Proceedings of SPIE Vol. 4675
Security and Watermarking
of Multimedia Contents IV
San Jose, Ca., Jan. 2002

A Communications Approach to Image Steganography

Joachim J. Eggers and R. Bäuml
Telecommunications Laboratory
University of Erlangen-Nuremberg
Cauerstrasse 7/NT, 91058 Erlangen, Germany

Bernd Girod
Information Systems Laboratory
Stanford University
Stanford, CA 94305-9510, USA

ABSTRACT

Steganography is the art of communicating a message by embedding it into multimedia data. It is desired to maximize the amount of hidden information (embedding rate) while preserving security against detection by unauthorized parties. An appropriate information-theoretic model for steganography has been proposed by Cachin. A steganographic system is perfectly secure when the statistics of the cover data and the stego data are identical, which means that the relative entropy between the cover data and the stego data is zero. For image data, another constraint is that the stego data must look like a "typical image." A tractable objective measure for this property is the (weighted) mean squared error between the cover image and the stego image (embedding distortion). Two different schemes are investigated. The first one is derived from a blind watermarking scheme. The second scheme is designed specifically for steganography such that perfect security is achieved, which means that the relative entropy between cover data and stego data tends to zero. In this case, a noiseless communication channel is assumed. Both schemes store the stego image in the popular JPEG format. The performance of the schemes is compared with respect to security, embedding distortion and embedding rate.

Keywords: steganography, scalar Costa scheme, histogram mapping

1. INTRODUCTION

Steganography is the art of communicating a message by embedding it into multimedia data (cover data), where the very existence of the embedded message should not be detectable by unauthorized parties. We consider steganography as secure information hiding in the presence of a "passive" adversary (warden) Fig. 1. A good illustration of this scenario is given by Simmons' "Prisoners' Problem".[1,2] Alice sends a message u over a channel controlled by Eve to the recipient Bob. Eve allows only restricted communication between Alice and Bob, that means some certain type of information should not be communicated. Therefore, Alice hides the message u within some cover data $\mathbf{x}$ that Eve usually allows to be transmitted. After information embedding, the cover data $\mathbf{x}$ is denoted as steganographic data $\mathbf{r}$. Note that the term "passive warden" means that the steganographic data $\mathbf{r}$ is not modified by Eve. Nevertheless, Eve can be active in the sense that communication between Alice and Bob is interrupted completely. In a typical scenario, Eve may belong to a company which tries to keep some information secret and Alice, being in this company, tries to transmit this secret to Bob, who is outside. Alice tries to maximize the hidden information (embedding rate) while preserving security against detection of the hidden information by Eve. Alice and Bob share a secret key K which is used for embedding the message u and for reception of the message $\hat{u}$. Ideally, we expect $\hat{u} = u$, however, as discussed below, some embedding schemes may not allow for entirely error-free reception, leading to $\hat{u} \neq u$ in general. In this paper, data sequences are represented by vectors, e.g., $\mathbf{x}$ for the cover data, with x_n being its nth element. Random variables are written in Sans Serif font, e.g., x for a scalar random variable and $\mathbf{\mathsf{x}}$ for a vector random variable.

The requirements of steganography and appropriate quality criteria are reviewed in Sec. 2. Then, two information hiding systems are investigated concerning their properties in case of a passive adversary. In both cases, only Alice knows the cover image. It has been shown[3–8] that Alice can exploit her side-information about the cover image to achieve high embedding rates for small embedding distortions. We investigate in Sec. 3 the usage of the watermarking technology ST-SCS ("*spread-transform scalar Costa scheme*") followed by JPEG compression, where JPEG compression is considered as an unavoidable "attack" on the steganographic data. The second system is explicitly designed to hide information in a JPEG compressed image format so that error-free communication becomes possible. The necessary data modifications for the information hiding are such that the probability mass functions (PMFs) of the quantized DCT coefficients are (almost) identical to those of the JPEG compressed image data without hidden information. In Sec. 5, the application of both systems to image steganography is analyzed with respect to the embedding distortion, embedding rate, and relative entropy between the cover image and the steganographic image.

Further author information: Send correspondence to J. Eggers. Email: eggers@LNT.de

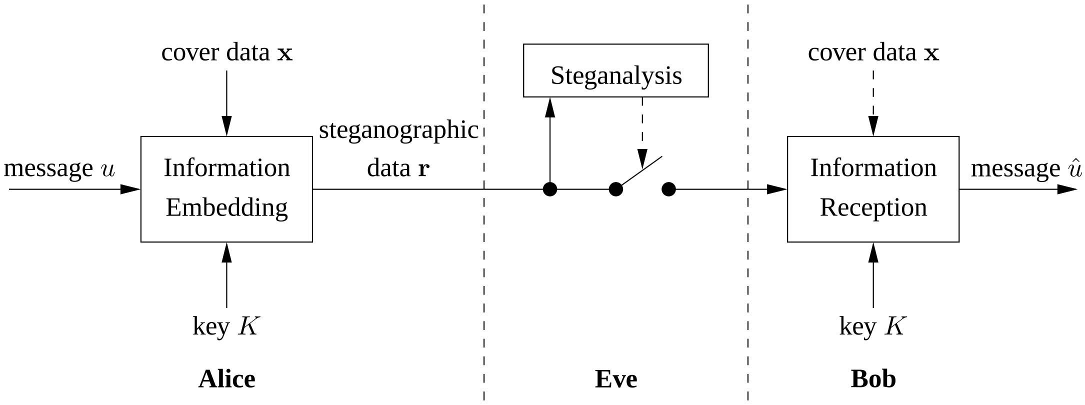

Figure 1. Dataflow for information hiding with a "passive" adversary (steganography).

2. DESIGN RULES FOR IMAGE STEGANOGRAPHY

The purpose of steganography is mainly to enable communication over the channel shown in Fig. 1. We have to find appropriate models of steganography in order to evaluate how close a specific steganographic scheme comes to this goal. However, one has to be careful with the interpretation of results obtained for such a model since steganalysis might exploit model inaccuracies. Thus, we summarize in this section the most important design rules for image steganography and emphasize the limitations of technical security measures.

2.1. General Considerations

Eve controls the communication between Alice and Bob and is willing to interrupt certain types of communication. Ideally, Eve would inspect each message and decide whether communication is allowed or not. Thus, encrypted data is not allowed since Eve cannot decipher the content. It is assumed that all "plain" data is examined by Eve, although this might become difficult if the "innocent" traffic between Alice and Bob is large. Thus, Alice is left with the attempt to hide "unallowed" messages within commonly accepted data that is also called cover data. One attractive type of cover data is natural image data, since

- images contain a significant amount of data, hopefully enabling high secret communication rates,
- natural image data can be modified slightly without leading to visible artifacts,
- images are in many scenarios "innocent" data types to Eve, e.g., Alice might be allowed to send some pictures from products of her company to Bob.

Due to these properties, image steganography has been investigated quite often within the last years. However, one should also consider the fact that a really paranoid Eve can flip the last argument and no longer allows the communication of image data between Alice and Bob if Eve learns that image steganography works. However, since this paper should be longer than two pages, we simply ignore this argumentation and consider natural image data as "innocent" even if image steganography works. Thus, Alice is left with embedding her message such that the steganographic image $\mathbf{r}$ does not look suspicious to Eve.

Eve can analyse the steganographic image $\mathbf{r}$ with respect to

- the size measured in bits per pixel,
- the subjective quality, and
- statistical properties.

The first item leads to the conclusion that uncompressed image data looks to Eve as suspicious as encrypted data. Thus, the steganographic image $\mathbf{r}$ has to be always in a compressed format. In this paper, we consider the popular JPEG format which involves lossy compression without degrading the image beyond acceptable quality. The subjective quality of $\mathbf{r}$ is difficult to analyze. Even more difficult is the question whether a certain image quality is natural or not. To get a hand at this problem, we measure the embedding distortion D_{Emb} by the mean squared error (MSE) between the image elements (pixels, transform coef£cients) of $\mathbf{x}$ and $\mathbf{r}$. In practice, Eve cannot evaluate D_{Emb}, however, small D_{Emb} ensures that the introduced modi£cations are visually imperceivable. It is not really required that no perceivable differences between $\mathbf{x}$ and $\mathbf{r}$ exist. However, small D_{Emb} ensures that $\mathbf{r}$ is close to common cover data and thus provides a tracktable quality measurement. The most dif£cult problem is to measure security against statistical steganalysis which is discussed in the following subsection.

