Available online at www.sciencedirect.com
ScienceDirect
Signal Processing 87 (2007) 1301–1313
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

Quality enhancement of printed-and-scanned images using distributed coding ☆

S. Voloshynovskiy*, O. Koval*, F. Deguillaume, T. Pun

University of Geneva - CUI, 24 rue du Général Dufour, CH-1211 Geneva 4, Switzerland

Received 20 December 2005; received in revised form 31 October 2006; accepted 7 November 2006
Available online 8 December 2006

Abstract

In this paper we address visual communications via print-and-scan channels from an information-theoretic point of view as communications with side information that targets quality enhancement of visual data at the output of this type of channels. The solution to this problem addresses important aspects of multimedia data processing and management. A practical approach to side information communications for printed documents based on Wyner–Ziv and Gray setups is analyzed in the paper that assumes two separated communications channels where an appropriate distributed coding should be elaborated. The printing channel is considered to be a direct visual channel for images (''analog'' channel with degradations). The ''digital channel'' is considered to be an appropriate auxiliary channel exploited to communicate the information exploited for quality enhancement of printed-and-scanned image. We demonstrate both theoretically and practically how one can benefit from this sort of ''distributed paper communications''.

Keywords: Data-hiding; Printed-and-scanned documents; Communications with side-information; Bar codes; Distributed coding; Joint source-channel coding; Gel'fand–Pinsker problem; Wyner–Ziv coding; Gray coding; Image enhancement

1. Introduction

Printed documents are still the most common, habitual and widely used form of visual information representation. Therefore, as with digital media, the important emerging issues for printed documents are copyright protection, authentication, integrity control, tracking and finally quality enhancement. The printed document can be considered as a part of visual communications where data are converted into some specific form suitable for reliable communications via printing channels, with some given constraints on perceptual quality. Digital halftoning is often used for the transformation of digital documents into printed form. Although, being very simple and having been used since 1855, this transformation unavoidably leads to the degradation of the original document quality. This is not the case for digital documents and has important consequences for document security and management.

☆ Parts of this paper appeared as S. Voloshynovskiy, O. Koval, F. Deguillaume and Thierry Pun, Visual communications with side information via distributed printing channels: extended multimedia and security perspectives, in: Proceedings of SPIE Photonics West, Electronic Imaging 2004, Multimedia Processing and Applications, San Jose, USA, January 18–22, 2004.

*Corresponding authors.
E-mail addresses: svolos@cui.unige.ch (S. Voloshynovskiy), Oleksiy.Koval@cui.unige.ch (O. Koval), Frederic.Deguillaume@cui.unige.ch (F. Deguillaume), Thierry.Pun@cui.unige.ch (T. Pun).
URL: http://watermarking.unige.ch/, http://sip.unige.ch.

0165-1684/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2006.11.003

Digital data-hiding [1–4] has been considered as a possible approach to address the above problems for digital documents. However, in most cases the traditional means of digital data-hiding are not always suitable for printed documents and new approaches are required. Moreover, the printing industry has already recognized a limit in the enhancement of printed/scanned image quality related to inverse halftoning. Further enhancement would imply the development of more expensive and complex printing/scanning devices, which is not appropriate for many low-cost and real-time applications. At the same time the security of printed documents (not necessarily printed on paper) including passports, visas, ID cards, driving licenses, diplomas, credit cards, banknotes or checks is usually accomplished by proprietary know-how technologies and digital data-hiding is far to be widely accepted in this field.

In this paper we consider some practical approaches and analyze several scenarios aiming at further extending the usage of distributed coding for printed documents. This problem is tackled from an information-theoretic point of view as communications with side information. Therefore, the goal of the paper is three-fold:

- introduce an information-theoretic formalism into visual communications via distributed printing channels;
- consider possible communications scenarios with side information for multimedia and security applications;
- establish corresponding bounds on system performance assuming side information available at the encoder or at the decoder depending on the particular scenario.

The main practical focus of this paper is on important problem of quality enhancement of documents at the output of print-and-scan channels. One of principle applications were such a problem is playing a crucial role is document authentication and person identification. In such a protocol one should verify if the photo contained in the identification document (passport, ID card, driver license, etc.) is authentic or not. Evidently, in such an application quality of the data at the output of the print-and-scan channel should be sufficiently high to guarantee errorless protocol performance.

As an another possible application one can consider the case of publicly distributed documents such as journals, newspapers, books and advertisement materials. For instance, Fig. 1 represents the basic setup where the image of concern (Lena) should be extracted from the printed document with the purpose of authentication, tracking or quality enhancement. If this image is simply scanned by *a public user* that does not have any access to specific information, the best quality of the resulting scanned image will be defined by the applied processing prior to printing (possibly lossy compression, resampling or blurring), by the quality of the printing and finally by the quality of the scanning (inverse halftoning). It is obvious that in some cases the degradation of quality can be quite considerable so that the scanned image cannot be further exploited for the targeted applications.

