

1 International Journal of Image and Graphics
 Vol. 5, No. 1 (2005) 1–31
 3 © World Scientific Publishing Company



5 **INFORMATION-THEORETIC DATA-HIDING:
 RECENT ACHIEVEMENTS AND OPEN PROBLEMS**

7 SVIATOSLAV VOLOSHYNOVSKIY*, FREDERIC DEGUILLAUME†,
 OLEKSIY KOVAL‡ and THIERRY PUN§

9 *Stochastic Image Processing (SIP) Group, Department of Computer Science,
 University of Geneva, 24 rue du Général Dufour, Geneva CH-1211, Switzerland*

*svolos@cui.unige.ch

11 †Frederic.Deguillaume@cui.unige.ch

‡Oleksiy.Koval@cui.unige.ch

13 §Thierry.Pun@cui.unige.ch

15 Received 4 September 2003

Revised 6 December 2003

Accepted 23 December 2003

17 In this paper we introduce and develop a framework for visual data-hiding technologies
 that aim at resolving emerging problems of modern multimedia networking. First, we
 19 introduce the main open issues of public network security, quality of services control and
 secure communications. Secondly, we formulate digital data-hiding into visual content
 21 as communications with side information and advocate an appropriate information-
 theoretic framework for the analysis of different data-hiding methods in various
 23 applications. In particular, Gel'fand-Pinsker channel coding with side information at
 the encoder and Wyner-Ziv source coding with side information at the decoder are
 25 used for this purpose. Finally, we demonstrate the possible extensions of this theory for
 watermark-assisted multimedia processing and indicate its perspectives for distributed
 27 communications.

29 *Keywords:* Data-hiding; robust watermarking; copyright protection; tamper proofing;
 secure communications; steganography; network security; network quality-of-service.

1. Introduction

31 The mass diffusion of digital media and the explosive growth of telecommunication
 are reshaping the lifestyles of ordinary people, research and industry. Over the last
 33 decade, the rise of digital telecommunication technologies (including ATM, PSTN,
 ISDN, ADSL, IP networks) has fundamentally altered how people work, think,
 35 communicate, and socialize. New emerging audio/visual applications have recently
 appeared thanks to the public multimedia networks. This growth is especially
 37 observed in networked video applications such as video phones, video conferencing,

*Corresponding author.

2 *S. Voloshynovskiy et al.*

1 video e-mail, video streaming, digital TV, high-definition TV (HDTV), video on
demand (VoD), distance learning, remote collaboration and surveillance.

3 However, despite the obvious progress of public networks, these developments
carry with them a number of risks such as copyright violation, prohibited usage
5 and distribution of digital media, secret communications, and network security.
Therefore, security, scalability and manageability amongst others have become
7 issues of serious concern, as current solutions do not satisfy anymore the growing
demands of multimedia communications.

9 In the scope of this paper, we will focus on a possible solution for public network
security in order to prevent unauthorized data exchange and to ensure secure
11 communications. Then two main objectives will be addressed: the first one is to in-
troduce and to overview a novel approach to multimedia security in public networks
13 that is based on data-hiding technologies. We will consider fundamentals of digital
data-hiding technologies in comparison with the traditional means of multimedia
15 security. A basic theoretical model of a data-hiding system will be analyzed, and we
will demonstrate the relevance of data-hiding problems to digital communications.
17 We will show the advantages of data-hiding based multimedia security protocols
over the traditional general means of security based on encryption, scrambling and
19 firewall systems. The second objective of the paper is to demonstrate some of the
main achievements in the field of digital data-hiding technologies for multimedia
21 network security (including copyright issues) and quality-of-service control. We will
present the state-of-the-art solutions for copyright protection of digital media, in-
23 tegrity verification, detection of modifications of the multimedia content, and secure
communications.

25 The paper is organized as follows: Section 2 outlines the main open issues in
public networks. Sections 3 and 4 introduce an alternative approach to public
27 network security and secure communications based on digital data-hiding and
present the main problems to be addressed. Section 5 is dedicated to robust
29 watermarking. Section 6 considers authentication, tamper proofing and watermark-
assisted communications, and Sec. 7 presents secure communications. Section 8
31 concludes the paper.

Notation

33 We use capital letters to denote scalar random variables X , bold capital letters to
denote vector random variables \mathbf{X} , corresponding small letters x and \mathbf{x} to denote
35 the realizations of scalar and vector random variables, respectively. The superscript
 N is used to denote length- N vectors $\mathbf{x} = x^N = \{x[1], x[2], \dots, x[N]\}$ with i th
37 element $x[i]$. We use $X \sim p_X(x)$ or simply $X \sim p(x)$ to indicate that a random
variable X is distributed according to $p_X(x)$. The mathematical expectation of a
39 random variable $X \sim p_X(x)$ is denoted by $E_{p_X}[X]$ or simply by $E[X]$ and $\mathbf{Var}[X]$
denotes the variance of X . Calligraphic fonts \mathcal{X} denote sets $X \in \mathcal{X}$ and $|\mathcal{X}|$ denotes
41 a cardinality of set.

1 **2. Open Issues of Public Networks**

3 An extremely rapid development of telecommunications has been observed in the
4 last decades. First of all, it concerns general public networks such as asynchronous
5 transfer mode (ATM), public switched telephone networks (PSTN), integrated
6 service digital network (ISDN), asynchronous digital subscriber line (ADSL) and
7 cable and Internet protocol (IP) service. In the recent years, the explosion of
8 audio/visual communications has caused additional interest towards multimedia
9 applications of public networks that are characterized by a huge amount of data
10 to be stored or communicated preferably in real time. The data storage become
11 extremely distributed, which leads to a number of problems related to reliable
12 and secure data transportation, distributed computation and data management in
13 distributed environments. Therefore, some questions of traditional communications
14 should be reconsidered to satisfy new growing requirements. An ideal multimedia
15 network is supposed to be capable of transferring reliably and securely any amount
16 of information without delay and loss. Unfortunately, fundamental practical limi-
17 tations and outdated main design principles of current networks do not meet these
18 demands. With respect to that, security and quality-of-services are the main issues
19 of concern (but not the only ones).

19 **2.1. Network quality-of-service control**

20 Practical public networks have a number of open issues related to quality-of-service
21 (QoS) control. In practice, most of public networks cannot guarantee quality-of-
22 service control due to many technical reasons, the main ones being³⁴:

- 23 • Network errors (bit errors), loss and deletions (packet or bit loss), and inser-
24 tions (cross talks): data can be lost in the network for a variety of reasons,
25 including congestion, rejection due to excessive delay, and network fault. These
26 problems can manifest themselves as unnatural artifacts in the video and images,
27 such as missing frames, lines, or blocks. Severe loss, such as from heavy network
28 congestion, can cause the video playback to be stopped until the receiver can
29 resynchronize. Moreover, errors arising from lost data can affect multiple video
30 frames by temporal error propagation.
- 31 • Delay (real time), jitter (timing errors) and latency: streaming multimedia should
32 fit delay constraints since the video must be decoded and played in real-time. If
33 the video data spends too much time in the network, it is useless even if it arrives
34 at the receiver. Buffering can reduce the effect of delay and jitter (timing errors).
35 Latency can also be an issue when two-way communication is necessary.
- 36 • Finite bandwidth (network sharing, limited resources): bandwidth is the amount
37 of data that can traverse the network or a part of the network at any given
38 time. Network bandwidth is a shared, limited resource and will vary with time.
39 A network may not be able to guarantee that the required bandwidth for
40 transporting multimedia data will be available.

4 *S. Voloshynovskiy et al.*

1 In addition heterogeneity and time-variance are important factors for public
2 network multimedia communications. A heterogeneous network is a network whose
3 parts (sub-networks) may have vastly unequal resources. For example, some parts
4 of a heterogeneous network may have abundant bandwidth and excellent congestion
5 control while other parts of the network are overloaded and congested by overuse
6 or by a lack of physical network resources. Different receivers on a heterogeneous
7 network can experience different performance characteristics. When streaming
8 video over a heterogeneous network, the video stream should be decodable at
9 optimal quality for users with a good network connection, and at usable quality
10 for users with a poor connection. Time-variance implies that bandwidth, delay,
11 loss, or other network characteristics can significantly vary over time, sometimes
12 drastically changing in a matter of seconds. The Internet (based on the under-
13 lying internet protocol, IP) main current flow control is the transmission control
14 protocol (TCP), which is based on a “best effort” type traffic,^{4,57} intended for
15 ordered reliable delivery of stream of data such as file transfer applications. The
16 second principal flow control of the Internet is the user datagram protocol (UDP),
17 typically used for audio or video streaming, but it does not guarantee the integral
18 delivery of the data being transmitted and offers no control over the data loss ratio.
19 This makes Internet a difficult network for transporting real-time multimedia data,
20 which is then an example of a heterogeneous, time-varying, network with unpre-
21 dictable congestion and delays and no QoS control. Some proposals, which we will
22 not discuss here, have been proposed to add QoS to the Internet, including: multiple
23 service classes,²⁹ usage-based schemes,¹³ priority pricing,³⁸ or a bandwidth/buffer
24 allocation equilibrium approach.³⁶ However if a high QoS is easy to obtain for
25 small networks, it appears then that a trade-off should be found between the lowest
26 service granularity of the QoS and the scalability of any proposed system.⁶⁰

27 **2.2. Network security**

28 Network security is no less important a problem of public networks.²⁸ Due to
29 the open nature of data communication protocols in public networks, special care
30 about access control, authentication, secure delivery and intrusion detection¹ should
31 be taken. Traditional means of network security such as firewalls, virtual private
32 networks (VPN) and intrusion detection systems (IDS) are well suited for specia-
33 lized applications where the above questions can be under control in the range of
34 some closed trusted environment. However, these solutions are not appropriate for
35 highly distributed public environments. Moreover, the definition of traffic *disre-*
36 *gards the nature of the multimedia content*, interpreting any kind of multimedia as
37 purely digital flow. This has a huge impact on many aspects of network security
38 since modern secret communication tools based on steganography, viruses based
39 on content embedding and content management systems are designed using com-
40 pletely different interpretation of the multimedia content. Moreover, contrarily to
41 the traditional approaches, new public distributed networks cannot rely anymore

1 on file headers, meta data or centralized bodies. Recent examples of distributed
2 networks based on peer-to-peer (P2P) communications²³ practically demonstrate
3 how easily huge amount of information can be exchanged in completely uncon-
4 trolled manner leading to secret communications, copyright violation and prohibited
5 content distribution.⁵² Traditional network security can hardly cope with these
6 requirements. The copyright/intellectual property issue is one of the major obstacle
7 to multimedia networking.²³

8 Internet as a public network is a very challenging environment for secure and
9 reliable transport of real-time multimedia data. Internet is a heterogeneous and
10 time-varying network, has no QoS control and no efficient and reliable mechanisms
11 for copyright protection, access control and secure communications. Therefore, we
12 will below consider potential approaches that make at least partially possible to
13 satisfy (or to complement existing solutions to) the above emerging problems.

3. Multimedia Security and Quality-of-Service: Main Open Issues

14 Multimedia content security has a number of specific requirements that should allow
15 to answer to the following questions:

- 16 • Who has issued the multimedia content?
- 17 • Who is the content owner?
- 18 • When was the content issued?
- 19 • Who has access right to the content?
- 20 • Is the content modified?
- 21 • Where was the content modified?
- 22 • What was the original content before modification?
- 23

24 It is obvious that specialized protocols or hardware alone are not able to resolve
25 all these questions. At the same time, new emerging requirements to secure Inter-
26 net communications essentially extended the horizons of traditional cryptography
27 protocols. A new paradigm has to answer not only the question of how to commu-
28 nicate in a secure way but also how to communicate in a completely undetectable
29 way over public networks. This subsequently leads to the extension of the concept
30 of covert communications and requires the creation of new “covert” multimedia
31 channels. Finally, the related open problem of quality-of-service control requires
32 one to provide the answer on the question of how to communicate in a robust way
33 providing end-to-end services?