2.2. Statistical Steganalysis

Different methods for statistical steganalysis have been proposed by several reasearchers and are often exploited to design improved steganographic schemes.[9–13] Common to all these approaches is the design of statistical tests that are used to distinguish original cover image data from steganographic image data which is basically a hypothesis testing problem. Cachin[2] proposed an appropriate information-theoretic model that allows to quantify the security of steganography in terms of the decision error probabilities of hypothesis testing. As already mentioned by Caching, such a formal security notion has to be interpreted with care since the adversary might exploit information that is not included into a certain model for steganography. The fundamental problem is that, in the long run, the information hider and the adversary can improve on their models. To obtain a fair analysis, we assume that both parties exploit the same statistical features of the cover data. Based on this assumption, it is reasonable to adopt Cachin's security measure for steganography which is basically the relative entropy[14] (Kullback Leibler distance) between the cover data and the steganographic data.

The relative entropy measures the "distance" between two probablity mass functions (PMFs) $p_{\mathsf{x}}[x]$ and $p_{\mathsf{r}}[x]$ of two discrete random variables x and r, both with support set $\mathcal{X}$. The formal de£nition is

$$D\left(p_{\mathsf{x}}[x]\,||\,p_{\mathsf{r}}[x]\right) = \sum_{x\in\mathcal{X}} p_{\mathsf{x}}[x] \log \frac{p_{\mathsf{x}}[x]}{p_{\mathsf{r}}[x]}, \tag{1}$$

with the convention that $0 \log \frac{0}{p_{\mathsf{r}}[x]} = 0$ and $p_{\mathsf{x}}[x] \log \frac{p_{\mathsf{x}}[x]}{0} = \infty$ (We set $p_{\mathsf{x}}[x] \log \frac{p_{\mathsf{x}}[x]}{0} = 0$ for measured $p_{\mathsf{r}}[x]$ due to the large estimation variance for small values of $p_{\mathsf{r}}[x]$). $D\left(p_{\mathsf{x}}[x]\,||\,p_{\mathsf{r}}[x]\right)$ is always non-negative and is zero iff $p_{\mathsf{x}}[x] = p_{\mathsf{r}}[x]$. Cachin de£nes that a steganographic system with cover data modeled by the random variable x and with steganographic data modeled by the random variable r is *ϵ-secure* if $D\left(p_{\mathsf{x}}[x]\,||\,p_{\mathsf{r}}[x]\right) \le \epsilon$ If $\epsilon = 0$, the steganographic system is called *perfectly secure*.

Note that the relative entropy provides only a security measure against statistical steganalysis. The embedding distortion D_{Emb} can be quite large, even if a steganographic system is perfectly secure according to Cachin's de£nition. Since D_{Emb} should be small as well, it is not possible to achieve a secure steganographic system by replacing the cover data $\mathbf{x}$ with any random data $\mathbf{r}$ having the same statistics. The relation of the embedding distortion and the security of a steganographic scheme is not included in Cachin's work, however, it is crucial for image steganography. Thus, we investigate in this work D_{Emb} and $D\left(p_{\mathsf{x}}[x]\,||\,p_{\mathsf{r}}[x]\right)$ to evaluate the security of a speci£c steganographic system.

3. STEGANOGRAPHY BASED ON (ST-)SCS WATERMARKING

Recently, information embedding has been investigated in particular in the context of digital watermarking. For digital watermarking, information embedding techniques have to be designed such that subsequent processing does not destroy the embedded information. This property makes digital watermarking technology also attractive for steganography when information embedding is followed by lossy compression. Fig. 2 shows a block diagram of watermark communication where lossy compression is modeled by a quantization attack.

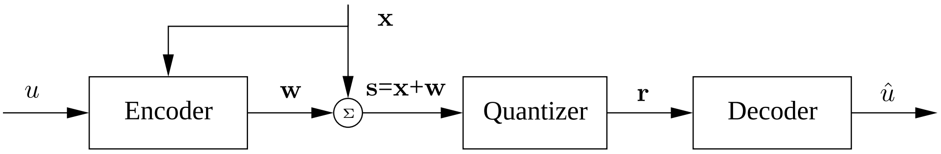

Figure 2. Information embedding with side information at the encoder. The communication channel is composed by simple quantization which models lossy compression of the data $\mathbf{s}$.

Chen and Wornell[15] and Cox, Miller and McKellips[16] realized that the scenario in Fig. 2 can be considered as *communication with side information at the encoder*. This perspective allows the design of watermarking schemes with high watermark rates. In this section, the application of such watermarking schemes to steganography is dicussed.

3.1. Costa's Scheme Exploiting Side-Information at the Encoder

Attacks by fine quantization can usually be modelled by an additive noise source. Costa[3] showed theoretically that the channel capacity for the communication scenario depicted in Fig. 2 with additive white Gaussian noise (AWGN) $\mathbf{v}$ instead of the quantization channel and independent identical distributed (IID) Gaussian cover data $\mathbf{x}$ is independent of the variance σ_{x}^2 of the cover data. The performance of an *ideal Costa scheme* (ICS) depends solely on the watermark-to-noise power ratio $\mathrm{WNR} = 10\log_{10}\sigma_{\mathsf{w}}^2/\sigma_{\mathsf{v}}^2$ [dB]. Costa's result is surprising since it shows that the original data $\mathbf{x}$ need not be considered as interference at the decoder although the decoder does not know $\mathbf{x}$. Costa presents a theoretic scheme which involves an L_x-dimensional random codebook $\mathcal{U}^{L_x}$ which is

$$\begin{aligned} \mathcal{U}^{L_x} \quad = \quad & \{\mathbf{u}_l = \mathbf{w}_l + \alpha\mathbf{x}_l \ \mid \ l \in \{1, 2, \dots, L_{\mathcal{U}}\}, \\ & \mathsf{w} \sim \mathcal{N}(0, \sigma_{\mathsf{w}}^2 I_{L_x}), \mathsf{x} \sim \mathcal{N}(0, \sigma_{\mathsf{x}}^2 I_{L_x})\}, \end{aligned} \tag{2}$$

where $\mathbf{w}$ and $\mathbf{x}$ are realizations of two L_x-dimensional independent random processes x and w with Gaussian PDF. $L_{\mathcal{U}}$ is the total number of codebook entries and I_{L_x} denotes the L_x-dimensional identity matrix. There exists at least one such codebook such that for $L_x \to \infty$ the capacity derived by Costa is achieved.

In ICS, the data $\mathbf{s}$ has a Gaussian probability density function PDF with variance $\sigma_{\mathsf{x}}^2 + \sigma_{\mathsf{w}}^2$ since x and w are statistically independent random variables with Gaussian PDF (Recall that the sum of two statistically independent Gaussian random variables produces again a Gaussian random variable). Amplitude scaling of $\mathbf{s}$ by $\sqrt{\sigma_{\mathsf{x}}^2/(\sigma_{\mathsf{x}}^2 + \sigma_{\mathsf{w}}^2)}$ delivers data $\mathbf{s}'$ with a PDF $p_{\mathsf{s}'}(s') = p_{\mathsf{x}}(s')$. Thus, the relative entropy between $\mathbf{x}$ and $\mathbf{s}'$ is zero which indicates a perfectly secure steganographic system. A more formal proof of this result has been given by J. K. Su.[17] Note that even additional Gaussian noise $\mathbf{v}$ cannot disturb this property when the amplitude scaling is adapted appropriately. However, the made assumptions about the cover data are not very realistic so that the obtained result is of little interest for practical steganography.

3.2. (ST-)SCS Watermarking

The ideal Costa scheme (ICS) is not practical due to the involved huge random codebook. Therefore, several suboptimal implementations of ICS have been proposed since 1999.[4,18,5,6,19,20] A natural simplification of ICS is the usage of a structured codebook $\mathcal{U}^{L_x}$, which in the most simple case can be constructed by a concatenation of scalar uniform quantizers leading to a sample-wise embedding and extraction rule. This approach has been independently proposed by several researchers.[18,5,6,20] We denote this approach as scalar Costa scheme (SCS).