Moreover, in some practical scenarios of person identification based on ID documents, the printed document cannot be reproduced in color, if cheap printing is used or when the physics of the reproduction process does not allow to achieve it One example consists in laser engraving of photos on plastic cards used in security printing. In this case, the color information can be communicated via data-hiding or some auxiliary channel exploiting the redundancy between color planes.

A different situation arises when the image contains some embedded data or when these data are communicated via some auxiliary channel that can assist the quality enhancement of the printed document. Assume here for simplicity that this auxiliary channel is a 2-D high capacity bar code shown at the bottom of the page (Fig. 1).[1] The bar code can be designed as a specific pattern or logo to avoid an annoying random-like appearance on the page. The bar code can be also attached to the last page of the journal or can be printed on some supplying layout. Access to the embedded data or to the content of the bar code is authorized and can be managed according to the selected distribution protocol. In the simplest case, it can be granted for an extra price. We call the person who has obtained the above access rights *a private user* or *an authorized user*. The private user can somehow benefit from the appropriately communicated extra information in enhancing the quality of the scanned document. Therefore, the overall goal is to design

[1]Note: More general auxiliary channel can be designed using different information carriers used in the printed (plastic) documents such as bar codes, invisible inks or crystals, magnetic strips, cheap electronic low-capacity chips.

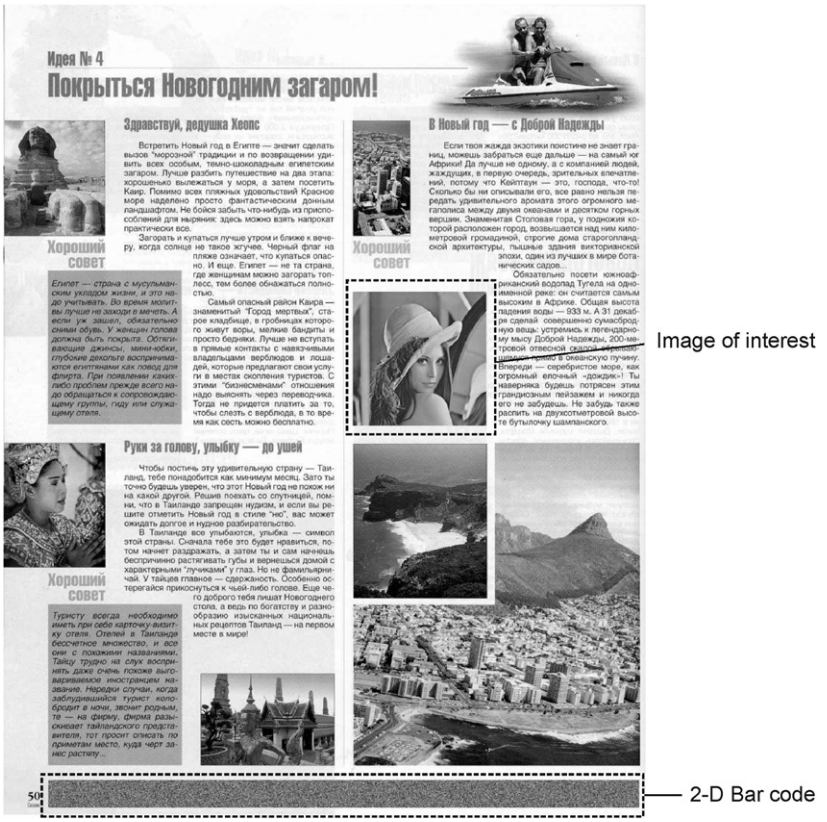

Fig. 1. Visual communications via printing channels: example of journal with side information communications in the form of 2-D bar code.

an optimal system that will provide the maximum possible quality of the reconstructed image with minimum needed rate of the extra information to be additionally communicated to the decoder.

Two alternative solutions to the formulated problem exist that differ by the physical channel structure of the channel conveying the quality update information. The first one is based on the high capacity data-hiding and exploits a self-embedding principle for this purpose [5]. The second one exploits high capacity 2-D bar codes, invisible inks or crystals, magnetic strips, cheap electronic low-capacity chips an auxiliary channel to transfer this update.

Up to our knowledge, only some theoretical results justifying the performance of protocol where information and host interference are simultaneously transmitted over a channel are available [5–7] and none existing practical high capacity data-hiding technique simultaneously satisfies the visibility and payload constraints and provides robustness to the digital-to-analogue-to-digital conversion. That is why in the scope of this paper we will mostly focus our analysis on the latter approach.

The paper is organized as follows. Section 2 outlines the fundamentals of rate-distortion theory with side information. Sections 3 presents a mathematical model of the printing process. Section 4 considers a possible protocol auxiliary channel index communications. Section 5 presents experimental results and Section 6 concludes the paper.