34 The list of these diverse problems seems to be very broad and from a traditional
35 point of view there does not seem to exist any common means of satisfying all
36 these requirements. However, there are some common aspects of secure and reliable
37 communications that could be addressed by novel technologies based on digital
data-hiding.

1 4. Multimedia Data-Hiding

3 Multimedia data-hiding represents an alternative concept for public network secu-
4 rity and secure communications and can be considered as an assisting functionality.
5 Multimedia data-hiding provides an additional “virtual” or covert channel of digital
6 communications through the embedding of some secret unperceived information
7 directly into the multimedia content without extra meta data, headers, sophisti-
8 cated specialized formats and attachments. This naturally leads to the concept of
9 a *smart media* where the features of the multimedia content are extended to extra
10 functionalities that can be exploited for multimedia processing, communications,
11 security and content management.

12 The concept of a smart media can be generally characterized as:

- 13 • self-sufficient or self-embedding (no extra headers, meta data and attachments
14 are needed);
- 15 • self-synchronizing (no synchronization information is used on the protocol side);
- 16 • self-authenticating (no access to the original data is required to establish the
17 content authenticity);
- 18 • self-correcting and offering error concealment (no format protocol modifications
19 related to forward/backward error corrections are necessary).

20 Besides the obvious advantages listed above, multimedia data-hiding should
21 additionally provide:

- 22 • perceptually invisible data embedding;
- 23 • robust and content independent extraction of embedded information;
- 24 • scalability of the content for different heterogeneous networks and applications;
- 25 • security provided by a proper key management and undetectability of the hidden
26 data presence by the existing detection tools.

27 We consider multimedia data-hiding with respect to three main applications
28 that should address the open issues presented in Sec. 3:

- 29 • robust watermarking;
- 30 • authentication and tamper proofing;
- 31 • secure communications.

32 Robust watermarking is mainly used for copyright protection, content tracking
33 and content self-labeling. Authentication and tamper proofing target the certifica-
34 tion of a given content originator, as well as the verification of its integrity and the
35 detection of local modifications in images, video and documents, or recovering of
36 the original content based on available copy of the modified or tampered content.
37 Finally secure communications address the issue of secure content delivery over
38 the public networks using two different possibilities. The first possibility is a visual
39 “encryption” or scrambling that should provide additional error resilience in the
40 case of wireless networks and networks with packet losses and erasures. The second

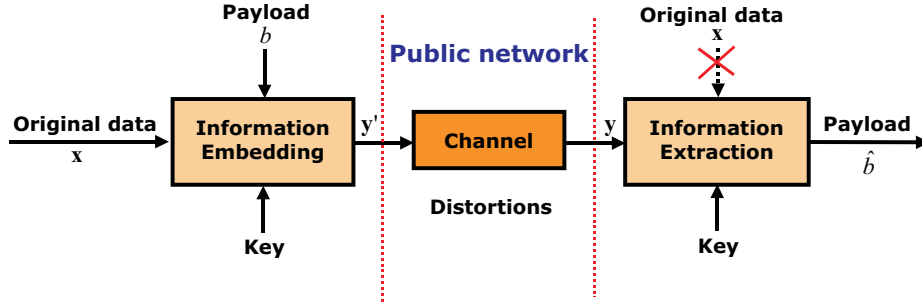


Fig. 1. Generalized diagram of robust watermarking.

1 possibility is steganography that guarantees secure content delivery by hiding the
 2 content to be securely communicated into the covert media whereas the presence
 3 of the hidden content presence should not be detected by various detection tools.

5. Robust Watermarking

5 Robust watermarking is one of the most challenging research directions of data-
 6 hiding combining a number of multidisciplinary issues ranging from information
 7 theory and digital communications, estimation and detection theory, to image
 8 processing and computer vision. Robust watermarking, from the information-
 9 theoretic perspective, should provide the reliable communication of some energy
 10 constrained payload m in the body of a multimedia content under a broad list
 11 of various intentional and unintentional attacks, those attacks constituting the
 12 resulting watermarking channel. The protocol describing robust watermarking can
 13 be schematically explained as in Fig. 1. This protocol consists of three main parts,
 14 i.e. information embedding or encoder, channel that represents the public network
 15 and the information extraction part or decoder.

The goal of the information embedder consists in the invisible “integration” of a specifically preprocessed payload m into the original (host) content \mathbf{x} based on some secret key K . We assume that the message M , uniformly distributed over the message set \mathcal{M} with the cardinality $|\mathcal{M}|$, is encoded based on a secret key into some watermark $\mathbf{w} = w^N = \{w[1], w[2], \dots, w[N]\}$ and embedded into a host data (image) $\mathbf{x} = x^N = \{x[1], x[2], \dots, x[N]\}$. We denote \mathbf{x} to be a two-dimensional sequence typically representing the luminance of the original image. The i th element of \mathbf{x} is denoted as $x[i]$ where $i = (n_1, n_2)$ and $\mathbf{x} \in \mathbb{R}^N$ and $N = N_1 \times N_2$ is the size of the host image. The message m typically has a length of 64 bits, i.e. $|\mathcal{M}| = 2^{64}$ and is content independent. In some cases, only a binary decision about the watermark presence/absence can be required: $|\mathcal{M}| = 2$ (so-called 1-bit watermarking, i.e. $\log_2 |\mathcal{M}| = 1$ bit). As another example, the printing industry only requires 16 bits for document tracking aiming at identifying the distribution channels. In any case, the payload for robust watermarking is relatively modest and

8 *S. Voloshynovskiy et al.*

rarely exceeds 100 bits. The embedding rule can be expressed as a mapping:

$$w[i] = f_i(m, x^i), \quad (1)$$

$$y'[i] = x[i] + w[i], \quad (2)$$

1 where m is a particular realization of the random message M , $y'_i, 1 \leq i \leq N$
 2 is the stego data and x^i can be used only partially as $x^i = \{x[1], x[2], \dots, x[i]\}$
 3 (so-called causal side information), or entirely $x^i = \{x[1], x[2], \dots, x[N]\}$ (so-called
 4 non-causal side information). We also include in this generalized model both spread
 5 spectrum schemes and host interference cancellation schemes based on pre-coding,
 6 according generally to Gel'fand-Pinsker,²⁰ and particularly to Costa coding⁸ de-
 7 veloped for Gaussian content. The setup (1) is quite general and can include
 8 different particular cases of watermarks. For example, watermark $\mathbf{w} \in \mathbb{R}^N$ and
 9 $W[i] \sim \mathcal{N}(0, \sigma_w^2)$, i.e. zero-mean Gaussian, $\mathbf{w} \in \{\pm 1\}^N$, i.e. pseudo random binary
 10 watermark, $\mathbf{w} \in \{-1\}^z \{+1\}^z$, i.e. pseudo random block-repeated watermark,
 11 $\mathbf{w} \in (-\Delta/2; +\Delta/2)^N$ and $W[i] \sim U(-\Delta/2; +\Delta/2)$, i.e. uniform watermark. The
 12 admissible distortion for watermark embedding is D_1 :

$$13 \quad E[d_1^N(\mathbf{X}, \mathbf{Y}')] \leq D_1, \quad (3)$$

14 where $d_1^N(\mathbf{X}, \mathbf{Y}') = \frac{1}{N} \sum_{i=1}^N d_1(x[i], y'[i])$ denotes N -vector distortion between
 15 vectors \mathbf{X} and \mathbf{Y}' and $d_1(x[i], y'[i])$ denotes element-wise distortion between i th
 16 elements $x[i]$ and $y'[i]$.

17 The channel is characterized as a transition probability $p(y|w, x)$, and can be
 18 quite general including both signal processing and geometrical distortions of the
 19 stego data. In the particular case of intentional attacks, the attacker aims at
 20 removing the watermark \mathbf{w} from \mathbf{y}' producing the attacked data \mathbf{y} . The admissible
 21 attacker distortion is D_2 that is defined in the same way as (3) between vectors \mathbf{y}'
 22 and \mathbf{y} :

$$23 \quad E[d_2^N(\mathbf{Y}', \mathbf{Y})] \leq D_2. \quad (4)$$

24 One should also note another possibility to define the attacker distortion between
 25 the original data \mathbf{x} and the attacked data \mathbf{y} . The decoder produces the estimate of
 26 \hat{M} based on \mathbf{y} using:

$$27 \quad \hat{m} = g(y^N), \quad (5)$$

28 where $g(\cdot)$ denotes the decoding rule and $\mathbf{y} = y^N = \{y[1], y[2], \dots, y[N]\}$ is the
 29 distorted stego data. The decoding error occurs when $\hat{M} \neq M$. A particular case of
 30 generalized decoding rule $g(\cdot)$ is the maximum a posteriori (MAP) decoding rule,
 31 which minimizes the probability of error:

$$32 \quad \hat{m} = \arg \max_{m \in M} p(m|y^N). \quad (6)$$

33 The cryptographic security of a robust watermarking system is considered as the
 system "immunity" against message removal or estimation using knowledge of the

1 algorithms (2) and (5). The blind cryptographic attack, that can be applied without
 2 the knowledge of the secret key K , can be simply designed as an exhaustive search
 3 procedure over all possible values of the watermark \mathbf{w} . The number of all possible
 4 watermarks to be tested in such a way is determined by the entropy of the water-
 5 mark. It is obvious that under the constraint $E[\mathbf{w}^2] \leq D_1$ the maximum entropy
 6 of a watermark with Gaussian p.d.f. is $h(W) = \frac{1}{2} \log_2(2\pi e D_1)$. In many practical
 7 watermarking algorithms such as quantization index modulation (QIM),⁶ scalar
 8 Costa scheme (SCS)¹⁴ or distortion compensated QIM (DC-QIM), the watermark
 9 code book is structured by some binning strategies that aims at host interference
 10 cancellation (or watermark invariance to the values of host images) and invari-
 11 ance to the geometrical transforms. This leads to scalar or vector quantization
 12 encoding strategies that represent the regular lattices aiming at overcoming the
 13 shaping loss (to be as close as possible to Gaussian p.d.f.), and periodical water-
 14 mark spatial tiling to resist against affine and projective transforms. This leads to
 15 the reduction of randomness or ambiguity at the watermark decoding. Obviously,
 16 code book structuring is known as a part of the algorithm for the attacker. However,
 17 it also reduces the entropy of the watermark as any conditioning $h(W) \geq h(W|SC)$
 18 where SC is a structure of the code book and gives more information leakage for
 19 the attacker. Therefore, some special care should be taken to prevent this leakage
 20 and to apply the data-hiding in such a way that this leakage will not be crucial
 21 for a given application. The first attempt to formalize the security of the robust
 22 watermarking technologies has been done by Barni, Bartolini and Furon.²

23 The generalized diagram from Fig. 1 can be considered in a simplified version
 24 shown in Fig. 2 when the channel is considered as an i.i.d. additive white Gaussian
 25 noise (AWGN) channel ($Z[i] \sim \mathcal{N}(0, \sigma_z^2)$). The channel consists of two sources:
 26 the host data $\mathbf{x}(X[i] \sim \mathcal{N}(0, \sigma_x^2))$ and the noise \mathbf{z} . Both sources are acting as an
 27 interference for the payload m . Shannon's theory states that if \mathbf{x} is known at both
 28 ends, then the channel capacity is equal to its theoretical upper bound⁹:

$$29 \quad C = \max_{p_X(x): E[\mathbf{w}^2] \leq \sigma_w^2} I(W; Y) = \frac{1}{2} \log_2 \left(1 + \frac{\sigma_w^2}{\sigma_z^2} \right). \quad (7)$$

30 If \mathbf{x} is not known at both ends, then it acts as a strong interference. In the
 31 case of watermarking, the host data \mathbf{x} is available at the encoder. Therefore, this
 scheme can be considered as communication with side information available at the

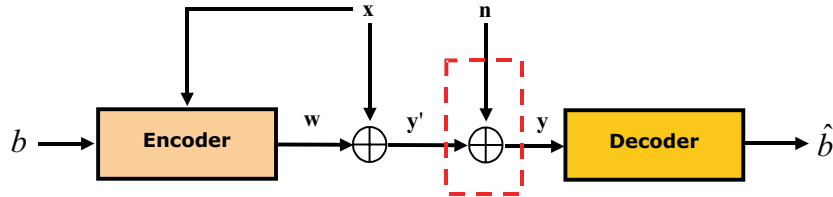


Fig. 2. Robust watermarking as communications with side information at the encoder.