In SCS, the message u is encoded into a sequence of letters $\mathbf{d}$, where $d_n \in \mathcal{D} = \{0, 1\}$ in case of binary SCS. Each of the letters is embedded into the corresponding cover data elements x_n. The embedding rule for the nth element is given by

$$s_n \quad = \quad x_n + \alpha\left(\mathcal{Q}_\Delta\left\{x_n - a_n\right\} - (x_n - a_n)\right), \quad \text{where} \ \ a_n = \Delta\left(\frac{d_n}{2} + k_n\right), \tag{3}$$

and $\mathcal{Q}_\Delta\{\cdot\}$ denotes scalar uniform quantization with step size Δ. The key $\mathbf{k}$ is a pseudo-random sequence with $k_n \in (0, 1]$ which has to be derived from the secret key K. This embedding scheme depends on two parameters: the quantizer step size Δ and the scale factor α. Both parameters can be jointly optimized to achieve a good trade-off between embedding distortion and detection reliability for a given noise variance of an AWGN attack.

Decoding of the message $\hat{u}$ from the received data $\mathbf{r}$ is based on the extracted received data $\mathbf{y}$. The extraction rule for the nth element is

$$y_n = \mathcal{Q}_\Delta\left\{r_n - k_n\Delta\right\} - (r_n - k_n\Delta), \tag{4}$$

where $|y_n| \leq \Delta/2$. y_n should be close to zero if $d_n = 0$ was sent, and close to $\pm\Delta/2$ for $d_n = 1$.

Although most of the work on SCS considers AWGN attacks, it has been shown as well that SCS is also robust against quantization attacks.[21] High watermark rates with low bit-error rates can be achieved by sophisticated error-correction coding of the embedded message u. Low watermark rates are unavoidable if the security of steganography requires low watermark power σ_{w}^2. In such a case, it might be useful to combine SCS with spread-transform (ST) watermarking, as proposed by Chen and Wornell.[15] In ST-SCS, the SCS watermark is embedded only into the projection of the cover data $\mathbf{x}$ onto a random spreading vector. A more detailed description and analysis of ST-SCS is given in our previous work.[8,21]

It can be shown[21] that the SCS watermark signal $\mathbf{w}$ has a uniform distribution of width $\alpha\Delta$ and is statistically independent from the cover data $\mathbf{x}$. Thus, the PDF of the watermarked signal $\mathbf{s}$ is given by

$$p_{\mathsf{s}}\left(s\right) = p_{\mathsf{x}}\left(s\right) * p_{\mathsf{w}}\left(s\right), \tag{5}$$

where "$*$" denotes linear convolution. In general, a non-zero relative entropy between x and s has to be accepted, meaning perfect security against statistical steganalysis cannot be achieved. However, we find $p_{\mathsf{s}}\left(s\right) \approx p_{\mathsf{x}}\left(s\right)$ for small watermark power σ_{w}^2 and a smooth cover data PDF $p_{\mathsf{x}}\left(x\right)$. The accuracy of this approximation is investigated for example image data in Sec. 5.

4. STEGANOGRAPHY WITH HISTOGRAM PRESERVING INFORMATION EMBEDDING

A new method for information embedding is proposed that preserves (in a statistical sense) the histogram of the cover data. The explanation of this new method is restricted to IID cover data $\mathbf{x}$, where each sample x_n can be considered a realization of a scalar random variable x. First, a data mapping method that allows a precise modification of the data histogram is described. Next, information embedding based on switched data mappings is introduced.

4.1. Data Mapping Achieving a Predefined Histogram

Histogram modification is an old problem, however, traditionally, elements of the input data having the same value are all mapped to the same value in the output data. Hence, the desired histogram can only be approximated, where the accuracy of the approximation depends strongly on the specific nature of the histogram of the input data. A predefined output histogram can be achieved perfectly if the data mapping allows that certain portions of data elements with identical value are mapped to different values in the output data. Such a mapping can be described by a matrix Γ with entries γ_{ij}, where γ_{ij} denotes the number of data elements being in the ith input histogram bin that have to be mapped to output values belonging to the j output histogram bin. Meşe and Vaidyanathan[22] propose a histogram modification method where the mean squared error (MSE) is minimized between the input and the output data. This is achieved by solving an integer linear programming problem to obtain the proper mapping matrix Γ. However, it appears that for many typical data types and a MSE distortion measure, a direct solution to the data mapping problem with predefined output histogram exists. Below, such a direct approach is described and a simple (approximative) implementation is outlined.

Let x denote a discrete random variable with the finite alphabet

$$\mathcal{X} = \{x^{(1)}, x^{(2)}, x^{(3)}, \ldots, x^{(N_x)}\}, \text{ where } N_x = |\mathcal{X}| < \infty \text{ and } x^{(1)} < x^{(2)} < x^{(3)} < \ldots < x^{(N_x)}. \tag{6}$$

The PMF $p_{\mathsf{x}}\left[x\right]$ of x can be estimated from an observation $\mathbf{x}$ of length L_x by normalizing its histogram $\hat{p}_{\mathsf{x}}\left[x\right]$ by the length L_x of the observed data sequence $\mathbf{x}$. Considering the histogram $\hat{p}_{\mathsf{x}}\left[x\right]$ is prefered to the PMF $p_{\mathsf{x}}\left[x\right]$ when an exact histogram modification is desired since all histogram values are integer numbers.

The data $\mathbf{x}$ shall be modified into data $\mathbf{y}$ with a predefined histogram $\hat{p}_{\mathsf{y}}\left[y\right]$. We assume $y_n \in \mathcal{X}$ without loss of generality. It appears that the derivation of the mapping $\mathbf{x} \rightarrow \mathbf{y}$ does not depend on the exact values of $x^{(1)}, x^{(2)}, x^{(3)}, \ldots, x^{(N_x)}$, but only on the histogram values $h_{\mathsf{x}}\left[i\right] = \hat{p}_{\mathsf{x}}\left[x^{(i)}\right]$ and $h_{\mathsf{y}}\left[i\right] = \hat{p}_{\mathsf{y}}\left[y^{(i)}\right]$ for $i \in \{1, 2, \ldots, N_x\}$. We demand that the mapping $\mathbf{x} \rightarrow \mathbf{y}$ introduces as little distortion as possible. Here, a mean squared error (MSE) distortion measure

$$D_{\text{Map}} = \frac{1}{L_x}\sum_{n=1}^{L_x} \mathrm{d}\left(x_n, y_n\right) = \frac{1}{L_x}\sum_{i=1}^{N_x}\sum_{j=1}^{N_x} \gamma_{ij}\mathrm{d}\left(x^{(i)}, x^{(j)}\right), \text{ with } \mathrm{d}\left(x_n, y_n\right) = \left(x_n - y_n\right)^2, \tag{7}$$

is assumed. An important consequence of this distortion measure is that the mapping $\mathbf{x} \rightarrow \mathbf{y}$ with minimum MSE must preserve the relation between different data elements. That is, for two input data elements x_n and x_m with $x_n > x_m$, the corresponding mapped data elements y_n and y_m have to satisfy $y_n \geq y_m$ for all $n, m \in \{1, 2, \ldots, L_x\}$. This property is stated more precisely in the following theorem, which has been proven by Tzschoppe et al..[23]

THEOREM 4.1. *Let $\mathbf{x}$ and $\mathbf{y}$ denote vectors of length L_x with elements sorted in increasing order so that $x_1 \leq x_2 \leq x_3 \leq \ldots \leq x_{L_x}$ and $y_1 \leq y_2 \leq y_3 \leq \ldots \leq y_{L_x}$. The MSE distortion $\mathrm{D} = (1/L_x)\sum_{n=1}^{L_x}(x_n - y_n)^2$ is never larger than the MSE $\mathrm{D}_\pi = (1/L_x)\sum_{n=1}^{L_x}(x_n - y_{\pi(n)})^2$, where $\pi(n)$ denotes an arbitrary permutation of the element indices.*

Now, consider the first histogram bin $i = 1$ of the output data which must contain $h_{\mathsf{y}}\left[1\right]$ data elements. Due to Theorem 4.1, all data elements y_n belonging into this histogram bin have to be derived from the $h_{\mathsf{y}}\left[1\right]$ smallest input data elements in order to achieve a mapping with minimum MSE distortion. Next, all output data elements y_n belonging to the second histogram bin have to be derived from the remaining $h_{\mathsf{y}}[2]$ smallest input data elements. Proceeding this argumentation shows that all bins of the output histogram have to be filled in increasing order by mapping the input data with values in increasing order. Note that the exact value x_n of the input data is irrelevant; only the relation between the values of the input data matters. The entries γ_{ij} of the mapping matrix Γ can be determined for increasing j by inspecting to which input histogram bin i those data elements x_n belong that have to be mapped to the value $x^{(j)}$ and thus fall into the jth output histogram bin. Knowing the mapping matrix Γ, the data mapping $\mathbf{x} \to \mathbf{y}$ is realized by randomly selecting γ_{ij} input data elements with value $x^{(i)}$ and mapping them to the corresponding output data elements with value $x^{(j)}$.