Notations: 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 the realizations of scalar and vector random variables, respectively. The superscript N is used to designate length-N vectors $\mathbf{x} = x^N = [x_1, x_2, \ldots, x_N]^{\mathrm{T}}$ with kth element x_k. 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 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. We use $I(X;Y)$ to denote a mutual information

between two random variables X and Y [8]. Calligraphic fonts $\mathcal{X}$ denote sets $X \in \mathcal{X}$ and $|\mathcal{X}|$ denotes a cardinality of set $\mathcal{X}$. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix.

2. Visual communications with side information

2.1. Asymmetric two-side information

Visual communications via printing channels can be considered from the point of view of communications with an *asymmetric side information* available at the encoder and decoder. The asymmetry of side information originates from the fact that the different types of prior information about the actual image X are available at the encoder and actual side information available a posteriori at the decoder. Let S_1 denotes the side information available at the encoder and S_2 the side information available at the decoder (Fig. 2). One can consider S_1 as a "virtually predicted" image after halftoning/inverse halftoning and S_2 as the physically printed-and-scanned image. Assume that samples of digital image and side information at encoder and at decoder $\{X_k, S_{1,k}, S_{2,k}\} \sim p_{X,S_1,S_2}(x, s_1, s_2)$ are independent drawings of jointly distributed random variables $X \in \mathcal{X}$, $S_1 \in \mathcal{S}_1$ and $S_2 \in \mathcal{S}_2$, where $\mathcal{X}, \mathcal{S}_1, \mathcal{S}_2$ are finite sets. A distortion measure is defined as $d(x, \hat{x})$ and the corresponding average distortion as $D = (1/N)\sum_{k=1}^{N} E[d(X_k, \hat{X}_k)]$. The image $\{X_k\}$ is encoded using a length-N blocks code i (a bar code, a watermark, etc.) into a binary stream of rate R, which should be decoded back as a reconstructed image $\{\hat{X}_k\}$.

Cover and Chiang [9] have shown that the rate R is achievable at the distortion level D, if there exists a sequence of $(2^{NR}, N)$ codes $i : \mathcal{X}^N \times \mathcal{S}_1^N \to \{1, 2, \ldots, 2^{NR}\}$, $\hat{X} : \{1, 2, \ldots, 2^{NR}\} \times \mathcal{S}_2^N \to \hat{\mathcal{X}}^N$ such that the average distortion $E[d(\mathbf{X}, \hat{\mathbf{X}}(i(\mathbf{X}, \mathbf{S}_1), \mathbf{S}_2))] \leqslant D$ and

$$R(D) = \min_{p(u|x,s_1)p(\hat{x}|u,s_2)} [I(U; S_1, X) - I(U; S_2)], \quad (1)$$

where the minimization is performed under the following distortion constraint $\sum_{x,u,s_1,s_2,\hat{x}} d(x, \hat{x}) p(x, s_1, s_2) p(u|x, s_1) p(\hat{x}|u, s_2)) \leqslant D$ and U is an auxiliary random variable. $S_2 \to (X, S_1) \to U$ and $(X, S_1) \to (U, S_2) \to \hat{X}$ form Markov chain.

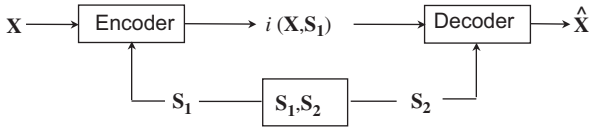

Fig. 2. Two-sided asymmetric side information available at both encoder and decoder.

In the following we will focus on two particular cases of the considered generalized setup from Fig. 2: symmetric side information and asymmetric side information available only at the decoder.

2.2. Symmetric two-side information ($S_1 = S_2 = S$)

To introduce the basic concept of visual communications via printing channels for the broad set of different application scenarios considered in the following sections, we will simplify the previous setup to a case with symmetric side information $S = S_1 = S_2$ that can be available at both encoder and decoder or only at the decoder according to Fig. 3.

The rate distortion with symmetric side information at both encoder and decoder $S = S_1 = S_2$ was first introduced by Gray [10] and further considered by Berger [11] and Flynn and Gray [12] (Fig. 3: dashed line corresponds to availability of S^N at the encoder). The corresponding rate-distortion function is defined as

$$R_{X|S}(D) = \min_{p(\hat{x}|x,s)} I(X; \hat{X}|S), \quad (2)$$

where the minimization is over all conditional distributions $p(\hat{x}|x, s)$ such that $\sum_{x,\hat{x},s} d(x, \hat{x}) p(x, s) p(\hat{x}|x, s) \leqslant D$. The above rate-distortion-function also follows from (1) assuming $S = S_1 = S_2$ [9].