10 *S. Voloshynovskiy et al.*

1 encoder. In this case, the problem of data-hiding can be reformulated as a reliable
 3 communication of the message m over the channel with noise \mathbf{z} and interference \mathbf{x}
 5 being known at the encoder but not at the decoder. The most general formulation
 of such type of communications was considered by Gel'fand and Pinsker in 1980 in
 non-watermarking applications and the capacity of this scheme was found as²⁰:

$$C = \max_{p(u,w|x)} [I(U; Y) - I(U; X)], \quad (8)$$

7 where U is an auxiliary random variable. The Gel'fand-Pinsker problem has a quite
 simple intuitive interpretation using a random binning argument. If we denote a set
 9 of all elements (codewords) as $2^{NI(U;Y)-\varepsilon}$ and apply a random binning technique
 we assume that each bin (subset) associated to a particular message has $2^{NI(U;X)-\varepsilon}$
 11 elements. It is then easy to find the total number of uniquely distinguished bins as
 $2^{NC+2\varepsilon} = \frac{2^{NI(U;Y)-\varepsilon}}{2^{NI(U;X)-\varepsilon}}$ that directly leads to $C = I(U; Y) - I(U; X)$.

Costa (1983) has considered the above problem in the Gaussian context and
 found that:

$$I(U; Y) = \frac{1}{2} \log_2 \frac{\sigma_w^2 + \sigma_x^2 + \sigma_z^2}{\frac{\sigma_w^2}{\sigma_w^2 + \sigma_x^2} (1 - \alpha)^2 \sigma_x^2 + \sigma_z^2}, \quad (9)$$

$$I(U; X) = \frac{1}{2} \log_2 \frac{\sigma_w^2 + \sigma_x^2}{\sigma_w^2} \quad (10)$$

13 and $C = \frac{1}{2} \log_2 (1 + \frac{\sigma_w^2}{\sigma_z^2})$ and where the auxiliary random variable has a form of
 $U = W + \alpha X$ and $\alpha = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_z^2}$ is chosen to provide independence of $W - \alpha(W + Z)$
 15 and $W + Z$.

Having considered the theoretical foundations of host interference cancellation
 17 using side information at the encoder, we concentrate on the practical data-hiding
 schemes. We will assume a binary representation of $m \equiv \mathbf{b}$ of length L_b , i.e. $\mathbf{b} \in$
 19 $\{0, 1\}^{L_b}$. \mathbf{b} is encoded into a sequence of letters \mathbf{d} of length L_x with $d[i] \in \mathcal{D}$,
 where \mathcal{D} can be binary $\mathcal{D} \in \{0, 1\}$ or multilevel $\mathcal{D} \in \{1, 2, \dots, D\}$ with $D = |\mathcal{D}|$,
 21 using some suitable error correction codes and proper spreading. Additionally, the
 sequence \mathbf{d} can be modulated. The simplest case of modulation is M-PAM signal
 23 constellation that consists of $M \geq 2$ equidistant real symbols centered on the origin,
 i.e. $\mathbf{d} = \frac{d_0}{2} \{-M + 1, -M + 3, \dots, M - 1\}$ (Fig. 3), where d_0 is the minimum
 25 distance between symbols. For equiprobable symbols the average symbol energy
 is $E_{\mathbf{d}} = E[\mathbf{d}^2] = (M^2 - 1)d_0^2/12$. The highest rate for the unencoded M-PAM is
 27 $R = \log_2 M$.

However, contrarily to digital communications where the sequence \mathbf{d} is directly
 29 used for the transmission over the noisy channel, the encoded message is combined
 with the host data in digital data-hiding applications according to the additive
 31 model (2). Depending on the different embedding rules (1), we can classify all
 existing data-hiding techniques as those that do not use side information about the
 33 host data at the encoder, and those that use side information.

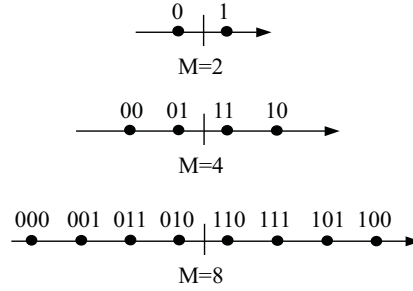


Fig. 3. M-PAM constellations.

1 Spread spectrum (SS) data embedding is historically the first and currently most
 2 often used technique in practice technique for data-hiding. The SS data-hiding does
 3 not directly use information about the host image for the watermark encoding and
 4 thus suffers from host interference:

$$5 \quad y'[i] = x[i] + w[i], \quad (11)$$

6 where \mathbf{w} can be any of the above techniques (a–d). However, in the most practical SS
 7 robust watermarking schemes proper spreading is applied for security, redundancy
 8 and geometrical attacks resistance reasons. This spreading is performed over the
 9 host data using a key-dependent spreading sequence $s \in \{\pm 1\}$ such that $w[j] =$
 10 $c[k]s[j], j \in S_k$, and where \mathbf{c} is the codeword of length L_c (particular case of \mathbf{d}
 11 encoding) that is mapped to 2-PAM, i.e. $\mathbf{c} \in \{\pm 1\}^{L_c}$ and S_k are non-overlapping
 12 subsites that are used for the allocation of each bit of codeword \mathbf{c} . Additionally, the
 13 watermark can be embedded exploiting particularities of the human visual system
 (HVS), as a so-called perceptually adapted watermarking:

$$15 \quad y'[i] = x[i] + \gamma[i]w[i], \quad (12)$$

16 where $\gamma[i]$ represents a mask adapted to the HVS in the coordinate or some
 17 transform domain such as DCT, DFT or wavelets.^{12,27,37,43,51} In most cases, an
 18 optimal ML-detector is used for the proper stochastic model of host data \mathbf{x} that
 19 takes the decision about the codeword $\hat{\mathbf{c}}$ with following appropriate soft-decoding.⁴²
 20 Moreover, channel state estimation can be applied to determine the distribution of
 21 channel noise, fading or erasures according to the model of binary symmetrical
 channels³⁰ or generalized channels⁴⁹ referring to diversity watermarking.

22 We consider three main variations of Costa's approach proposed for host inter-
 23 ference cancellation assuming data-hiding with side information: Least Significant
 24 Bit Modulation (LSBM),⁴⁷ QIM,⁶ SCS.¹⁴

The LSBM encoder embeds the data according to the next rule:

$$27 \quad y'[i] = Q(x[i]) + d[i] = x[i] + d[i] + (Q(x[i]) - x[i]) = x[i] + w[i]. \quad (13)$$

28 The image is first precoded based on an uniform quantizer $Q(x)$ with a step Δ and
 29 then the M-PAM watermark \mathbf{d} is added to this image. The embedding distortion

12 *S. Voloshynovskiy et al.*

1 is:

$$D_{y'x} = E[|y' - \mathbf{x}|^2] = \frac{\Delta^2}{12} + E[\mathbf{b}^2] = \frac{\Delta^2}{12} + \frac{(M^2 - 1)d_0^2}{12}. \quad (14)$$

3 The LSBM decoder performs the direct estimation of the message:

$$\hat{d}[i] = y[i] - Q(y[i]). \quad (15)$$

5 The binary QIM encoder performs the quantization of the host image using two
7 sets of quantizers $Q_{-1}(\cdot)$ and $Q_{+1}(\cdot)$ that are shifted by Δ with respect to each
other:

$$y'[i] = Q_d(x[i]) = x[i] + (Q_d(x[i]) - x[i]) = x[i] + w[i], \quad (16)$$

9 where $Q_d(\cdot)$ denotes the quantizer for $d = -1$ and $d = +1$. The QIM embedding
distortion is:

$$11 \quad D_{y'x} = E[|y' - \mathbf{x}|^2] = \frac{\Delta^2}{12}. \quad (17)$$

13 Therefore, the embedding distortion for the LSBM is higher than that for the QIM.
15 However, it is necessary to note that the LSBM can have the same embedding
17 distortion as the QIM, if one applies a distortion minimization procedure choosing
the resulting quantization bin with the minimum possible distortion after final
addition of the M-PAM watermark. This will not affect the capacity, but it will
decrease the embedding distortion. The QIM decoder performs the ML-estimation:

$$\hat{d} = \arg \min_{d \in \{\pm 1\}} \|y[i] - Q_d(y[i])\|^2. \quad (18)$$

19 In contrary to the LSBM and the QIM, which do not use any prior information
21 about the attacking channel state, the SCS exploits the knowledge of the AWGN
channel statistics at the encoder. The SCS encoder generates the stego data based
on the rule:

$$23 \quad y'[i] = x[i] + \alpha(Q_d(x[i]) - x[i]) = x[i] + \alpha w[i]. \quad (19)$$

25 The parameter α is optimized to resist against the AWGN attack. It should be
27 noted that when $\alpha = 1$ the SCS corresponds to the QIM, as well as to the case when
 $\sigma_z^2 \rightarrow 0$ — the high watermark-to-noise-ratio (WNR) regime. Also, a useful analogy
with the channel state information about the statistics of the AWGN channel could
be outlined. In particular, the variance of the noise should be known at the encoder
29 in advance. Since it is not the case in practical applications, Eggers *et al.* propose
to optimize α for the working dynamic range of the $\mathbf{WNR} \in [-20, +20]$ dB.¹⁴ This
31 leads to a slight decrease of the performance. The SCS embedding distortion is:

$$D_{y'x} = E[|y' - \mathbf{x}|^2] = \alpha^2 \frac{\Delta^2}{12}. \quad (20)$$

33 The probabilities of error in spread spectrum and quantization-based data-
35 hiding techniques have different characters, that drastically influence the perfor-
mance of these methods in different regimes. In particular, the probability of

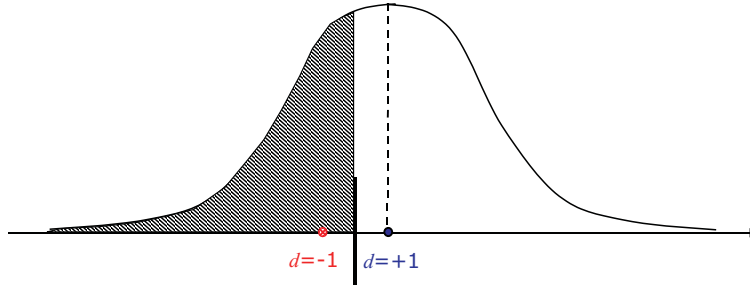


Fig. 4. Probability of error computation for $p(y|d = +1)$ and binary spread-spectrum-based embedding. The highlighted region indicates the error.