The described data mapping requires that the entire input data $\mathbf{x}$ is known before the mapping can be designed and applied. However, in many practical cases, the input process x is only described by its PMF $p_{\mathsf{x}}[x]$ and elements should be mapped directly so that the output process y achieves a certain target PMF $p_{\mathsf{y}}[y]$. Such a mapping becomes possible when considering $\gamma_{ij}/h_{\mathsf{x}}[i]$ as the probability with that a value $x_n = x^{(i)}$ is mapped to the output value $y_n = x^{(j)}$. Hence, the mapping $x_n \to y_n$ is no longer deterministic, but depends on the probabilities $\gamma_{ij}/h_{\mathsf{x}}[i]$. For large data sequences $\mathbf{x}$, the normalized histogram of the output data $\mathbf{y}$ tends to the desired PMF $p_{\mathsf{y}}[y]$.

We propose an efficient implementation of the random mapping $x_n \to y_n$ which does not require the computation of the mapping matrix Γ. However, it suffcies to randomize the input data x_n and to quantize this randomized input data to the output data y_n. The mapping can be characterized completely by the required scalar quantizer $\mathcal{Q}_t$, which itself is characterized by the set $\mathcal{T} = \{t_1, t_2, \dots, t_{N_x-1}\}$ of decision thresholds. We describe a mapping algorithm that operates on the randomized index i of the possible input symbols $x^{(i)}$ so that the case of unequally spaced symbols is covered more easily. Fig. 3 illustrates the derivation of the set $\mathcal{T}$ for an example with $N_x = 10$. The upper diagrams show the given input histogram $h_{\mathsf{x}}[i]$ and the desired output histogram $h_{\mathsf{y}}[j]$.

Let i_n denote the symbol index in $\{1, 2, \dots, N_x\}$ of a given input data element x_n. This index is randomly mapped on a continuous valued random variable t with

$$t_n = i_n - a_n, \tag{8}$$

where a_n is drawn from a continuous valued random variable a with uniform support over the range $[0, 1)$. The PDF of the random variable t is proportional to the function

$$\tilde{h}_{\mathsf{x}}(t) = \sum_{i=1}^{N_x} h_{\mathsf{x}}[i] \operatorname{rect}\left(t + \frac{1}{2} - i\right), \quad \text{with } \operatorname{rect}(x) = \begin{cases} 1 & ; \quad -1/2 < x \le 1/2 \\ 0 & ; \quad \text{else.} \end{cases} \tag{9}$$

$\tilde{h}_{\mathsf{x}}(t)$ is shown for the given example by a dotted line in the upper right diagram in Fig. 3. We introduce

$$h_{\mathsf{x}}(t) = \sum_{i=1}^{N_x} h_{\mathsf{x}}[i]\, \delta(t - i) \quad \text{and} \quad h_{\mathsf{y}}(t) = \sum_{j=1}^{N_x} h_{\mathsf{y}}[j]\, \delta(t - j), \tag{10}$$

with $\delta(x)$ being a Dirac impulse, in order to obtain well defined integrals of the histograms $h_{\mathsf{x}}[i]$ and $h_{\mathsf{y}}[j]$. The lower diagrams in Fig. 3 show the integrals $\int_0^t h_{\mathsf{y}}(\tau)\,\mathrm{d}\tau$, $\int_0^t h_{\mathsf{x}}(\tau)\,\mathrm{d}\tau$ and $\int_0^t \tilde{h}_{\mathsf{x}}(\tau)\,\mathrm{d}\tau$, for the given example histograms.

The quantizer $\mathcal{Q}_t$ is defined as the function

$$\mathcal{Q}_t(t) = \begin{cases} x^{(1)} & ; \quad t \le t_1 \\ x^{(j)} & ; \quad t_{j-1} < t \le t_j \quad \forall j \in \{2, 3, \dots, N_x - 1\} \\ x^{(N_x)} & ; \quad t_{N_x-1} < t \end{cases} \tag{11}$$

so that the operation

$$y_n = \mathcal{Q}_t(t_n) = \mathcal{Q}_t(i_n - a_n) \tag{12}$$

produces the desired output data y_n if the set $\mathcal{T}$ of decision thresholds fulfills the integral equations

$$\int_0^1 h_{\mathsf{y}}(\tau)\,\mathrm{d}\tau = \int_0^{t_1} \tilde{h}_{\mathsf{x}}(\tau)\,\mathrm{d}\tau \quad \text{and} \quad \int_{j-1}^{j} h_{\mathsf{y}}(\tau)\,\mathrm{d}\tau = \int_{t_{j-1}}^{t_j} \tilde{h}_{\mathsf{x}}(\tau)\,\mathrm{d}\tau \quad \text{for } j \in \{2, 3, \dots, N_x - 1\}. \tag{13}$$

The meaning of these integral equations is illustrated in Fig. 3 by the dashed lines in the lower diagrams. Note that the highest threshold t_{10} is redundant.

We summarize that input data $\mathbf{x}$ with a given histogram $h_{\mathsf{x}}[i]$ can be mapped to output data $\mathbf{y}$ which achieves (in a statistical sense; for long data sequences) a predefined histogram $h_{\mathsf{y}}[j]$ with minimum MSE distortion using the operations defined in (11) and (12), where the quantizer decision thresholds have to be computed from (13).

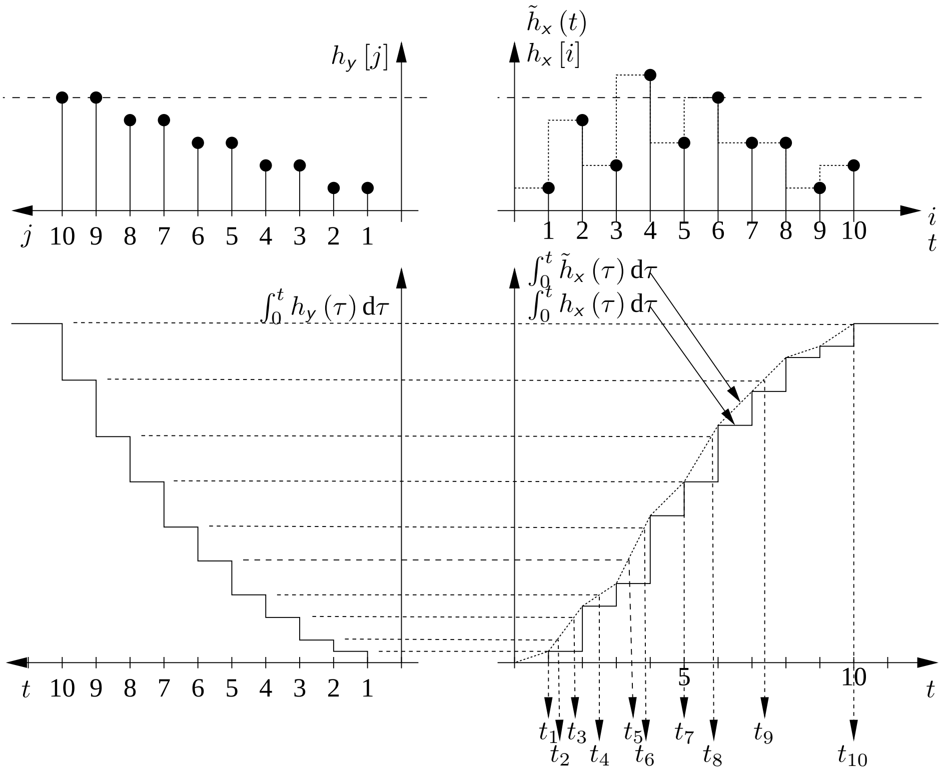

Figure 3. Derivation of thresholds $\{t_1, t_2, \ldots, t_{N_x}\}$ for the data mapping $x \to y$. The input histogram $h_x\,[i]$ is mapped onto the predefined output histogram $h_y\,[j]$.