2.3. Asymmetric side information at decoder

The rate-distortion problem with side information available at the decoder was considered by Wyner and Ziv [13]. It also follows from (1), if $S_1 = \emptyset$, $S_2 = S$ and $(S_1, X) = X$:

$$\begin{aligned} R(D)_{X|S}^{WZ} &= \min_{p(u|x)p(\hat{x}|u,s)} [I(U; X) - I(U; S)] \\ &= \min_{p(u|x)p(\hat{x}|u,s)} H(U) - H(U|X) \\ &\quad - H(U) + H(U|S) \\ &= \min_{p(u|x)p(\hat{x}|u,s)} H(U|S) - H(U|X, S) \\ &= \min_{p(u|x)p(\hat{x}|u,s)} I(U; X|S), \end{aligned} \quad (3)$$

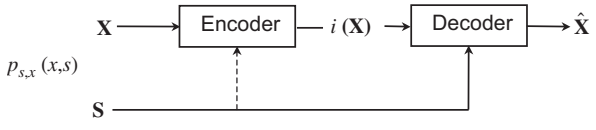

Fig. 3. Symmetric side information $S_1 = S_2 = S$.

where the minimization is performed over all $p(u|x)p(\hat{x}|u,s)$, i.e., $p(x,s,u,\hat{x}) = p(x,s)p(u|x)p(\hat{x}|u,s)$ for fixed $p(x,s)$, and all decoder functions $\hat{X}(i(X),S)$ such that $\sum_{x,\hat{x},u,s} p(x,s)\, p(u|x)p(\hat{x}|u,s)d(x,\hat{x}) \leqslant D$ and $|\mathcal{U}| \leqslant |\mathcal{X}| + 1$. The third equality in (3) follows from the fact that $S \to X \to U$ form a Markov chain and $H(U|X) = H(U|X,S)$, consequently, while the second Markov chain is formed as $X \to (U;S) \to \hat{X}$.

In the general case, $I(\hat{X};X|S) \leqslant I(U;X|S)$ and therefore it can be a *loss* in performance of the Wyner–Ziv setup in comparison to the rate-distortion function with side information available at both encoder and decoder [14].

We will denote the classic rate-distortion function without any side information ($S_1 = \emptyset$, $S_2 = \emptyset$) as

$$R(D) = \min_{p(\hat{x}|x)} I(\hat{X};X), \tag{4}$$

where the minimization is performed over all $p(\hat{x}|x)$ such that $\sum_{x,\hat{x}} p(x)p(\hat{x}|x)d(x,\hat{x}) \leqslant D$.

In the subsequent subsections we will consider the case of a Gaussian formulation of the rate-distortion problem. A methodology allowing to extend the results obtained for the discrete alphabets to continuous formulation can be found in [15].

2.4. Shannon lower bound (SLB): Gaussian source and MSE distortion

The classic rate-distortion function without any side information (4) for a Gaussian $\mathcal{N}(0,\sigma_X^2)$ source with squared error distortion is represented by the SLB [8]:

$$R(D) = \frac{1}{2}\log_2^+ \frac{\sigma_X^2}{D}, \tag{5}$$

where $\log_2^+ x = \max\{0, \log_2(x)\}$. The corresponding rate-distortion function simultaneously describing N independent Gaussian random variables (*a parallel Gaussian source model*) $X_k \sim \mathcal{N}(0,\sigma_{X_k}^2)$, $k = 1,2,\ldots,N$ with a rate R bits/source vector is given by

$$R(D) = \sum_{k=1}^{N} \frac{1}{2}\log_2 \frac{\sigma_{X_k}^2}{D_k}, \tag{6}$$

where

$$D_k = \begin{cases} \lambda & \text{if } \lambda < \sigma_{X_k}^2, \\ \sigma_{X_k}^2 & \text{if } \lambda \geqslant \sigma_{X_k}^2, \end{cases} \tag{7}$$

where λ is chosen to satisfy $\sum_{i=1}^{N} D_k = D$.

2.4.1. Side information at encoder and decoder

Let X be a sample generated by a zero mean and stationary Gaussian memoryless source, i.e., $X \sim \mathcal{N}(0,\sigma_X^2)$. The side information is given by a noisy version of source, i.e., $S = X + Z'$ where X and Z' are independent and $Z' \sim \mathcal{N}(0,\sigma_{Z'}^2)$.

The rate-distortion function (2) in this case will have the same character as for the stationary Gaussian source:

$$R(D)_{X|S} = \frac{1}{2}\log_2^+ \frac{\sigma_{X|S}^2}{D}, \tag{8}$$

where $\sigma_{X|S}^2 = \sigma_X^2(1-\rho^2)$ is conditional variance, $\rho = E[(X-\bar{X})(S-\bar{S})^{\mathrm{T}}]/\sqrt{\mathbf{Var}[X]\mathbf{Var}[S]} = \sigma_X/\sigma_S$ is the correlation coefficient and $\sigma_S^2 = \sigma_X^2 + \sigma_{Z'}^2$, thus, providing $\sigma_{X|S}^2 = \sigma_Z^2\sigma_X^2/(\sigma_Z^2 + \sigma_X^2)$.