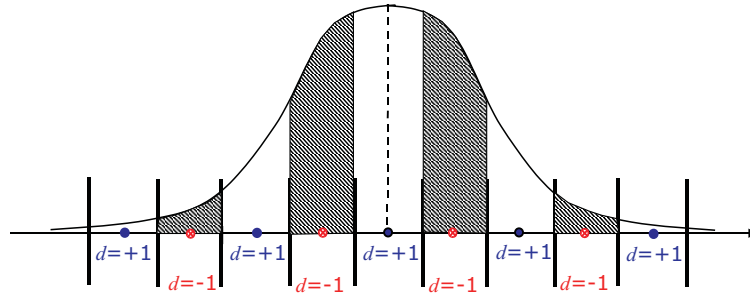


Fig. 5. Probability of error computation for $p(y|d = +1)$ and binary QIM-based embedding.

1 error for the spread-spectrum based data-hiding methods is computed as in digital
 2 communications for M-PAM modulation (Fig. 4) including the additional impact
 3 of the host data expressed as the convolution $p_Y(y) = p_X(x) * p_Z(z) * p_D(d)$.
 4 The fundamental difference with the quantization-based techniques consists in the
 5 periodic integration of the probability of the error over all bins that do not coincide
 6 with the transmitted symbol constellation (Fig. 5). Therefore, for high variance of
 7 AWGN attack, this error can increase more rapidly in contrast to the SS-based
 8 data-hiding. The reader is referred to Ref. 41 for more details about performance
 9 analysis of different data-hiding techniques under various additive attacks.

10 The performance of different watermark encoding and modulation (embed-
 11 ding) techniques can be considered depending on an operational WNR: $\mathbf{WNR} =$
 12 $10 \log_{10} \frac{\sigma_w^2}{\sigma_n^2}$ for AWGN channels. Low-rate data-hiding achieves AWGN channel
 13 capacity in low-WNR regime while high-rate data-hiding is possible for relatively
 14 high-WNR regime. The low-WNR regime is typical for robust digital watermarking
 15 when the attack is aiming at removing the watermark. In this case, the variance
 16 of the attack might be higher than the variance of the watermark. In this case,
 17 the host interference is not crucial for approaching channel capacity and spread
 18 spectrum based data-hiding can be sufficient. In this regime, two approaches to
 19 watermark encoding are mostly used in practice:

- 1 • binary watermark encoding using binary low-rate error correcting codes with soft decoding (Turbo codes, or LDPC codes);
- 3 • binary watermark encoding using the above binary codes with higher rates and following replication.

5 The first approach is characterized by a lower probability of error while the second one has a higher resistance against the cropping attack. It should be also
 7 noticed that the error correction codes for erasure channels can also be used to withstand cropping attack. Additionally, properly designed repetitive watermark
 9 can be also used to recover from geometrical attacks that are characterized by an affine transform, using self-synchronization.^{10,31,50}

11 The high-WNR regime makes it possible to increase the watermark embedding rate for the same embedding distortions. In the general case, unencoded M-ary
 13 pulse amplitude modulation (PAM) can be used to approach channel capacity within 1.53 dB (that is related to a shaping loss of uniform vs. Gaussian p.d.f.
 15 of watermark). Finally, a coded modulation is used in practice that combines M-ary signaling with binary error correcting codes such as Turbo codes or LDPC
 17 codes. The host interference cancellation embedding is also necessary for this regime. Therefore, methods based on Gel'fand-Pinsker (Costa) framework should
 19 be used. However, their performance will be very poor for the low-WNR regime. Keeping in mind relatively low watermark embedding rate required for the robust
 21 watermarking applications, the SS-based methods could be recommended. Moreover, as discussed, Costa-based methods have good performance for the AWGN
 23 attack while the spectrum of all possible attacks and corresponding noise distributions is much broader in practical applications. The simple change of noise
 25 from AWGN to additive uniform noise demonstrates the very low robustness of Costa-based methods while SS-based methods are practically insensitive to the
 27 change of noise statistics even without special adaptivity of the matched detector part.⁴¹ In addition, the question of the worst case noise (attack) is still open for the
 29 Costa-based methods and additional research is required. It is possible to foresee that the worst case attack against Costa-based methods could be based on noise
 31 with non-periodical trains of pulses that should trick the unique watermark detection. Additional attacks can be imagined where the attacker can estimate the
 33 parameters of the used quantizers in LSBM, QIM or SCS and requantize data to the original centroids $Q[x]$ providing even an enhancement of stego data quality.

35 A natural question arises then: where is the limit in the game between the data-hider and the attacker? Consequentially, what are the maximum data-hiding
 37 capacity and the worst attacking strategy? The estimation of data-hiding capacity for real images based on game between data hider and attacker was performed
 39 by Moulin and Mihcak³⁹ based on the estimation-quantization image model of LoPresto³⁵ and the spike model of Weidmann and Vetterli.⁵⁹ Recently, the same
 41 framework was applied to edge-process stochastic image models proposed by Voloshynovskiy *et al.*⁵⁴ where not so optimistic results were reported concerning
 43 the actual data-hiding capacity for robust watermarking.

1 Most of the robust watermarking schemes are vulnerable to the *copy attack*,³² a
3 protocol attack which consists of the estimation of the watermark from a protected
image and its re-embedding into another media, creating an ambiguity about the
5 hold copyright. It should also be mentioned that additional specific requirements
of robust data-hiding for medical and military applications exist, where the robust
7 watermark should be reversible (i.e. completely removable with the corresponding
authorization). This preserves the quality of the content and allows personalizing
9 the content for different users. The first practical schemes of reversible water-
marking were proposed by Fridrich *et al.*¹⁹ using LSBM-based data embedding.
11 Obviously this sort of embedding had very low robustness. A practical scheme
based on QIM embedding was proposed by Eggers *et al.*¹⁶ Kalker and Willems have
performed the estimation of capacity bounds for invertible robust watermarking.⁴⁸

13 Summarizing the above discussion, we can point out the main requirements
of robust watermarking. Robust watermarking requires the embedding of a 64-bit
15 content independent message (low capacity) into the original image in an invisible
manner specified by a proper distortion criteria. It is also required to have high
17 robustness to all intentional and unintentional attacks and distortions both coming
from signal processing and geometrical transformations. The security requirement
19 calls for a proper resistance against message removal that would be based on the
knowledge of the algorithm.

21 The design of practical algorithms that take into account all these conflicting
requirements is a very challenging task. One of the possible examples of a
23 practical technique based on the SS-based embedding with soft-decoding that meets
these requirements is the *Berkut 1.0* technology developed at the University of
25 Geneva.⁴⁰ *Berkut 1.0* has the best so far reported benchmarking scores according to
Stirmark 3.1 (0.996). The watermark embedding is performed in a critically
27 sampled wavelet domain with appropriate anisotropic perceptual mask that takes
into account texture and luminance masking in each sub band. The robustness
29 against geometrical attacks is ensured based on a special periodical watermark. An
important issue of watermark security of periodical watermark is also resolved by
31 increasing the watermark entropy and thus reducing the watermark predictability.
The watermark is not simply repeated like it is done in the majority of SS-based
33 techniques but it is tiled with some key-dependent predistortions in such a way
that a non-authorized averaging of all tiles leads to the watermark self-destruction
35 and thus watermark prediction is not very accurate. However, we should note that
although this technology is practically very robust to all attacks from the Stirmark
37 benchmark, there is always the possibility that soon new more powerful attacks
might appear in the never ending data-hiding/attacker game.

39 **6. Integrity Control, Tamper Proofing and Watermark-Assisted Communications**

41 The goals of integrity control and tamper proofing consists of the verification
of content integrity, the detection of local modifications in images, video and

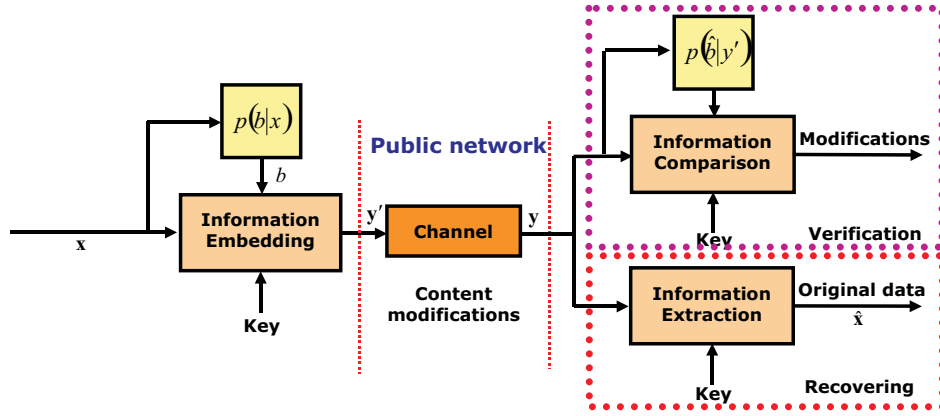


Fig. 6. Generalized diagram of integrity control, tamper proofing and self-recovering systems.

1 documents, and the recovering of the original content based on the available copy
 2 of modified or tampered content. The generalized block-diagram of an integrity
 3 control and verification system is shown in Fig. 6. This protocol is quite similar to
 4 the robust watermarking (Fig. 1) and consists of three main parts. The first part,
 5 information embedding, has the same purpose as robust watermarking, i.e. embed-
 6 ding of the payload \mathbf{b} into the original data \mathbf{x} with the specified distortion that
 7 should not exceed D_1 . However, a fundamental difference exists between these two
 8 applications that consists in the nature of the payload \mathbf{b} . The payload \mathbf{b} has a higher
 9 rate (about 5–10 Kbits depending on the size of the original data). Moreover, the
 10 payload \mathbf{b} is content dependent and related to the original data by some mapping
 11 rule $p(\mathbf{b}|\mathbf{x})$ that might represent some hashing, features or even compressed version
 12 of the original content. Therefore, depending on the final requirements it might be
 13 necessary to provide the embedding rate in the range of $R = 1 - 2$ bpp (bits per
 14 pixels).

15 The second element of the protocol is the public network represented by some
 16 channel with the transition matrix $p(y|y')$. The behavior of this channel also
 17 shows significant differences with the corresponding robust watermarking channel.
 18 Contrary to the robust watermarking channel, where the attacker is interested in
 19 impairing the reliable watermark detection/decoding subject to the constraint of
 20 the maximum allowable distortion D_2 , the protocol attacker in this application
 21 has completely different objectives. The main goal of the attacker is to modify
 22 or to counterfeit the content with the purpose of producing a new content with a
 23 modified visual appearance. For example, the content could be modified by replacing
 24 objects, objects features, human faces, bodies, different authentication attributes,
 25 background, or any identification data targeting some sensational, misleading or
 26 illegal purposes. Obviously, in this case the document is either partially modified or
 27 a fraction of the document is copied into another document. Therefore, the global
 introduced distortion D_2 is of secondary importance for the evaluation of the degree

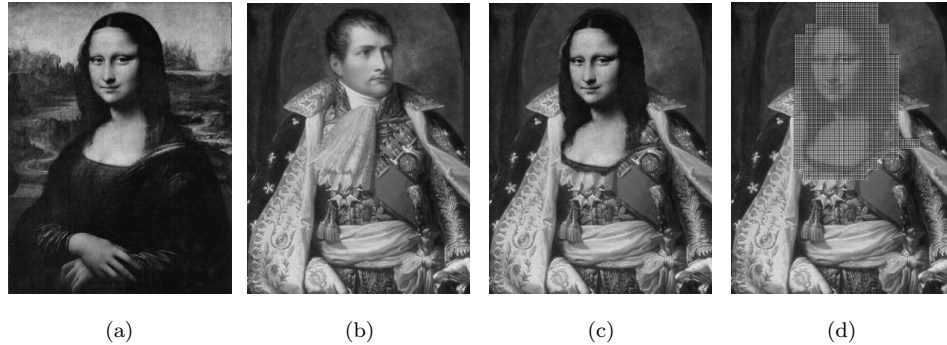


Fig. 7. Example of tamper proofing: (a) and (b) original images, (c) modified content as the result of collage between (a) and (b), (d) highlighted regions indicate the content modifications.