4.2. Secure Information Embedding Based on Switched Data Mapping

The previously described data mapping can be used for information embedding if different mappings $x_n \to r_n$ are defined and the data to be embedded is used to switch between the possible mapping rules. Although this principle is quite general, we focus here on the case of IID input data $\mathbf{x}$ where each data element is characterized by the discrete valued random variable x with the support set $\mathcal{X}$. In order to render successful steganalysis impossible, it is required that the PMFs of the cover data **x** and the steganographic data **r** are (almost) identical, that is $p_{\mathbf{r}}\,[\mathbf{r}] = p_{\mathbf{x}}\,[\mathbf{r}]$. Due to the IID assumption, this statement can be reduced to the equality $p_r\,[r] = p_x\,[r]$.

We assume that the receiver of the steganographic data $\mathbf{r}$ has no access to the original cover data $\mathbf{s}$. Thus, it must be possible to decode the message $\hat{u}$ simply by inspection of $\mathbf{r}$. Further, if error-free communication is desired, the set of all possible steganographic data $\mathbf{r}$ that deliver a certain message $\hat{u} = u_0$ must be disjoint from the set of possible steganographic data delivering $\hat{u} \neq u_0$. Thus, for information embedding, the cover data $\mathbf{x}$ has to be mapped onto members from disjoints sets for the different possible secret messages u. Such an information embedding principle is already known in the digital watermarking communicty as quantization index modulation (QIM), as proposed by Chen and Wornell.[24] QIM allows for error-free transmission in the case of noise-less channels. For watermarking applications, QIM turned out to be not appropriate since QIM is not very robust against channel noise. However, there is no channel noise in steganography if the used quantizer constellation fits to the quantizers for lossy compression of the cover data. A general QIM scheme does not necessarily preserve the PMF of the cover data, however, this can be achieved when the generalized data mapping introduced in the previous subsection is used to map the cover data onto the different sets of quantizer representatives.

The required message dependent data mapping operates sample-wise in the simple case of IID cover data $\mathbf{x}$. To enable binary embedding, two disjoint sets $\mathcal{X}_0$ and $\mathcal{X}_1$ have to be defined, where

$$\mathcal{X}_0 \cup \mathcal{X}_1 = \mathcal{X} \quad \text{and} \quad \mathcal{X}_0 \cap \mathcal{X}_1 = \emptyset. \tag{14}$$

These sets $\mathcal{X}_0$ and $\mathcal{X}_1$ can be interpreted as the representatives of two different quantizers, which emphasises the relationship of the new information embedding technique to QIM. The message u is encoded into a stream $\tilde{\mathbf{b}}$ with binary elements $\tilde{b}_n \in \{0, 1\}$.

Next, $\tilde{\mathbf{b}}$ is embedded into $\mathbf{x}$ by the mapping of $x_n \rightarrow r_n$ using the mappings Map$(\mathcal{X}, \mathcal{X}_0)$ and Map$(\mathcal{X}, \mathcal{X}_1)$ for $\tilde{b}_n = 0$ and $\tilde{b}_n = 1$, respectively.

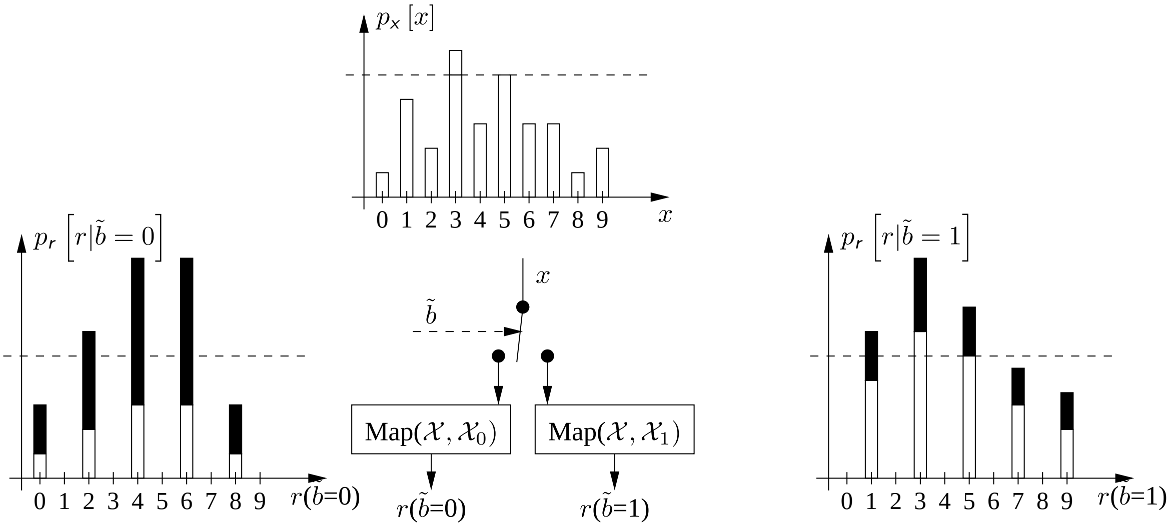

Figure 4. Illustration of switched data mapping for an entropy coded message $\tilde{\mathbf{b}}$.

Fig. 4 depicts the influence of information embedding by switched data mapping on the conditional PMFs of the steganographic data $\mathbf{r}$ for an example with $\mathcal{X} = \{0, 1, 2, \dots, 9\}$, $\mathcal{X}_0 = \{0, 2, 4, 6, 8\}$ and $\mathcal{X}_1 = \{1, 3, 5, 7, 9\}$. The data mapping rules Map$(\mathcal{X}, \mathcal{X}_0)$ and Map$(\mathcal{X}, \mathcal{X}_1)$ have to be designed such that the conditional PMFs $p_r\left[r|\tilde{b}=0\right]$ and $p_r\left[r|\tilde{b}=1\right]$ are scaled proportianal to the cover PDF $p_x\,[x]$ for all members of the set $\mathcal{X}_0$ and $\mathcal{X}_1$, respectively, and zero elsewhere. The black part of the bars in the leftmost and rightmost graph in Fig. 4 indicates the amount of data that has been mapped from values of the set $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. Formally, the conditional PMFs are given by

$$p_r\left[r|\tilde{b}=i\right] \;=\; \begin{cases} \frac{p_x[r]}{\mathrm{Prob}(x\in\mathcal{X}_i)} & ; r \in \mathcal{X}_i \\ 0 & ; r \notin \mathcal{X}_i \end{cases} \qquad \text{for } i \in \{0, 1\}. \tag{15}$$

The unconditional PMF $p_r\,[r]$ of the steganographic data r is given by

$$\begin{aligned} p_r\,[r] &= \mathrm{Prob}\left(\tilde{b}=0\right)\cdot p_r\left[r|\tilde{b}=0\right] + \mathrm{Prob}\left(\tilde{b}=1\right)\cdot p_r\left[r|\tilde{b}=1\right] \\ &= \begin{cases} \frac{\mathrm{Prob}\left(\tilde{b}=0\right)}{\mathrm{Prob}(x\in\mathcal{X}_0)} p_x\,[r] & ; r \in \mathcal{X}_0 \\ \frac{\mathrm{Prob}\left(\tilde{b}=1\right)}{\mathrm{Prob}(x\in\mathcal{X}_1)} p_x\,[r] & ; r \in \mathcal{X}_1 \end{cases} \\ &= p_x\,[r] \quad \text{iff } \mathrm{Prob}\left(\tilde{b}=i\right) = \mathrm{Prob}\left(x \in \mathcal{X}_i\right) \quad \forall\, i \in \{0,1\}. \end{aligned} \tag{16}$$