Consequently, the rate saving due to the available side information at the encoder and the decoder with respect to the independent coding of X is

$$R(D) - R(D)_{X|S} = \frac{1}{2}\log_2 \frac{1}{1-\rho^2}. \tag{9}$$

If $\rho^2 \to 0$, i.e., the variance of the noise Z' is increasing and X and S become less correlated, $R(D)_{X|S} \to R(D)$ and no gain can be expected from the joint decoding.

In the case of non-stationary Gaussian vector (parallel Gaussian source), $\mathbf{X} = \{X_1, X_2, \ldots, X_N\}$, where $X_k \sim \mathcal{N}(0,\sigma_{X_k}^2)$, we have similarly to (6):

$$R(D)_{X|S} = \sum_{k=1}^{N} \frac{1}{2}\log_2 \frac{\sigma_{X_k}^2\sigma_Z^2}{(\sigma_{X_k}^2 + \sigma_Z^2)D_k}, \tag{10}$$

where

$$D_k = \begin{cases} \lambda & \text{if } \lambda < \dfrac{\sigma_{X_k}^2\sigma_Z^2}{\sigma_{X_k}^2 + \sigma_Z^2}, \\ \dfrac{\sigma_{X_k}^2\sigma_Z^2}{\sigma_{X_k}^2 + \sigma_Z^2} & \text{if } \dfrac{\sigma_{X_k}^2\sigma_Z^2}{\sigma_{X_k}^2 + \sigma_Z^2}, \end{cases} \tag{11}$$

where λ is chosen to satisfy $\sum_{i=1}^{N} D_k = D$.

2.4.2. Side information at decoder

Wyner and Ziv [13,15] demonstrated that for the Gaussian case $R(D)_{X|S}^{WZ} = R(D)_{X|S}$. However, generally $R(D)_{X|S}^{WZ} \geqslant R(D)_{X|S}$ [14]. This means that there is a rate loss with the Wyner–Ziv coding. Also it is known that for discrete sources when $D = 0$, the Wyner–Ziv problem reduces to the Slepian-Wolf problem [16] with $R(0)_{X|S}^{WZ} = R(0)_{X|S} = H(X|S)$ assuming that S is communicated to the decoder at the rate $H(S)$.

2.4.3. Approaching theoretical rate-distortion function with symmetric side-information (Wyner–Ziv vs Gray setups)

For practical reasons, we will perform a comparative analysis of the non-collaborative encoding setup (Wyner–Ziv problem) with collaborative one (Gray problem) for distributed coding in printing applications.

In the general case, the following "inequality sandwich" is valid:

$$R(D) \geqslant R(D)_{X|S}^{WZ} \geqslant R(D)_{X|S}. \tag{12}$$

Therefore, the rate-distortion function $R(D)_{X|S}^{WZ}$ is above the rate-distortion function $R(D)_{X|S}$. In our particular case of distributed coding with side information in printing applications, the side information at the encoder is naturally given by the physical nature of the problem. Therefore, the printing application should definitively benefit from the availability of side information.

2.4.4. Compressed channel index communication

In printed documents, the side information $\mathbf{S}$ is directly available as the printed image (possibly after inverse halftoning). However, the index i should be somehow communicated to the decoder in practical systems. According to the justification, given in Introduction, we will assume that it is communicated via an *auxiliary channel* that can include different information carriers used in the printed (plastic) documents such as bar codes, invisible inks or crystals, magnetic strips, cheap electronic low-capacity chips.

3. Mathematical model of printing

To investigate the "quality" of side information, i.e., to establish how close S is correlated with X from one side and to investigate the possibility of self-embedded communications from another side, one should consider a mathematical model of a printing channel.

3.1. Halftoning and inverse halftoning: image processing perspectives

Digital halftoning is a process that converts a gray-scale image ($X \in \{0, 1, \ldots, 255\}$) into a bi-level image ($X \in \{0, 255\}$) that can easily be reproduced by a simple printing device. There are two types of halftoning, i.e., ordered and random (or so-called error diffusion) halftoning. The ordered halftoning is mostly used in laser printers while error diffusion represents a broad class of ink jet printing devices. We will focus on the later for demonstration purposes. The error diffusion takes the error from the quantized gray-scale pixel to bi-level pixel and diffuses the quantization error over a causal neighborhood The inverse halftoning of error diffusion should ensure that the spatially localized average pixels values of the halftone and original gray-scale image are similar. Kite et al. [17] proposed an accurate linear model for error diffusion halftoning. This model accurately predicts the "blue noise" (high-frequency noise) and edge sharpening effects observed in various error-diffused halftones. The resulting halftone image $\mathbf{s}$ is obtained as

$$\mathbf{s} = H\mathbf{x} + Q\mathbf{w} = H\mathbf{x} + \mathbf{z}', \tag{13}$$

where H represents a linear shift-invariant (LTI) halftone filter, Q is a LTI filter corresponding to the error diffusion coloring, $\mathbf{w}$ is white noise and $\mathbf{z}' = Q\mathbf{w}$ is colored noise.