1 of document modification for this application. The following example demonstrates
 2 the possible content modifications of two source images [Figs. 7(a), 7(b)]; such
 3 alterations could change the historical, artistic or even criminalistic conclusions
 4 that could be drawn. Figure 7(c) is the result of a collage between the two previous
 5 images. The goal of the decoder of a tamper proofing system is thus to reliably
 6 detect the intentional or unintentional modifications [Fig. 7(d)], and to point out
 7 the modified areas or preferably reconstruct the original content.

8 Therefore, from the attacker perspective the integrity of the document should
 9 be preserved in such a way that the authentication watermark will not be capable
 10 enough to detect the introduced modifications. This is a very challenging task
 11 that has recently resulted in a lot of attention in the watermarking commu-
 12 nity with respect to the investigation of new protocol attacks against tamper
 13 proofing systems¹¹: such attacks are mostly advanced substitution attacks includ-
 14 ing the *cut-and-paste* attack,³ the *vector-quantization (VQ)* or *Holliman-Memon*
 15 attack,²⁶ image compositions and the *collage attack*, as well as cryptographic attacks
 16 targeting the used hashing function.

17 To withstand the above attacks one should properly design a data-hiding
 18 scheme that should resolve two related problems: the first one is the *detection of*
 19 *modifications*; the second one is the *recovering of the original data \mathbf{x}* after content
 20 modifications. According to the payload and the targeted objectives, the existing
 21 practical systems can be divided in two large groups: *authentication and tamper*
 22 *proofing* watermarks, and *self-recovering* watermarks.

23 6.1. Authentication and tamper proofing

24 While robust watermarking for copyright protection was clearly the first main
 25 research direction of the 1990s, watermarks for authentication and tamper
 26 proofing have then been rapidly proposed for authentication and verification
 27 of integrity. Authentication aims at checking the authenticity of a document
 and of its source, while tamper proofing detects unauthorized modifications.

1 These authentication/tamper proofing watermarks aim at making falsifications
2 and unauthorized modifications easy to detect and characterize. Early authenti-
3 cation watermarks are the Yeung-Mintzer scheme⁶⁴ which authenticates each pixel
4 with respect to a binary logo, and Wong⁶¹ and Coppersmith *et al.*⁷ schemes which
5 divide the image in blocks and attach cryptographic hash-codes or signatures within
6 blocks.

7 However it is known that most of schemes based on block-wise independent
8 hashing are vulnerable to substitution attacks which exploit databases of images
9 all protected with the same key. The cut-and-paste attack takes parts from two or
10 more of these protected images and paste them together (preserving the watermark
11 synchronization) to form a new image. The collage attack is a cut-and-paste attack
12 which uses rather large parts: in that case these parts are individually validated
13 by the decoder and only the boundaries between them are indicated as tampered.
14 Then a collage attack cannot be distinguished with certainty from simple local
15 tampering, and can be seen as a protocol attack which creates the ambiguity
16 about the actual authenticity of the image. The even more powerful VQ attack²⁵
17 allows the construction of completely arbitrary images, usually of very good
18 quality and which are wrongly authenticated by the decoder, by pasting blocks
19 from already watermarked images (the same blocks as those used by the tamper
20 proofing scheme). Moreover, regarding robust watermarking, most existing robust
21 schemes are vulnerable to the copy attack. This is a potential problem for many
22 real-world applications: if the watermark can be copied, how to be sure that the
23 document actually holds the decoded copyright?

24 Various methods of blocks or hash-codes/signatures chaining, undeterministic
25 signatures, etc. have been proposed for authentication watermarks against
26 substitution attacks.³ Fridrich¹⁷ proposed to embed unique identifiers (ID) or
27 “time-stamps” within individual images or even within individual blocks, a method
28 which efficiently and conveniently defeats collage attacks. While several solutions
29 have been proposed to make robust watermarks resistant against the copy attack,
30 one possibility is to include host related data into the watermark. A powerful and
31 elegant solution exists which consists of joining robust and authentication water-
32 marks in a *hybrid* scheme. Therefore the hybrid scheme can resolve problems related
33 to copyright, authenticity and integrity in an integrated framework, and further-
34 more it is also able to defeat both protocol attacks above: the copy attack is made
35 impossible since local signatures mismatch if the watermark is copied from one
36 image to another; and regarding the collage attack, the robust part of the hybrid
37 watermark can help us to identify the areas coming from different sources since
38 they hold different robust watermarks. Hybrid schemes were recently proposed by
39 Fridrich¹⁷ and by Deguillaume *et al.*¹¹ Moreover diagnostics coming from robust
40 and authentication parts can be interpreted together in order to detect copy and
41 collage attacks. This solution is integrated in a prototype *Berkut 2.0*.⁴⁰

42 Most of proposed authentication watermarks are strictly sensitive to any change,
43 in the sense that any modification, even of a single pixel, is detected: they are said

1 *fragile* watermarks. Therefore they are not suitable for compression nor for digital/
 2 analogue conversion. Media-conversion compatible schemes called *semi-fragile*
 3 watermarks have then been proposed, based on robust or *visual hashing* instead
 4 of standard hash-codes as well as on an embedding approach which resists against
 5 a certain level of “acceptable and non-malicious” distortions. The idea of visual
 6 hashing is to generate a key-dependent secure digest which continuously changes
 7 with the input, differing only by a small number of bits for two perceptually
 8 equivalent inputs. On the other hand, completely different inputs or different keys
 9 should generate uncorrelated codes. The verification is then done by comparing
 10 the percentage of mismatching bits with a threshold representing the amount of
 11 allowed distortions. For example visual hashing functions have been proposed using
 12 geometric hashing based on salient points and a voting algorithm,²⁴ or derived from
 13 low-pass DCT coefficients and made invariant to translation, scaling and rotation
 with the Fourier-Mellin transform.¹⁸

15 **6.2. Self-recovering watermarking**

Error resilient coding and error concealment have recently received a lot of attention
 16 in different applications related to wireless networks and networks with no QoS
 17 control. The main existing techniques can be characterized as⁵⁸: layered coding with
 18 transport coding, multiple description coding, joint source/channel coding, robust
 19 waveform coding, robust entropy coding and post-processing at the decoder. Data-
 20 hiding techniques could be used for this purpose. Moreover, intentional content
 21 modifications could be considered as channel degradation and appropriate strategies
 22 can be applied for self-recovering data.

The main practical approaches to self-recovering watermarking differ depending
 23 on the used payload (features, edges, compressed coefficients), data-embedding
 24 technique and final reconstruction/recovering procedure. For instance, an example
 25 of edge directivity embedding is proposed by Yin *et al.*⁶⁵ while downsampled and
 26 lossy compressed (JPEG QF = 25) payload is used in SARI technique.³³

27 Generally, any sort of image features can be embedded into the image itself
 28 that should help reconstruct the resulting image after channel degradations. If the
 29 quadratic distortion is used as a fidelity measure and the channel distortions are
 30 interpreted as AWGN, and if the content is assumed to be stationary Gaussian,
 31 one can apply the MMSE estimator to reconstruct the original image (as post-
 32 processing). The resulting variance of the obtained estimate σ_{MMSE}^2 depends on the
 33 variance of the content σ_x^2 and the variance of the channel noise σ_z^2 , as: $\sigma_{MMSE}^2 =$
 34 $\frac{\sigma_x^2 \sigma_z^2}{\sigma_x^2 + \sigma_z^2}$. Therefore, the higher the variance of the image, the lower is the accuracy
 35 of the estimate. A maximum likelihood (ML) estimate is most often applied to real
 36 images to estimate the local image variance. It is a known fact that real images
 37 are highly non-stationary processes. The variance in the vicinity of the edges and
 38 textures will thus be highly overestimated,⁵⁵ σ_{MMSE}^2 will then be very large in
 39 these regions and no reliable estimate will be possible. Therefore, one can embed
 40
 41

20 *S. Voloshynovskiy et al.*

1 additional information about edges and textures in the watermark to reduce the
 2 variance of the estimation, while flat regions that are characterized by relatively low
 3 variance can easily be reconstructed at the decoder. This approach to error-resilient
 4 coding can be efficiently used for the QoS control and enhanced scalability of public
 5 networks.

6.3. Joint source/channel coding with side information

7 All the above schemes can be considered in a generalized setup of joint
 8 source/channel coding with side information. Since we have already considered
 9 channel coding with side information at the encoder in Sec. 5, we will focus in this
 10 section on source coding with side information at the decoder.

11 We start by considering the problem of lossy compression with side information
 12 (Fig. 8). This problem was first introduced by Wyner-Ziv (1976)⁶³ with \mathbf{X} and \mathbf{Y}
 13 being continuous correlated i.i.d. sources with joint p.d.f. $p_{XY}(x, y)$. The problem
 14 is to compress \mathbf{X} in a lossy way with \mathbf{Y} being known at the decoder but not at
 15 the encoder. First, we consider this problem for a Gaussian content and quadratic
 16 distortion. The source \mathbf{X} produces N samples with $X[i] \sim \mathcal{N}(0, \sigma_x^2)$. The encoder
 17 and the decoder communicate without error at a rate R bits per source symbol.
 18 At the same time, the decoder has access to $\mathbf{Y} = \mathbf{X} + \mathbf{Z}$ (Fig. 9), where \mathbf{X} and \mathbf{Z}
 19 are independent and $Z[i] \sim \mathcal{N}(0, \sigma_z^2)$. The goal is to communicate with the lowest
 20 possible rate R such that $E[d^N(\mathbf{X}, \hat{\mathbf{X}})] \leq D$. This lower bound will be denoted as
 21 $R(D)_{X|Y}^{WZ}$.

22 This communication protocol is based on the fact that if both encoder and
 23 decoder have access to \mathbf{Y} , then they can compute a minimum mean square
 error (MMSE) estimate of \mathbf{X} as $\mathbf{X}_{MMSE} = E[\mathbf{X}|\mathbf{Y}]$. In this case one can only

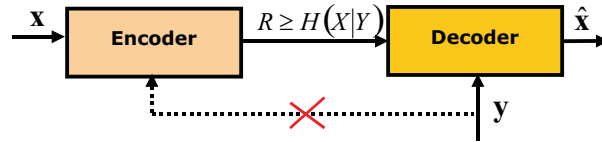


Fig. 8. Generalized diagram of lossy source coding with side information \mathbf{y} .

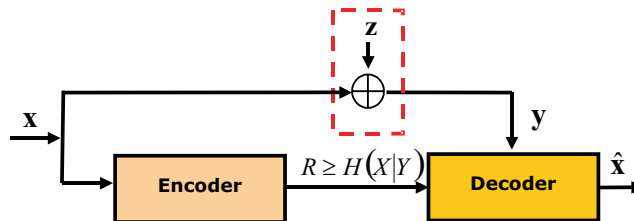


Fig. 9. Generalized diagram of lossy source coding with side information given in the form of a noisy version of $\mathbf{x} : \mathbf{y} = \mathbf{x} + \mathbf{z}$.