We observe that the PMF of the cover data is not modified by the proposed information embedding scheme if the probability $\mathrm{Prob}\left(\tilde{b}=1\right)$ of "1"-bits in the encoded message $\tilde{\mathbf{b}}$ is equal to the probability $\mathrm{Prob}\left(x \in \mathcal{X}_1\right)$ that the elements of the cover data $\mathbf{x}$ belong to the set $\mathcal{X}_1$. Note that the condition $\mathrm{Prob}\left(\tilde{b}=0\right) = \mathrm{Prob}\left(x \in \mathcal{X}_0\right)$ is fulfilled as soon as $\mathrm{Prob}\left(\tilde{b}=1\right) = \mathrm{Prob}\left(x \in \mathcal{X}_1\right)$ is valid since $\mathrm{Prob}\left(\tilde{b}=0\right) = 1 - \mathrm{Prob}\left(\tilde{b}=1\right)$. We denote $P^1 = \mathrm{Prob}\left(x \in \mathcal{X}_1\right)$ as the channel state for the given cover data $\mathbf{r}$. P^1 is a property of the cover data that can not be modified by the information embedder. Since security against steganalysis can be achieved only for $\mathrm{Prob}\left(\tilde{b}=1\right) = P^1$, the amount of information that can be embedded per cover data element (steganographic capacity) is given by the binary entropy function

$$H(P^1) = P^1 \log_2(P^1) + (1 - P^1)\log_2(1 - P^1) \quad \text{[bits/cover element]}. \tag{17}$$

The steganographic capacity is 1 bit/(cover element) for $P^1 = 0.5$ and decreases for $P^1 \neq 0.5$. The capacity is zero for $P^1 = 0$ and $P^1 = 1$.

So far, the histogram preserving embedding method has been described in terms of an encoded binary message $\tilde{\mathbf{b}}$. It appears that security can be achieved only for $\mathrm{Prob}\left(\tilde{b} = 1\right) = P^1$. However, usually it is assumed that a binary encoded message u contains as many zeros as ones. This is particularly true if binary encoding of u is combined with encryption to ensure that only authorized parties are able to decode the embedded information. However, in general, the cover data might have $P^1 \neq 0.5$ which requires an unequal distribution of zeros and ones within the encoded message $\tilde{\mathbf{b}}$. In order to solve this problem, the message encoding is separated into two steps. First, u is transformed into a binary sequence and encrypted into a binary message $\mathbf{b}$ using the secure key K. Any secure encryption algorithm can be used. Second, $\mathbf{b}$ is processed by an entropy decoder (e.g. Huffman decoder), where the entropy code (e.g. Huffman code) has been designed for a binary source with probablity P^1 for the source symbol "1". The output of the entropy decoder is the binary sequence $\tilde{\mathbf{b}}$ with $\mathrm{Prob}\left(\tilde{b} = 1\right) = P^1$ as desired. Such an encoding process is depicted in Fig. 5.

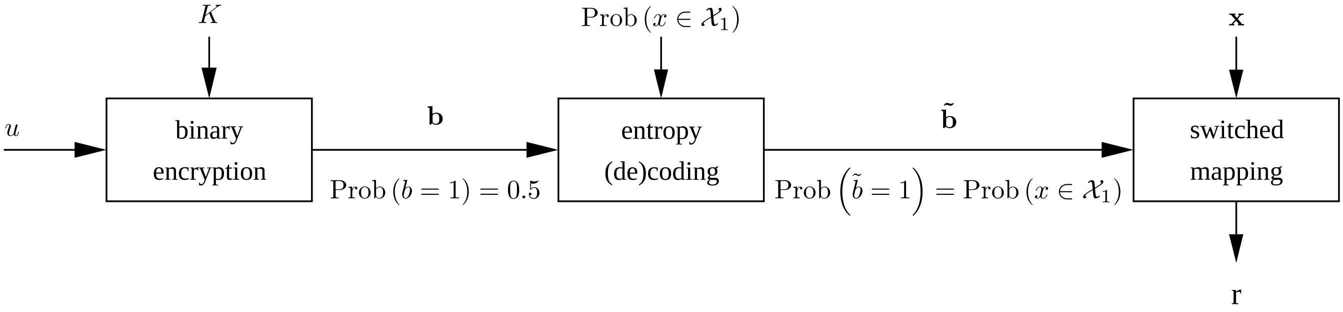
K
$\mathrm{Prob}\,(x \in \mathcal{X}_1)$
$\mathbf{x}$
u
binary encryption
$\mathbf{b}$
$\mathrm{Prob}\,(b = 1) = 0.5$
entropy (de)coding
$\tilde{\mathbf{b}}$
$\mathrm{Prob}\left(\tilde{b} = 1\right) = \mathrm{Prob}\,(x \in \mathcal{X}_1)$
switched mapping
$\mathbf{r}$

Figure 5. Encryption and entropy coding of the message u for histogram preserving steganography.

The presented information embedding method is designed such, that the relative entropy between the cover data $\mathbf{x}$ and the steganographic data $\mathbf{r}$ tends to zero. Note that this limit can be achieved only for long data sequences. A perfect match of the histograms of the cover data and the steganographic data can never be achieved since the randomness of the information to be embedded does not allow a deterministic mapping $\mathbf{x} \rightarrow \mathbf{r}$ as proposed by Meşe and Vaidyanathan.[22] However, the random mapping described in the second part of the previous subsection provides a simple and elegant method to produce steganographic data $\mathbf{r}$ having the same statistics as the cover data $\mathbf{x}$.

As already described in Sec. 2, secure image steganography requires also that the distortion introduced by the information embedding algorithm is not too large. The MSE embedding distortion is determined by the mapping distortion of both mapping rules Map$(\mathcal{X}, \mathcal{X}_0)$ and Map$(\mathcal{X}, \mathcal{X}_1)$ which can be computed from (7). Unfortunately, the mapping distortion is mainly determined by the sets $\mathcal{X}_0$ and $\mathcal{X}_1$ and the cover data histogram. In this paper, we do not consider the adaption of the sets $\mathcal{X}_0$ and $\mathcal{X}_1$ to the distortion constraint since the decoder of the steganographic data must know $\mathcal{X}_0$ and $\mathcal{X}_1$, too. Thus, the distortion can be controlled only by applying the described information embedding to a fraction of all cover data elements.

5. IMAGE STEGANOGRAPHY

We designed two systems for image steganography with JPEG compressed steganographic images $\mathbf{r}$. The £rst system is based on ST-SCS watermarking and the second system is based on histogram-preserving data mappings (HPDM). Due to space constraints it is impossible to describe both systems in detail. Thus, only a rough outline of the design concepts is given.

5.1. Outline of the Implemented Systems for Image Steganography

We exploit only a very simple stochastic model of the cover data based on a two-dimensional Discrete Cosine Transform (DCT) of non-overlapping 8×8 blocks of the image pixels. Fig. 6 illustrates the 8×8 block DCT, which is denoted as BDCT subsequently. The ith 8×8 block in row-scan is transformed into 64 DCT coef£cients $\{x_{i,1}^{\mathrm{BDCT}}, x_{i,2}^{\mathrm{BDCT}}, \ldots, x_{i,j}^{\mathrm{BDCT}}, \ldots, x_{i,64}^{\mathrm{BDCT}}\}$. Next, the coef£cients with identical frequency index j from all 8×8 blocks compose the signal $\mathbf{x}_j^{\mathrm{BDCT}}$, which can be considered a subchannel. Thus, there are 64 subchannels, all having the same length $L_{x^{\mathrm{BDCT}}}$ which is identical to the number of 8×8 blocks in the given image $\mathbf{x}$. The common zig-zag scan[25] is used for labeling the 64 signals $\mathbf{x}_j^{\mathrm{BDCT}}$.