Therefore, the *inverse halftoning* according to (13) can be posed as a deconvolution or restoration problem (in image processing) or channel equalization problem (in digital communications) where operator H represents the equivalent blurring or channel and $\mathbf{z}'$ corresponds to the image acquisition or channel noise.

The distributed coding system with communications via printing channels is shown in Fig. 4. The system consists of two parts. The first part is the physically printed image $\mathbf{S}$ and the second part is the auxiliary channel dedicated to the communication of index i that can be considered as a bar code. The physical printing channel is represented as a sequential set of operations comprising halftoning, printing and scanning. Here, we assume that printing itself, as the black point reproduction process with a specified resolution, and scanning do not introduce any distortions. In more general case one would need to take into account some optical blurring and aberrations as well as sensor non-linearity.

Considering the halftoning problem according to the vector channel (13), assuming a vector Gaussian setup $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_X)$ and $\mathbf{Z}' \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{Z'})$, the deconvolution problem can be formulated as the MMSE estimation of $\mathbf{x}$ given the conditional mean $\hat{\mathbf{X}}' = E[\mathbf{X}|\mathbf{S}]$:

$$\hat{\mathbf{x}}' = (H^{\mathrm{T}}\mathbf{C}_{Z'}^{-1}H + \mathbf{C}_X^{-1})^{-1}H^{\mathrm{T}}\mathbf{C}_{Z'}^{-1}\mathbf{s} = H_W\mathbf{s}, \tag{14}$$

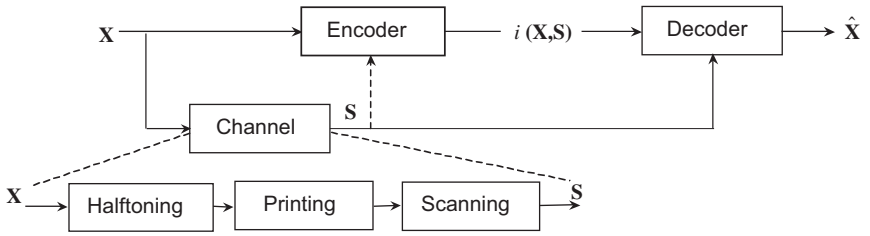

Fig. 4. Printing channel communication setup.

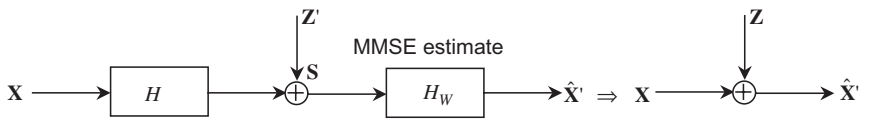

Fig. 5. Equivalent halftoning/inverse halftoning channel.

where $H_W = (H^{\mathrm{T}}\mathbf{C}_{Z'}^{-1}H + \mathbf{C}_X^{-1})^{-1}H^{\mathrm{T}}\mathbf{C}_{Z'}^{-1}$ represents a Wiener "restoration-equalization" filter.

The estimation error $\mathbf{Z} = \hat{\mathbf{X}}' - \mathbf{X} = E[\mathbf{X}|\mathbf{S}] - \mathbf{X}$ is also Gaussian $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_Z)$ due to linear nature of the MMSE estimate where the covariance matrix is defined as

$$\mathbf{C}_Z = \mathbf{C}_{X|S} = E[(\hat{\mathbf{X}}' - \mathbf{X})(\hat{\mathbf{X}}' - \mathbf{X})^{\mathrm{T}}] = (H^{\mathrm{T}}\mathbf{C}_{Z'}^{-1}H + \mathbf{C}_X^{-1})^{-1}. \quad (15)$$

The halftoning/inverse halftoning printing channel can be sequentially represented by an equivalent channel shown in Fig. 5.

3.2. Inverse halftoning: practical low-complexity algorithm

In this section we will extend the results obtained for i.i.d. Gaussian signals to real images. The key issue is the availability of an accurate, simple and tractable stochastic image model leading to a close form solution. Although stochastic inverse halftoning is a stand-alone important problem, we will use here the state-of-the-art inverse halftoning method proposed by Neelamani et al. [18] for wavelet-based inverse halftoning via deconvolution (WInHD). This approach is based on the original idea of Donoho [19] called *wavelet-vaguelette decomposition*. The wavelet-vaguelette decomposition splits the deconvolution problem into two parts. In the first part, the operator H in (13) is inverted leading to the "deblurred" estimate:

$$\tilde{\mathbf{x}} = H^{-1}\mathbf{s} = \mathbf{x} + H^{-1}Q\mathbf{z}' \quad (16)$$

that produces the clean image $\mathbf{x}$ with noisy part $H^{-1}Q\mathbf{z}'$ assuming that the inverse operator H^{-1} exists (that it is not the case in general for all image restoration problems).