1 communicate the error of estimate $\tilde{\mathbf{X}} = \mathbf{X}_{MMSE} - \mathbf{X}$ ($\tilde{X}[i] \sim \mathcal{N}(0, \sigma_{X|Y}^2)$), where
 2 $\sigma_{X|Y}^2 = \frac{\sigma_x^2 \sigma_z^2}{\sigma_x^2 + \sigma_z^2}$, with the specified distortion D , that corresponds to a rate distortion
 3 function $R(D)_{X|Y}$ which assumes \mathbf{Y} to be available at both encoder and decoder:

$$R(D)_{X|Y} = \begin{cases} \frac{1}{2} \log_2 \frac{\sigma_{X|Y}^2}{D}, & \text{if } 0 \leq D < \sigma_{X|Y}^2, \\ 0, & D > \sigma_{X|Y}^2. \end{cases} \quad (21)$$

5 Wyner and Ziv^{62,63} demonstrated that generally $R(D)_{X|Y}^{WZ} \geq R(D)_{X|Y}$. This
 6 means that there is in general a rate loss with Wyner-Ziv coding. Zamir⁶⁶ also
 7 shown that this loss can be as close as $R(D)_{X|Y}^{WZ} - R(D)_{X|Y} \leq \frac{1}{2}$ bit. Note that for
 8 discrete sources when $D = 0$, the Wyner-Ziv problem reduces to the Slepian-Wolf
 9 problem⁴⁵ with $R(0)_{X|Y}^{WZ} = R(0)_{X|Y} = H(X|Y)$. The generalized rate distortion
 10 with side information at the decoder for discrete memoryless sources with distortion
 11 measure $d(\cdot, \cdot)$ was considered in Refs. 9, 62 and 63:

$$R(D)_{X|Y}^{WZ} = \min_{p(u|x), g(\cdot)} [I(X; U) - I(Y; U)]. \quad (22)$$

13 The minimization is performed over all $p(x, y, u) = p(x, y)p(u|x)$ and all decoder
 functions $\hat{x} = g(u, y)$ such that $\sum_{x, u, y} p(x, y)p(u|x)d(x, g(u, y)) \leq D$.

Su *et al.*⁴⁶ shown that for $U^*[i] \sim \mathcal{N}(0, \sigma_u^2)$ where $\sigma_u^2 = (\sigma_x^2 - \frac{\sigma_x^2 + \sigma_z^2}{\sigma_z^2} D) \frac{\sigma_z^2 - D}{\sigma_z^2}$
 one can obtain:

$$I(X; U^*) = \frac{1}{2} \log_2 \frac{(\sigma_z^2 - D)\sigma_x^2}{\sigma_z^2 D}, \quad (23)$$

$$I(Y; U^*) = \frac{1}{2} \log_2 \frac{(\sigma_z^2 - D)(\sigma_x^2 + \sigma_z^2)}{(\sigma_z^2)^2}, \quad (24)$$

15 such that the rate $R^* = I(X; U^*) - I(Y; U^*) = R(D)_{X|Y}$. The design of the code
 16 book \mathcal{U} is quite similar to the Gel'fand-Pinsker code book construction and includes
 17 about $2^{NI(X; U^*) - \varepsilon}$ vectors coming from $\mathcal{N}(\mathbf{0}, \sigma_u^{*2} \mathbf{I})$. The vectors are equiprobably
 18 randomly assigned to $2^{N(R^* + 2\varepsilon)}$ distinct bins with some indexes and each bin
 19 contains about $2^{NI(Y; U^*) - \varepsilon}$ vectors.

20 Having considered the main results of source and channel coding with side
 21 information, we can design a joint source/channel coding (JSCC) system that
 22 generalizes the above approaches. The block diagram of JSCC system with side
 23 information is shown in Fig. 10. For simplicity, assume that the channel is charac-
 24 terized either by AWGN or by some high-rate lossy compression that degrade the
 25 image quality. The allowable channel distortion is D_2 while the degradation due to
 26 the watermark embedding should not exceed D_1 . The ratio D_1/D_2 determines the
 27 maximum data-hiding rate according to the Gel'fand-Pinsker (8), Costa (9), and
 28 (10) channel coding setups. The source coding in the above JSCC scheme is designed
 29 assuming the Wyner-Ziv formulation, that is availability of the degraded stego data
 \mathbf{y} as the side information at the decoder. The watermark (payload) represents the

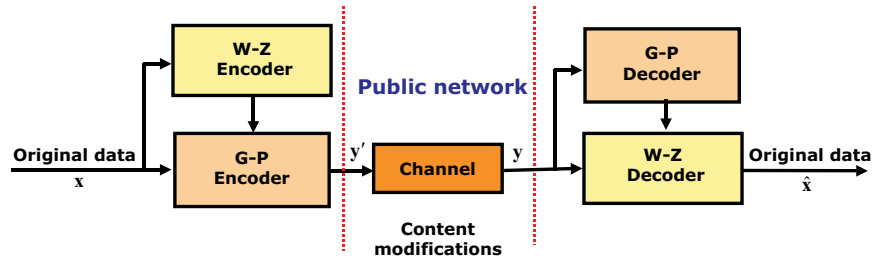


Fig. 10. Generalized block-diagram of joint source/channel coding based on side information.

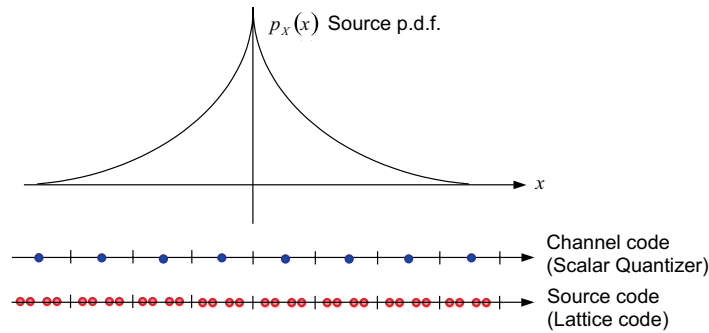


Fig. 11. Joint source/channel coding based on side information.

1 direct compressed channel transmission in the Wyner-Ziv problem communicated
 2 with the rate $R \geq H(X|Y)$. The simplest 1D interpretation of the above scheme
 3 is shown in Fig. 11. The channel code is represented by a uniform scalar quantizer
 4 that at the same time corresponds to a “coarse” approximation of the source data
 5 \mathbf{x} . The source code is designed based on a lattice code with respect to the above
 6 scalar quantizer, assuming amplitude limited attacks or distortions. In particular,
 7 4-PAM is used within each bin of lattice code providing a total embedding rate of
 8 2 bits per pixel. It is also assumed that the watermark is encoded according to the
 9 Wyner-Ziv problem, using multilevel codes and properly allocated over the image.

10 The decoder extracts the watermark based on the Gel’fand-Pinsker decoding.
 11 Then the Wyner-Ziv decoder reconstructs the original data \mathbf{x} using the estimated
 12 payload $\hat{\mathbf{b}}$ and the side information \mathbf{y} . In the noise-free case, the quality of the
 13 reconstructed image will be determined by the specified distortion D of the lossy
 14 Wyner-Ziv coding. In the case of AWGN, the ratio of half of the distance between
 15 lattice constellations ($d_0/2$) in the source code to the noise standard deviation
 16 (σ_z^2) will determine the corresponding probability of decoding error. If $d_0/2 \gg$
 17 σ_z^2 , one can expect error-free communications typical for QIM-like schemes and
 18 corresponding reconstruction with the distortion not exceeding D . Otherwise, errors
 19 might occur and the system fails to reconstruct the original data with the given
 fidelity.

1 The above JSCC scheme can also be used as a self-recovering system for networks
with bit, block or packet losses or erasures. In these case, the payload should be
3 decreased to enable appropriate encoding using error correction codes suited to
the erasure channels, or simply using appropriate data allocation with repetitions.
5 This sort of communications represents a form of error resilient coding that can
assist the problem of QoS control in public networks. It should also be mentioned
7 the possibility of hybrid analog/digital transmission where the digital counterpart
(embedded watermark) can be used to provide additional quality of transmission
9 for authorized users.⁴⁴

11 An extension to the above JSCC is possible in the case of data delivery with
a given targeted data quality to different users, that are divided on public users
(those who do not know the key) and private users (those who have access to
13 the key). Increasing the embedding distortion D_1 , one can considerably degrade
the image/video/audio quality thus making it uninteresting to public users, while
15 private users can still decode the encoded data. Therefore, this scheme can also
be used in applications that require “partial data encryption” to enable the secure
17 content delivery to the target users.

7. Secure Communications

19 The goal of secure communications is to securely deliver some content over the
public networks. There are several possibilities for secure communications. The
21 first one is a visual “encryption” or scrambling that should provide additional error
resilience in the case of wireless networks and networks with packet losses and
23 erasures. The second possibility is steganography that guarantees secure content
delivery by hiding the content to be securely communicated into the covert media,
25 whereas the presence of the hidden content presence should not be detected by
various detection tools.

27 7.1. Visual scrambling

29 The goal of visual scrambling consists in the enciphering of visual content in a way
suitable for reliable communications over public networks; this prevents access of
a third unauthorized party to the enciphered content. The block diagram of visual
31 scrambling is shown in Fig. 12. The content that should be securely communicated
over public networks is scrambled at the encoder based on the private key in such
33 a way that it cannot be anymore visually recognized. Contrary to traditional data
enciphering, it is required here to ensure both visual scrambling and error resilient
35 coding. Moreover, to provide a secure solution it is required to avoid additional
redundancy in headers, meta data and attachments. It is also preferable to provide
37 format independency and to ensure high efficiency towards erasure channels and
channels with varying parameters. Obviously, traditional means cannot completely
39 satisfy these requirements. The network part of communications is concerned in
this application with two different issues, i.e. security and robustness. The security

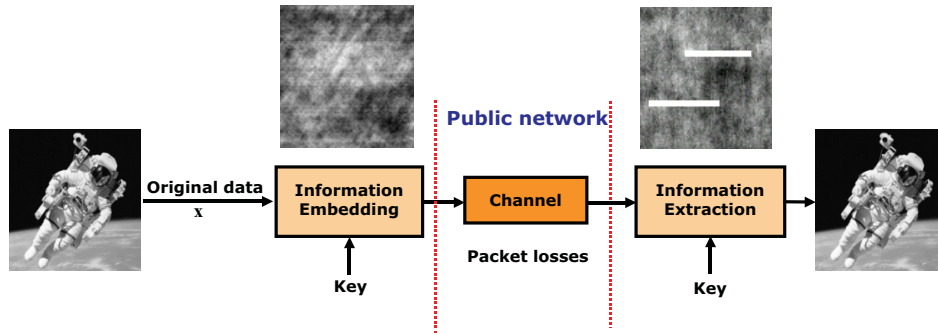


Fig. 12. Generalized diagram of visual scrambling.

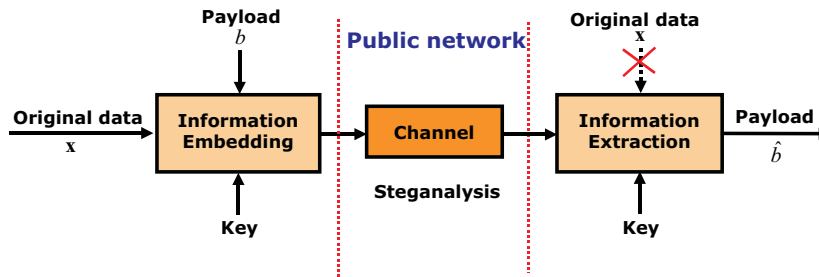


Fig. 13. Generalized diagram of secure communications based on steganography.