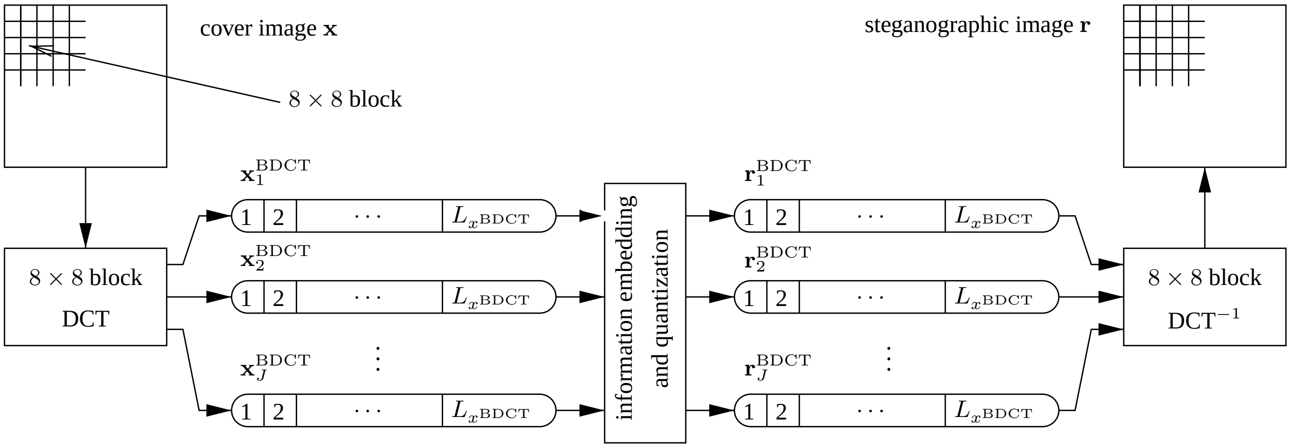
cover image $\mathbf{x}$
8×8 block
steganographic image $\mathbf{r}$
$\mathbf{x}_1^{\mathrm{BDCT}}$
$\mathbf{x}_2^{\mathrm{BDCT}}$
$\mathbf{x}_J^{\mathrm{BDCT}}$
8×8 block DCT
1 2 ⋯ $L_{x^{\mathrm{BDCT}}}$
information embedding and quantization
$\mathbf{r}_1^{\mathrm{BDCT}}$
$\mathbf{r}_2^{\mathrm{BDCT}}$
$\mathbf{r}_J^{\mathrm{BDCT}}$
8×8 block DCT^{-1}

Figure 6. Image steganography using the 8×8 block DCT (BDCT).

We model each component $\mathbf{x}_j^{\mathrm{BDCT}}$ by an IID random process. This model is not always very accurate, nevertheless this assumption is made in order to keep the complexity of both systems at an acceptable level. In natural image data, statistical dependencies between different subchannels and different elements within one subchannel have to be considered by the information embedder and by the adversary. However, the focus of this paper is on highlighting the general properties of ST-SCS and HPDM, which can be done most easily for the described simple stochastic image model. Each subchannel is quantized according to JPEG compression with quality factor 75. We assume that the cover image $\mathbf{x}$ has not been compressed before, which improves the quality of the steganographic image $\mathbf{r}$. Nevertheless, both systems would work as well for an already compressed cover image $\mathbf{x}$. Only the £rst 21 subchannels in zigzag-scan were used for information embedding, since the very high frequency subchannels are quantized too strongly so that almost no space for information hiding is left. The total embedding distortion is distributed over all 21 used subchannels so that the MSE embedding distortion per subchannels is roughly proportional to the respective distortion of simple quantization due to JPEG compression.

JPEG quantization is considered in the design of the ST-SCS based scheme as a simple additive noise source. The data for all used subchannels is error correction encoded with a rate 1/3 Turbo code and the embedding strength is chosen so that the bit-error-rate after Turbo decoding is below 10^{-5}. The embedding rate can be modi£ed by the spreading factor of the spread-transform. A spreading factor of 1 provides the highest data rate, but produces also the largest embedding distortion. The SCS quantizer step size has a £xed relation to the step size of JPEG quantization. Thus, no side information has to be transmitted to the decoder.

The HPDM based scheme provides error-free communication, since JPEG quantization is already considered within the embedding algorithm. The sets $\mathcal{X}_0$ and $\mathcal{X}_1$ contain all possible even and all possible odd coef£cient per subchannel, respectively. This simple rule ensures that no side information about the choice of $\mathcal{X}_0$ and $\mathcal{X}_1$ has to be transmitted to the decoder. However, due to this arrangement, each subchannel can have a different channel state P^1 and the distortion D_{Map} can be controlled only by varying the fraction ρ of subchannel elements used for information embedding. P^1 and ρ are quantized to 16 possible values and transmitted to the decoder as side information. This side information is always transmitted in the £rst subchannel, where $P^1 = 0.5$ is assumed for the transmission of the side information. Note that encryption of the side information is required to resist steganalysis.

5.2. Experimental Results

Experiments have been made for both schemes with several grayscale images, different rates of hidden information and different embedding distortions. Here, we discuss only example results for the grayscale test image “Lenna” of size 512×512 which re¤ect the most important results obtained by all experiments. Standard JPEG compression with quality factor 75 of the cover image $\mathbf{r}$ gives a compressed image $\mathbf{s}$ with $\mathrm{PSNR} = 38.08$ dBand size of 254576 bits = 31.08 kB. Experiments with 100 different messages have been made with the emedding systems ST-SCS and HPDM, where the subsequently presented results are averaged over all 100 simulations. The spreading factor of ST-SCS has been set to one, thus, plain SCS has been applied. An embedding distortion of 36.42 dB compared to the cover image $\mathbf{r}$ has been achieved for error-free communication of the hidden message. The parameters ρ of HPDM have been adapted such that the same embedding distortion of 36.42 dB results. The quality loss of about 1.66 dB compared to the directly JPEG compressed image $\mathbf{s}$ is low enough so that no difference between the images $\mathbf{s}$ and $\mathbf{r}$ can be perceived subjectively.

	PNSR	size of steganographic image $\mathbf{r}$	size of hidden message	ratio (size of message)/(size of $\mathbf{s}$)
HPDM	36.42 dB	266118 Bits = 32.49 kB	32096 Bits = 3.92 kB	12.61 %
ST-SCS	36.42 dB	292755 Bits = 35.74 kB	28668 Bits = 3.50 kB	11.26 %

Table 1. Experimental results for 100 simulations with the cover image “Lenna” of size 512×512.

Table 1 shows the resulting size of the steganographic image $\mathbf{r}$ and the size of the hidden message $\mathbf{b}$ for HPDM and ST-SCS. HPDM enables a slightly larger amount of hidden information, where both systems allow at the given embedding distortion a message length of more than 11 % of the directly JPEG compressed image. Note that the size of the steganographic image $\mathbf{r}$ is much larger for ST-SCS than for HPDM. This effect occurs since ST-SCS does not preserve the PMF of the DCT coef£cients for higher frequencies very well, which can be concluded from the measured relative entropies shown in Fig. 7. Thus, the entropy coder included within JPEG compression performs worse than in the case of the directly quantized image data. The size of the steganographic image $\mathbf{r}$ in case of HPDM is also sligthly larger than the size of $\mathbf{s}$, although the relative entropy for HPDM is very small for all subchannels. This effect shows that the assumption about independent DCT subchannels is not accurate. Entropy encoding within JPEG compression takes advantage of dependencies between different DCT coef£cients. However, HPDM breaks these dependencies which results in the sligthly increased £le size.

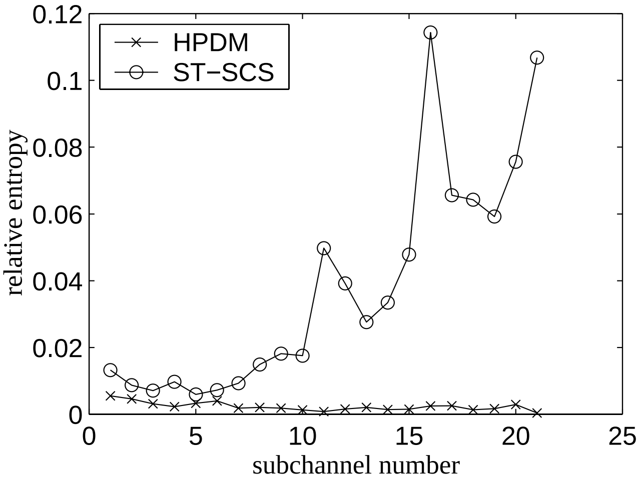

Figure 7. Relative entropy between $\mathbf{r}_i^{\mathrm{BDCT}}$ and $\mathbf{s}_i^{\mathrm{BDCT}}$ for $i = 1, 2, \ldots, 21$ for both embedding schemes HPDM and ST-SCS.