The second part of wavelet-vaguelette decomposition tries to remove the noisy part $H^{-1}Q\mathbf{z}'$ using wavelet-based denoising. Therefore, the relevant problem of stochastic image modeling is reduced to the proper modeling of wavelet coefficients. Hjorungnes et al. [20] have proposed a simple and tractable model that has found a number of applications in state-of-the-art image coding (EQ-coder [21]), denoising (wavelet Wiener filter [22]) and watermarking for evaluation of data-hiding capacity [3] and content-adaptive data-hiding [23]. According to the EQ framework, image coefficients are considered coming from a realization of independent Gaussian distribution with a "slowly" varying variance. This represents a class of local Gaussian models and justifies the application of local operations. It should be mentioned that the relationship between the global image statistics and locally Gaussian assumption was elegantly demonstrated based on mixture model where the variance also becomes a random parameter drawn from the exponential distribution [20]. This also justifies a doubly stochastic modeling of image coefficients and provides a logical connection to the parallel Gaussian source model considered in Section 2.4.

The denoiser of WInHD is also based on the same model that leads to the Wiener filter as the MMSE (MAP) estimate for Gaussian case. To enhance the performance of the denoiser, the WInHD

implements the Wiener filter in an overcomplete transform domain (complex wavelets) to overcome the problems of the critically sampled wavelet transform. Thus the final estimate is obtained as

$$\hat{x}'_{j,k} = \frac{\sigma^2_{j,k}}{\sigma^2_{j,k} + \sigma^2_j} \tilde{x}_{j,k}, \tag{17}$$

where j denotes the scale, $\sigma^2_{j,k}$ the local variance of image coefficients for scale j, σ^2_j the variance of the residual noisy term in the transform domain and all variables are considered in the transform domain.

It is not our primary goal to enhance the WInHD, we will use it in its original form. It should, however, be mentioned that more powerful stochastic image models could be used, taking into account the stationarity of transform image coefficients along the edges and transitions. Such models should bring additional enhancement of the inverse halftone image quality and corresponding higher reliability of side information at the decoder [24].

4. System implementation: index communication via auxiliary physical channel

The approach based on the index communication via auxiliary physical channel can be designed either using the Gray setup, i.e., symmetric two-sided information, or the Wyner–Ziv setup, i.e., side information available only at the decoder. The basic system block-diagram for both setups is shown in Fig. 6 with the difference consisting in the availability of side information S^N at the encoder, i.e., Gray setup.

The practical design of the auxiliary channel index communication system includes two main functional blocks, namely source coding for the residual compression and channel coding (2-D bar code) for index communication.

4.1. Source coding

To establish the lower bound for the rate-distortion function we will consider here only the Gray setup assuming that the Wyner–Ziv system will do the best to approach this bound using multidimensional lattice/coset codes. According to the analysis presented in Sections 2.4 and 3, we assume that the side information, which is given in the form of analog (halftone) image is also available at the encoder due to the complete knowledge of the printing process (in other words, we assume that the inverse halftone image produced at the encoder is very accurate approximation of the real scanned image obtained at the decoder). Thus, the encoder should only communicate the difference $\mathbf{Z} = E[\mathbf{X}|\mathbf{S}] - \mathbf{X}$ to the decoder with the given bounded distortion D choosing corresponding index $i(\mathbf{X}, \mathbf{S})$.

From an information-theoretic perspective, an optimal source code $(2^{NR}, N)$ for this setup consists of an encoding map $i : \mathcal{X}^N \times \mathcal{S}^N \to \{1, 2, \ldots, 2^{NR}\}$ and a reconstruction map $\hat{X} : \{1, 2, \ldots, 2^{NR}\} \times \mathcal{S}^N \to \hat{\mathcal{X}}^N$, where both encoder $i(.)$ and decoder $\hat{X}(.)$ are designed based on the state information S^N. In the general case, rate-distortion performance of this encoding scheme is lower bounded by (2) and in the Gaussian setup the conditional rate distortion function is given by (8). The details of a particular practical implementations of the source encoding/decoding are presented in Section 5.

4.2. Channel coding

The task of the channel code is to communicate errorlessly $(\Pr[i \neq \hat{i}(\mathbf{S})] \to 0)$ the index i via some analog auxiliary channel. A natural auxiliary channel for printed documents consists of 2-D bar codes [25], although as was mentioned in the introduction other types of information carriers with sufficient capacity can be used for this purpose.