1 assumes that unauthorized deciphering can be applied and the robustness issue
 2 refers to different imperfections of networks explained in Sec. 2. Finally, the decoder
 3 should provide reliable descrambling of the content even if some bits, blocks or
 4 packets have been corrupted during transmission. One of possible solutions of
 5 this problem was proposed by Grytskiv *et al.*,²¹ based on phase encryption. This
 6 approach, while being very simple, demonstrates high efficiency for both content
 7 scrambling and error resilience.

7.2. Steganography

9 Steganography, originally designed for cover or hidden communications, should
 10 provide a certain level of security for public communications. The block diagram of
 11 steganographic communications is shown in Fig. 13. The encoding/decoding part
 12 of steganography systems has a lot in common with robust watermarking based on
 13 Costa's scheme. However, it has reduced requirements towards attacks aiming at
 14 the watermark removal and thus it can provide higher embedding rate. It essentially
 15 corresponds to high WNR-regime of data-hiding meaning that normally it should
 16 withstand unintentional attacks such as format conversion, slight lossy compres-
 17 sion and in some special cases analog to digital conversion. While most existing
 steganographic tools can provide perceptually invisible data-hiding, the stochastic

1 visibility or unauthorized detectability of hidden data still remains a challenging
 2 task. Therefore, to be secure, the steganographic system should satisfy a set of
 3 requirements. The main requirement consists in providing the statistical indistin-
 4 guishability between the cover data and the host data. A possible information-
 5 theoretic measure of stochastic closeness is the relative entropy or Kullback-Leibler
 6 distance (KLD) between two distributions under test, which was first proposed by
 7 Cachin.⁵ More generally, the stochastic visibility can be considered as the possibility
 8 of unauthorized detection to differentiate between the cover and the host data based
 9 on a hypothesis testing.

10 The basic scheme for steganography communications requires high-rate commu-
 11 nications. Thus, the host interference cancellation issue should be resolved. The first
 12 steganographic techniques have been mostly built based on the LSBM embedding.
 13 The QIM and SCS based embedding for steganographic purposes, proposed by
 14 Eggers *et al.*¹⁵ and Guillon *et al.*,²² have proven that the SCS-based steganography
 15 is secure according to the Cachin's criteria of ϵ -security⁵ which requires:

$$D(p_X \| p_Y) < \epsilon, \quad (25)$$

16 where $D(p_X \| p_Y) = \sum_{x \in \mathcal{X}} p_X(x) \log \frac{p_X(x)}{p_Y(x)}$ denotes relative entropy (or Kullback-
 17 Leibler distance) between the cover data X and the stego data Y .

18 However, the relative entropy is a global criteria and does not reflect the local
 19 content modifications. This means that the content can be modified locally in such
 20 a way that the attacker can detect it either visually or using some specially de-
 21 signed statistical tests, while the relative entropy can be tuned to be very low. This
 22 also valid for the estimation of image quality using the MSE criterion. It was also
 23 shown that images can be considerably locally distorted while the MSE indicated
 24 acceptable image quality.⁵³

25 At the same time, it is obvious that the higher the data embedding rate, the more
 26 modifications are introduced into the original content and consequently the easier
 27 the task of steganalysis. The trade-off between data-hiding capacity and security of
 28 data-hiding technologies expressed as the possibility of unauthorized detection was
 29 performed by Voloshynovskiy and Pun.⁵⁶ Different data-hiding techniques have
 30 been considered for both low-WNR and high-WNR regimes and corresponding
 31 statistical detection strategies have been proposed. It was emphasized that the
 32 crucial role in the unauthorized detection of hidden data belongs to the proper
 33 stochastic modeling of the cover content. The commonly used EQ model pro-
 34 vides overestimated values of local image variances indicating "high steganographic
 35 security of specific image regions". Therefore, the corresponding conclusion is that
 36 one should embed the payload in edges and textures due to their high variance and
 37 corresponding high probability of error of unauthorized detection. Recent results
 38 obtained for capacity of robust data-hiding techniques question the validity of these
 39 assumptions.⁵⁴ It is also likely that the appearance of new more powerful and
 40 accurate stochastic image models can change this belief and inspire new secure
 41 data-hiding strategies.

1 8. Conclusion

3 The goal of this paper was to demonstrate how network security, QoS control
5 and secure communications over the public networks can benefit from data-hiding
7 technologies. It was shown that such technologies can play an important assisting
9 role in public networks characterized to be heterogeneous, time-varying, and in
11 networks with no QoS control. We have also indicated that traditional means of
13 multimedia security can hardly cope with novel emerging requirements of multi-
15 media communications. The data-hiding technologies, being a possible alternative,
17 do not require considerable investment, protocol modifications and are compatible
19 with the existing standards of multimedia compression and communications. We
21 have also developed a unified theoretical basis of digital data-hiding as well as shown
23 its major applications. In particular, digital communication with side information
was demonstrated to be an appropriate theoretical basis for considering different
aspects of channel coding and source coding in data-hiding applications. The recent
interest towards the Gel'fand-Pinsker and Wyner-Ziv problems makes it possible
to consider digital data-hiding in the scope of multi-terminal and multi-user net-
works and communications. This makes the analysis even more attractive with a
huge future potential of extensions towards multimodal data storage, communica-
tions and management in highly distributed environments. State-of-the-art robust
watermarking, tamper proofing, watermark-assisted multimedia processing and
secure communications are considered among others based on a unified theoretical
basis. The main requirements, design principles, generalizations, as well as future
perspectives are underlined in the paper.

Acknowledgment

25 This paper was partially supported by SNF grant No 21-064837.01, SNF Professor-
27 ship grant No PP002-68653/1, CTI-CRYMEDA-SA and IM2 projects. The authors
are thankful to Yuriy Rytsar for many helpful and interesting discussions.

References

- 29 1. J. Allen, A. Christie, W. Fithen, J. McHugh, J. Pickel and E. Stoner, "State of the
31 practice of intrusion detection technologies," Technical Report CMU/SEI-99TR-028,
Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA, USA
(2000).
- 33 2. M. Barni, F. Bartolini and T. Furon, "A general framework for robust watermarking
35 security," *Signal Processing, Special Issue on Security of Data Hiding Technologies*
(2003).
- 37 3. P. S. L. M. Barreto, H. Y. Kim and V. Rijmen, "Toward a secure public-key blockwise
fragile authentication watermarking," In *IEEE ICIP2001*, pp. 494–497, Thessaloniki,
39 Greece (October 2001).
- 41 4. D. Bertsekas and R. Gallager, *Data Networks*, Englewood Cliffs, NJ, Prentice-Hall
(1987).
5. C. Cachin, "An information-theoretic model for steganography," In *Information
Hiding: Second International Workshop IHW'98*, Portland, Oregon, USA (April 1998).

- 1 6. B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably
3 good methods for digital watermarking and information embedding," *IEEE Trans.
on Information Theory*. **47**, 1423–1443 (May 2001).
- 5 7. D. Coppersmith, F. Mintzer, C. Tresser, C. W. Wu and M. M. Yeung, "Fragile
imperceptible digital watermark with privacy control," In *IS&T/SPIE Electronic
7 Imaging99, Session: Security and Watermarking of Multimedia Contents*, San Jose,
CA, USA (January 1999).
- 9 8. M. Costa, "Writing on dirty paper," *IEEE Trans. on Information Theory* **29**(3),
439–441 (May 1983).
- 11 9. T. Cover and J. Thomas, *Elements of Information Theory*, Wiley and Sons, New York
(1991).
- 13 10. F. Deguillaume, S. Voloshynovskiy and T. Pun, "Method for the estimation and
recovering of general affine transforms in digital watermarking applications," In
15 *IS&T/SPIE's 14th Annual Symposium, Electronic Imaging 2002: Security and
Watermarking of Multimedia Content IV*, Vol. 4675, pp. 313–322, San-Jose, CA, USA
(January 20–25, 2002).
- 17 11. F. Deguillaume, S. Voloshynovskiy and T. Pun, "Secure hybrid robust watermarking
resistant against tampering and copy attack," *Signal Processing* **83**(10), 2133–2170
19 (2003).
- 21 12. J. F. Delaigle, C. De Vleeschouwer and B. Macq, Watermarking algorithm based on
a human visual model," *Signal Processing* **66**, 319–335 (1998).
- 23 13. R. J. Edell, N. McKeown and P. P. Varaiya, "Billing users and pricing for TCP,"
IEEE J. Select. Areas Communications **13**, 1162–1175 (September 1995).
- 25 14. J. Eggers, J. Su and B. Girod, "A blind watermarking scheme based on structured
codebooks," In *Secure Images and Image Authentication, IEE Colloquium*, pp. 4/1–
4/6, London, UK (April 2000).
- 27 15. J. J. Eggers, R. Buml and B. Girod, "A communications approach to image steganog-
raphy, In *Proceedings of SPIE: Electronic Imaging 2002, Security and Watermarking
29 of Multimedia Contents IV*, Vol. 4675, pp. 26–37, San Jose, CA, USA (January 2002).
- 31 16. J. J. Eggers, R. Buml, R. Tzschoppe and B. Girod, "Inverse mapping of
SCS-watermarked data," In *Eleventh European Signal Processing Conference
(EUSIPCO'2002)*, Toulouse, France (September 3–6, 2002).
- 33 17. J. Fridrich, "A hybrid watermark for tamper detection in digital images," In *ISSPA '99
Conference*, Brisbane, Australia (August 1999).
- 35 18. J. Fridrich, "Visual hash for oblivious watermarking," In *IS&T/SPIE Proceedings
3971*, San Jose, California, USA (January 2000).
- 37 19. J. Fridrich and M. Goljan, "Images with self-correcting capabilities," In *IEEE Inter-
national Conference on Image Processing Proceedings ICIP'99*, pp. 792–796, Kobe,
39 Japan (October 25–28, 1999).
- 41 20. S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters,"
Problems of Control and Information Theory **9**(1), 19–31 (1980).
- 43 21. Z. Grytskiv, S. Voloshynovskiy and Y. Rytsar, "Cryptography and steganography
of video information in modern communications," In *Proc. 3rd International Con-
ference on Telecommunications in Modern Satellite, Cable and Broadcasting Services
45 TELSIKS'97 1*, 164–167, Nis, Yugoslavia (October 1997).
- 47 22. P. Guillon, T. Furon and P. Duhamel, "Applied public-key steganography," In *Pro-
ceedings of SPIE: Electronic Imaging 2002, Security and Watermarking of Multimedia
Contents IV 4675*, San Jose, CA, USA (January 2002).
- 49 23. S. Hansell, "E-Msic sites settle on prices, it's a start," *NY Times (Business Day)*
(March 3, 2003).