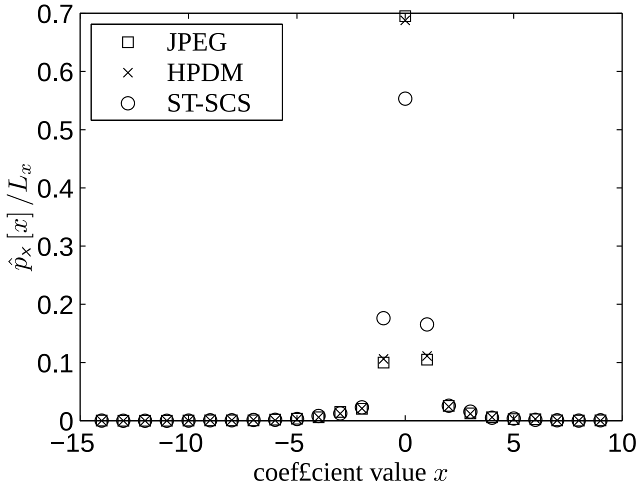

Figure 8. Example histograms taken from $\mathbf{s}_{15}^{\mathrm{BDCT}}$, $\mathbf{r}_{15}^{\mathrm{BDCT}}$ with HPDM embedding, and $\mathbf{r}_{15}^{\mathrm{BDCT}}$ with ST-SCS embedding.

Fig. 7 shows that the measured relative entropy between $\mathbf{r}_i^{\mathrm{BDCT}}$ and $\mathbf{s}_i^{\mathrm{BDCT}}$ is very small for all subchannels when HPDM based steganography is used. Thus, HPDM can be considered as a rather secure system. ST-SCS produces signi£cantly larger relative entropies in subchannels having a relatively “peaky” PMF of the quantized DCT coef£cients. In these cases, the convolution of the cover PDF with that of the ST-SCS watermark signal leads to signi£cant modi£cations of the PMF of the quantized DCT coef£cients with hidden information bits. Fig. 8 illustrates this effect for measured example PMFs taken from the $15th$ DCT subchannel. Direct JPEG compression and HPDM give almost the same PMF. However, ST-SCS reduces the amount of zero coef£cients while increasing the number of coef£cients with value ± 1.

6. CONCLUSIONS

Steganography based on blind ST-SCS watermarking and based on histogram preserving data mappings (HPDM) has been investigated. The new HPDM scheme gives in the limit of long cover data sequences a zero relative entropy between the cover data and the steganographic data which proves security of the system within a given stochastic data model. ST-SCS and HPDM based image steganography allows for almost the same rate of hidden information at a £xed embedding distortion. However, ST-SCS leaves significant traces in the statistics of the DCT coefficients of the steganographic image as soon as the PDF of the DCT coefficients of the cover image is not very smooth. The presented image steganography based on HPDM can be considered secure within the exploited simple stochastic image model. Further, we believe that the extension of a HPDM based steganographic system to more complicated stochastic image models is straight forward. Such an improved system should give security even against very sophisticated steganalysis.

REFERENCES

1. G. J. Simmons, “The prisoners’ problem and the subliminal channel,” in *Advances in Cryptology: Proceeedings of CRYPTO 83*, D. Chaum, ed., Plenum Press, 1983.
2. C. Cachin, “An information-theoretic model for steganography,” in *Proceedings of 2nd Workshop on Information Hiding*, D. Aucsmith, ed., vol. 1525, Lecture Notes in Computer Science, Springer, (Portland, Oregon, USA), May 1998.
3. M. H. M. Costa, “Writing on dirty paper,” *IEEE Transactions on Information Theory* **29**, pp. 439–441, May 1983.
4. B. Chen and G. W. Wornell, “Provably robust digital watermarking,” in *Proceedings of SPIE: Multimedia Systems and Applications II (part of Photonics East ’99)*, vol. 3845, pp. 43–54, (Boston, MA, USA), September 1999.
5. M. Ramkumar, *Data Hiding in Multimedia: Theory and Applications.* PhD thesis, Dep. of Electrical and Computer Engineering, New Jersey Institute of Technology, Kearny, NJ, USA, November 1999.
6. J. J. Eggers, J. K. Su, and B. Girod, “A blind watermarking scheme based on structured codebooks,” in *Secure Images and Image Authentication, Proc. IEE Colloquium*, pp. 4/1–4/6, (London, UK), April 2000.
7. J. J. Eggers, J. K. Su, and B. Girod, “Robustness of a blind image watermarking scheme,” in *Proceedings of the IEEE Intl. Conference on Image Processing 2000 (ICIP 2000)*, (Vancouver, Canada), September 2000.
8. J. J. Eggers, J. K. Su, and B. Girod, “Performance of a practical blind watermarking scheme,” in *Proc. of SPIE Vol. 4314: Security and Watermarking of Multimedia Contents III*, (San Jose, Ca, USA), January 2001.
9. H. Farid, “Detecting steganographic messages in digital images,” tech. rep., Dartmouth College, Computer Science, 2001. TR2001-412.
10. N. Provos, “Defending against statistical steganalysis,” tech. rep., Center for Information Technology Integration, University of Michigan, 2001. CITI Techreport 01-4.
11. N. Provos, “Defending against statistical steganalysis,” in *10th USENIX Security Symposium*, (Washington DC, USA), August 2001.
12. J. Fridrich, M. Goljan, and R. Du, “Detecting LSB steganography in color and gray-scale images,” *IEEE Multimedia (Multimedia and Security)* , Oct-Dec 2001.
13. A. Westfeld, “High capacity despite better steganalysis: F5 – a steganographic algorithm,” in *Proceedings of 4th Information Hiding Workshop 2001*, I. S. Moskowitz, ed., vol. 2137, pp. 301–314, Lecture Notes in Computer Science, Springer, (Pittsburgh, PA, USA), April 2001.
14. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.
15. B. Chen and G. W. Wornell, “Achievable performance of digital watermarking systems,” in *Proceedings of the IEEE Intl. Conference on Multimedia Computing and Systems (ICMCS ’99)*, vol. 1, pp. 13–18, pp. 13–18, (Florence, Italy), June 1999.
16. I. J. Cox, M. L. Miller, and A. L. McKellips, “Watermarking as communications with side information,” *Proceedings of the IEEE, Special Issue on Identification and Protection of Multimedia Information* **87**, pp. 1127–1141, July 1999.
17. J. K. Su, “On the information-hiding capacity of a memoryless gaussian watermarking game.” preprint, November 2001.
18. B. Chen and G. Wornell, “Preprocessed and postprocessed quantization index modulation methods for digital watermarking,” in *Proc. of SPIE Vol. 3971: Security and Watermarking of Multimedia Contents II*, pp. 48–59, (San Jose, Ca, USA), January 2000.
19. J. Chou, S. Pradhan, L. E. Ghaoui, and K. Ramchandran, “A robust optimization solution to the data hiding problem using distributed source coding principles,” in *Proc. of SPIE Vol. 3974: Image and Video Communications and Processing 2000*, (San Jose, Ca, USA), January 2000.
20. M. Kesal, M. K. Mihçak, R. Kötter, and P. Moulin, “Iterative decoding of digital watermarks,” in *Proc. 2nd Symp. on Turbo Codes and Related Topics*, (Brest, France), September 2000.
21. J. J. Eggers, *Information Embedding and Digital Watermarking as Communication with Side Information.* PhD thesis, Lehrstuhl für Nachrichtentechnik I, Universität Erlangen-Nürnberg, Erlangen, Germany, November 2001. preprint.
22. M. Meşe and P. P. Vaidyanathan, “Optimal histogram modification with MSE metric,” in *Proceedings of the IEEE Intl. Conference on Speech and Signal Processing 2001 (ICASSP 2001)*, (Salt Lake City, Utah, USA), May 2001.
23. R. Tzschoppe, R. Bäuml, and J. Eggers, “Histogram modification with minimum MSE distortion,” tech. rep., Telecommunications Laboratory, Universtiy of Erlangen-Nuremberg, December 2001.
24. B. Chen and G. W. Wornell, “Digital watermarking and information embedding using dither modulation,” in *Proc. of IEEE Workshop on Multimedia Signal Processing (MMSP-98)*, pp. 273–278, (Redondo Beach, CA, USA), Dec. 1998.
25. G. K. Wallace, “The JPEG still picture compression standard.,” *Communications of the ACM* **34**, pp. 31–44, April 1991.