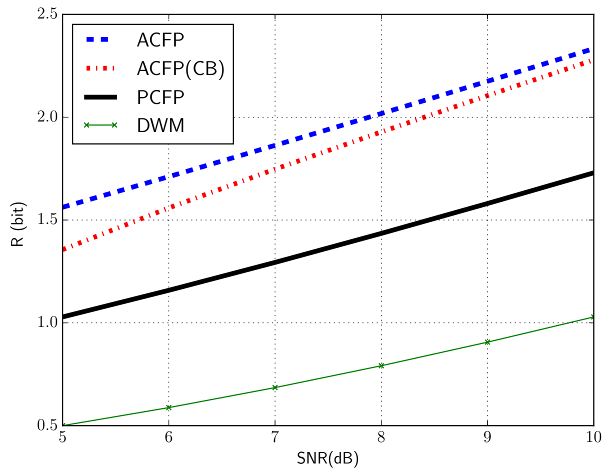
$$\begin{aligned} C_{\mathrm{PCFP}} &= I(X;Z) \\ &= \frac{1}{2}\log_2\left(1+\frac{V_X}{V_N}\right). \end{aligned}$$

The capacity of the content-based identification based on DWM, in which a unique message is embedded into each content-sequence, and under the same distortion constraint as ACFP is given by [5]

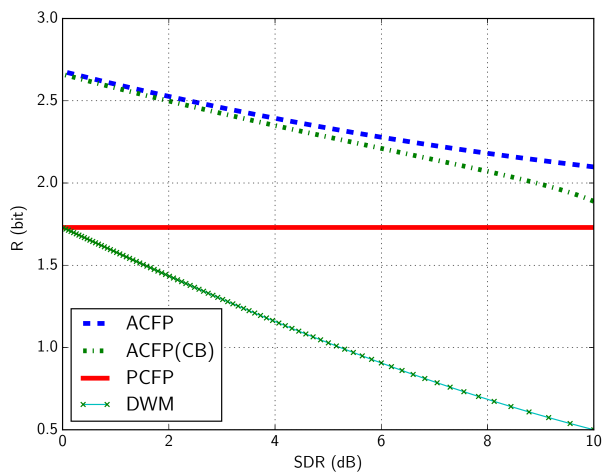
$$C_{\mathrm{DWM}} = \frac{1}{2}\log_2\left(1+\frac{\Delta}{V_N}\right).$$

Fig. 3 shows the maximum identification rates based on ACFP versus PCFP and DWM for various Signal-to-Noise Ratios (SNRs), and the Signal-to-Distortion Ratios (SDRs). As shown in Fig. 3, DWM has the worst performance with respect to other identification schemes due to the fact that in the DWM setup a message independent from a content-sequence is embedded into the content-sequence. As such the identification performance depends on the message power which is equal to distortion Δ [5]. Contrary to DWM in ACFP the introduced modification is "correlated" with the host that leads to the higher identification performance in terms of the identification rate.

From Fig. 3, one can conclude that ACFP using arbitrary modification or coding-based scheme outperforms PCFP and



(a) SDR = 5 dB



(b) SNR = 10 dB

Fig. 3. The maximum identification rates based on ACFP using arbitrary modification and coding-based schemes, PCFP and DWM for various Signal-to-Noise Ratios (SNRs) and Signal-to-Distortion Ratios (SDRs) under Gaussian assumptions.

DWM. Moreover, the general ACFP outperforms the coding-based ACFP scheme due to the first constraint in (23). However, as shown in Fig. 3 the gap between these two ACFP schemes reduces either as SNR increases (See Fig. 3a), since increasing SNR makes the first constraint in (23) being valid for any value of $0 \leq \rho < 1$, or as SDR decreases, which is mentioned in Remark 3 (See Fig. 3b).

V. Conclusions

In this paper, we analyzed the performance of content-based identification based on two different active content fingerprinting schemes. We evaluated the capacity of identification based on active content fingerprinting using arbitrary modification and determined the optimal modification scheme under Gaussian assumptions.

Moreover, to investigate the possibility of reducing the search complexity in identification systems using ACFP, we analyzed and could determine the achievable region of identification using coding-based model. We specifically investigated the optimal coding scheme under Gaussian assumptions. The theoretical findings under Gaussian assumptions showed that using ACFP can significantly improve the identification rate w.r.t. conventional content fingerprinting and digital watermarking. Although using the ACFP coding-based scheme leads to drop of identification rate w.r.t. the arbitrary modification scheme, using structured codes can tremendously reduce the search complexity.

Acknowledgement

This paper was supported by Swiss NSF project P2GEP2_151836.

References

[1] F. Willems, T. Kalker, J. Goseling, and J.-P. Linnartz, “On the capacity of a biometrical identification system,” in *IEEE ISIT*, 2003.

[2] J. Fridrich, “Robust bit extraction from images,” in *MCS*, vol. 2, 1999.

[3] J. Haitsma and T. Kalker, “A highly robust audio fingerprinting system.” in *ICM Information Retrieval,*, 2002.

[4] S. I. Gel’fand and M. S. Pinsker, “Coding for Channel with Random Parameters,” *Problems of Control Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[5] M. Costa, “Writing on dirty paper,” *IEEE tarns. on IT*, vol. 29, no. 3, pp. 439–441, May 1983.

[6] S. Voloshynovskiy, F. Farhadzadeh, O. Koval, and T. Holotyak, “Active content fingerprinting: A marriage of digital watermarking and content fingerprinting,” in *IEEE WIFS*, 2012.

[7] F. Farhadzadeh, S. Voloshynovskiy, T. Holotyak, and F. Beekhof, “Active content fingerprinting: Shrinkage and lattice based modulations,” in *IEEE ICASSP*, 2013.

[8] F. Farhadzadeh, F. Willems, and S. Voloshynovskiy, “Information theoretical analysis of identification based on active content fingerprinting,” in *SIT in Benelux*, May 2014, pp. 134–139.

[9] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York, NY, USA: Wiley-Interscience, 1991.