28 *S. Voloshynovskiy et al.*

- 1 24. H. Hel-Or, Y. Yitzhaki and Y. Hel-Or, "Geometric hashing techniques for water-
marking," In *ICIP 2001* (2001).
- 3 25. M. Holliman and N. Memon, "Counterfeiting attacks on linear watermarking
systems," In *Proc. IEEE Multimedia Systems 98, Workshop on Security Issues in*
5 *Multimedia Systems*, Austin, Texas (June 1998).
- 7 26. M. Holliman and N. Memon, "Counterfeiting attacks on oblivious block-wise inde-
pendant invisible watermarking schemes," In *IEEE Trans. on Image Processing* **9**,
pp. 432–441 (March 2000).
- 9 27. M. S. Kankanhalli and R. K. R. Ramakrishnan, "Content-based watermarking of
images," In *Multimedia and Security Workshop at ACM Multimedia'98, Bristol, UK*
11 (September 1998).
- 13 28. R. A. Kemmerer, *Secure Computing on the Internet* (1998).
- 15 29. J. N. Keon, *Pricing in Telecommunications Networks Offering Multiple Services and*
Quality of Service Guarantees, PhD thesis, Systems Eng. Dept., Univ. of Pennsylvania,
Philadelphia, PA, USA (August 2000).
- 17 30. D. Kundur and D. Hatzinakos, "Improved robust watermarking through attack char-
acterization," In *Optics Express* **3**, 485–490 (December 1998).
- 19 31. M. Kutter, "Watermarking resistant to translation, rotation and scaling," In *Proc.*
SPIE Int. Symp. on Voice, Video, and Data Communication **3528**, 423–431, Boston,
USA (November 1998).
- 21 32. M. Kutter, S. Voloshynovskiy and A. Herrigel, "Watermark copy attack," In
IS&T/SPIE's 12th Annual Symposium, Electronic Imaging 2000: Security and
23 *Watermarking of Multimedia Content II* **3971**, San Jose, California USA (23–28 Jan
2000).
- 25 33. C. Lin, D. Sow and S. Chang, "Using self-authentication-and-recovery images for error
concealment in wireless environments," In *SPIE ITCOM/OptiComm* **4518**, Denver,
27 CO, USA (August 2001).
- 29 34. E. T. Lin, C. I. Podilchuk, T. Kalker and E. J. Delp, "Streaming video and rate
scalable compression: What are the challenges for watermarking?" In *Proceedings*
of the SPIE International Conference on Security and Watermarking of Multimedia
31 *Contents III* **4314**, San Jose, CA, USA (January 2002).
- 33 35. S. LoPresto, K. Ramchandran and M. Orhard, "Image coding based on mixture
modeling of wavelet coefficients and a fast estimation-quantization framework," In
Data Compression Conference 97, pp. 221–230, Snowbird, Utah, USA (1997).
- 35 36. S. H. Low, "Equilibrium bandwidth and buffer allocation for elastic traffics,"
IEEE/ACM Trans. Networking **8**, 373–383 (June 2000).
- 37 37. J.-F. Delaigle, M. Bertran and B. Macq, "Some improvements to HVS models for
fingerprinting in perceptual decompressors," In *IEEE Int. Conf. on Image Processing*
39 *ICIP2001*, 1039–1042, Thessaloniki, Greece (October 2001).
- 41 38. J. K. Mackie-Mason and H. Varian, "Pricing the Internet," In B. Kahin and MA,
MIT Press, J. Keller (Eds). Cambridge, *Public Access to the Internet*, pp. 269–314,
San Jose, CA, USA (January 1995).
- 43 39. P. Moulin and M. K. Mihcak, "A framework for evaluating the data-hiding capacity
of image sources," *IEEE Trans. on Image Processing* **11**(9), 1029–1042 (September
45 2002).
- 47 40. University of Geneva Stochastic Image Processing (SIP) Group. SIP Watermarking
Technology. http://watermark.unige.ch/wmg_technology.html.
- 49 41. F. Pérez-González, F. Balado and J. R. Hernández, "Performance analysis of existing
and new methods for data hiding with known-host information in additive channels,"
IEEE Trans. on Signal Processing, Special Issue on Signal Processing for Data Hiding
51 *in Digital Media and Secure Content Delivery* **51**(4), (April 2003).

- 1 42. F. Pérez-González, J. R. Hernández and F. Balado, "Approaching the capacity limit
3 in image watermarking: A perspective on coding techniques for data hiding appli-
cations," *Signal Processing, Special Issue on Information Theoretic Issues in Digital
Watermarking* **81**(6), 1215–1238 (2001).
- 5 43. C. I. Podilchuk and W. Zeng, "Image-adaptive watermarking using visual models,"
IEEE Journal on Selected Areas in Communications **16**(4), 525–539 (May 1998).
- 7 44. R. Puri, K. Ramchandran and S. S. Pradhan, "On seamless digital upgrade of
9 analog transmission systems using coding with side information," In *Allerton Conf.
Communication, Control, and Computing*, Allerton, IL, USA (October 2002).
- 11 45. D. Slepian and J. K. Wolf, "Noiseless encoding of correlated information sourcea,"
IEEE Trans. Information Theory **19**, 471–480 (July 1973).
- 13 46. J. K. Su, J. J. Eggers and B. Girod, "Channel coding and rate distortion with side
information: Geometric interpretation and illustration of duality," *Submitted to IEEE
Trans. on Information Theory* (May 2000).
- 15 47. M. D. Swanson, B. Zhu and A. H. Tewfik, "Data hiding for video-in-video," In *IEEE
Int. Conf. on Image Processing ICIP1997* **2**, 676–679, Piscataway, USA (October
17 1997).
- 19 48. F. M. Willems and T. Kalker, "Capacity bounds and code constructions for reversible
data-hiding," In *IS&T/SPIE Proceedings, Security and Watermarking of Multimedia
Contents V* **5020**, Santa Clara, California, USA (January 2003).
- 21 49. S. Voloshynovskiy, F. Deguillaume, S. Pereira and Thierry Pun, "Optimal diversity
watermarking with channel state estimation," In *IS&T/SPIE's Annual Symposium,
23 Electronic Imaging 2001: Security and Watermarking of Multimedia Content III*
4134, 23–27, San Jose, California USA (21–26 January 2001).
- 25 50. S. Voloshynovskiy, F. Deguillaume and T. Pun, "Multibit digital watermarking
robust against local nonlinear geometrical distortions," In *IEEE Int. Conf. on Image
27 Processing ICIP2001*, pp. 999–1002, Thessaloniki, Greece (October 2001).
- 29 51. S. Voloshynovskiy, A. Herrigel, N. Baumgaertner and T. Pun, "A stochastic approach
to content adaptive digital image watermarking," In *Third International Workshop
on Information Hiding*, pp. 212–236 (September 29–October 1st 1999).
- 31 52. S. Voloshynovskiy, A. Herrigel, Y. Rytsar and T. Pun, "Stegowall: Blind statistical
33 detection of hidden data," In E. J. Delp and P. W. Wong (eds.), *Proceedings of SPIE
Photonics West, Electronic Imaging 2002, Security and Watermarking of Multimedia
Contents IV*, San Jose, CA, USA (January 2002).
- 35 53. S. Voloshynovskiy, S. Pereira V. Iquise and T. Pun, "Towards a second genera-
tion benchmark," *Signal Processing, Special Issue on Information Theoretic Issues
37 in Digital Watermarking* **81**, 1177–1214 (June 2001).
- 39 54. S. Voloshynovskiy, O. Koval, F. Deguillaume and T. Pun, "Data hiding capacity-
security analysis for real images based on stochastic non-stationary geometrical
41 models," In *IS&T/SPIE's Annual Symposium, Electronic Imaging 2003: Image
and Video Communications and Processing V*, Santa Clara, California USA (20–24
43 January 2003).
- 45 55. S. Voloshynovskiy, O. Koval and T. Pun, "Wavelet-based image denoising using non-
stationary stochastic geometrical image priors," In *IS&T/SPIE's Annual Symposium,
47 Electronic Imaging 2003: Image and Video Communications and Processing V*, Santa
Clara, California USA (20–24 January 2003).
- 49 56. S. Voloshynovskiy and T. Pun, "Capacity-security analysis of data hiding tech-
nologies," In *IEEE International Conference on Multimedia and Expo ICME2002*,
Lausanne, Switzerland (August 26–29, 2002).
- 51 57. J. Walrand and P. Varaiya, *High-Performance Communication Networks*, San
Francisco, CA: Morgan Kaufmann (1996).

30 *S. Voloshynovskiy et al.*

- 1 58. Y. Wang and Q.-F. Zhum, “Error control and concealment for video communications:
A review,” *Proc. IEEE* **86**(5), (1998).
- 3 59. C. Weidmann and M. Vetterli, “Rate-distortion analysis of spike processes,” In *Data
Compression Conference*, Snowbird, USA (March 1999).
- 5 60. M. Welzl and M. Mühlhäuser, “Scalability and quality of service: A trade-off,” *IEEE
Communications Magazine — Scalability in IP-Oriented Networks* **41**(6).
- 7 61. P. W. Wong, “A public key watermark for image verification and authentication,”
In *IEEE International Conference on Image Processing '98 (ICIP'98) Proceedings* **1**
9 (1998).
62. A. Wyner, The rate-distortion function for source coding with side information at
the decoder-ii: General sources,” *Information and Control* **38**, 60–80 (1978).
- 11 63. A. Wyner and J. Ziv, “The rate-distortion function for source coding with side in-
formation at the decoder,” *IEEE Trans. Information Theory* **22**(1), 1–10 (January
13 1976).
- 15 64. M. M. Yeung and F. C. Mintzer, “An invisible watermarking technique for image
verification,” In *1997 International Conference on Image Processing (ICIP '97)* **2**,
17 680–683, Washington, DC, USA (October 26–29, 1997).
- 19 65. P. Yin, B. Liu and H. Yu, “Error concealment using data hiding,” In *IEEE Int. Conf.
On ASSP*, Salt Lake City, USA (May 2001).
- 21 66. R. Zamir, “The rate loss in the wyner-ziv problem,” *IEEE Trans. Information Theory*
19, 2073–2084 (November 1996).



Sviatoslav Voloshynovskiy received the Radio Engineer Degree from the Lviv Polytechnic Institute in 1993 and the PhD degree in Electrical Engineering from State University “Lvivska Polytechnika”, Lviv, Ukraine in 1996.

In 1998–1999 he has been with the University of Illinois at Urbana-Champaign, USA, as a visiting scholar. Since 1999, he has been with University of Geneva, Switzerland, where he is currently an Associate Professor with the Department of Computer Science.

His current research interests are in information-theoretic aspects of digital data-hiding, visual communications with side information and stochastic image modeling for denoising, compression and restoration. He has served as a consultant to private industries in the above areas.



Frédéric Deguillaume graduated in 1997 from the “Computer Vision and Multimedia Laboratory”, University of Geneva, Switzerland.

He has been working since 1998 in this laboratory and obtained his PhD in October 2002 in the field of robust and authentication digital watermarking. He participated in the EU ACTS project DVP in 1996–1997, and in the EU ESPRIT OMI project JEDI-FIRE in 1998–2000.

23

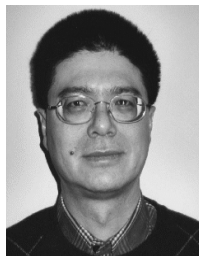
1 He is now collaborating as a post-doc in the Swiss CTI project CRYMEDA-
3 SA since January 2003, which addresses the security of authentication documents
based on high-capacity data-hiding, barcodes, and luminescent crystal-based analog
watermarking.



Oleksiy Koval received his Master degree in Electrical Engineering from the National University “Lvivska Politechnika” (Lviv, Ukraine) in 1996.

In 1996–2001 he was with the Department of Synthesis, Processing and Identification of Images, Institute of Physics and Mechanics (Lviv, Ukraine) as a research engineer.

5 Since 2002 he has been in affiliation with the Stochastic Image Processing Group, Computer Vision and Multimedia Laboratory, University of Geneva where he is currently pursuing his PhD in Stochastic image modeling. His research interests are: information theory, image processing with side information.



Thierry Pun received his PhD in image processing in 1982, from the Swiss Federal Institute of Technology in Lausanne (EPFL).

He joined the University of Geneva, Switzerland in 1986, where he is currently a full professor at the Computer Science Department and the head of the Computer Vision and Multimedia Laboratory. Since 1979 he has been active in various domains of image processing, image analysis and computer vision. He has authored or co-authored over 200 journal and conference papers in these areas as well as six patents, and led or participated to a number of national and European research projects.

His current research interests, related to the design of multimedia information systems and multimodal interaction, focus on: image and video content-based information retrieval systems, brain-computer interaction, multimodal computer interfaces for blind users, data-hiding, image and video watermarking.