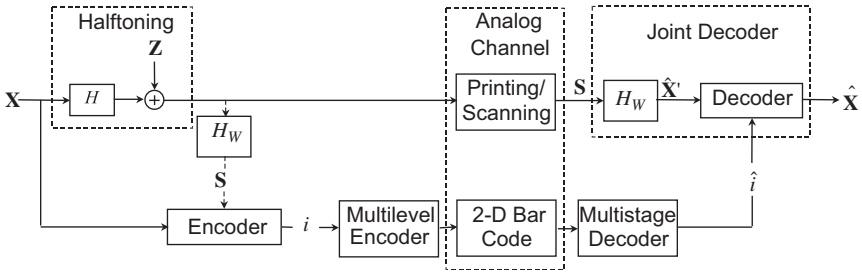

Fig. 6. Index communication via auxiliary (2-D bar code) channel.

In the case of 2-D bar codes, *multilevel codes* (MLC) [26,27] can be used for errorless index communication. The MLC are known to closely approach capacity for band-limited channels The main idea of multilevel coding consists in the protection of each address bit of the signal point by an individual binary code $\mathcal{C}^{\ell}$ at level ℓ. One can use irregular low-density parity-check (LDPC) codes as the binary code $\mathcal{C}^{\ell}$ for each level ℓ followed by 8-PAM chosen for simplicity reasons and easy integration into printed 2-D bar code. One can also apply precoding for bar-codes to avoid intersymbol interference caused by the halftoning filter H. Finally, *multistage decoding* is applied based on multiple access channel decomposition of MLC using chain rule for mutual information between channel output and each level. More details about a particular implementation of the MLC/MSD channel code will be given in the next section.

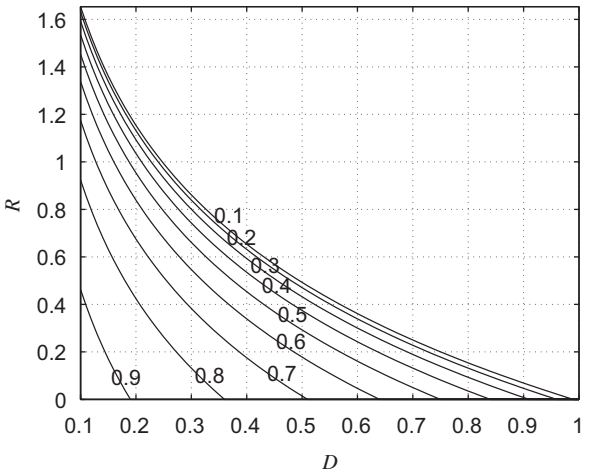

Fig. 7. Conditional rate-distortion function for different ρ and $X \sim \mathcal{N}(0,1)$.

5. Results

First, to investigate the theoretical limits of source coding with side information available at both encoder and decoder, we have modeled the rate-distortion function $R(D)_{X|S}$ for a Gaussian source and the MSE distortion measure for different values of ρ^2 (Fig. 7), where ρ is a correlation coefficient. The highest gain with respect to the classic coding without side information can be expected for highly correlated sources X and S (reliable side information) when $\rho^2 \to 1$. If $\rho^2 \to 0$, there is no gain with respect to $R(D)$. Therefore, the printing channel (and particularly a joint halftoning/inverse halftoning) should provide an accurate estimate of the source X, i.e., to reduce the impact of the halftoning process. In real applications, ρ^2 is expected to be in the range of $0.2 \leqslant \rho^2 \leqslant 0.6$ and therefore for the stationary Gaussian source the maximum gain can be as much as 0.26–0.66 bit per sample.

Secondly, we have performed the comparison of different coding strategies for a number of real images based on Floyd halftoning and wavelet-based inverse haftoning via deconvolution [18]. Here we only report results for the images Lena and Baboon of size of 512×512. The "*lower empirical bound*" is estimated for the Gray setup described in Section 4. The prediction of the inverse halftone images $\mathbf{s}$ based on the MMSE estimate ($\hat{\mathbf{X}}' = E[\mathbf{X}|\mathbf{S}]$) is performed at the encoder and the residual $\mathbf{Z} = \hat{\mathbf{X}}' - \mathbf{X}$ is compressed at various rates and communicated to the decoder via 2-D bar code. At the decoder, the MMSE estimate $\mathbf{S}$ is computed from the printed image and combined with the decoded residual $Q(\mathbf{Z})$.

Three different lossy coders are used to compress the residual $Q(\mathbf{Z})$, i.e., the DCT-based JPEG, wavelet-based WaveConvert [28] and wavelet-based EQ coder. The last two codecs are wavelet transform-based. In the case of WaveConvert [29] the encoding algorithm is a progressive zero-tree-based wavelet encoder exploiting progressive quantization followed by the adaptive arithmetic entropy coder.

The EQ lossy image codec [21] exploits an intraband correlations among coefficients in the high-frequency wavelet subbands. The main underlying idea of this method is that the samples in these subbands are distributed according to locally Gaussian distribution with slowly varying variance that might be locally estimated using a maximum likelihood strategy based on the causal neighborhood and potentially a parent sample located in the same spatial position.

At the end of the estimation step, the coefficients are quantized using a uniform threshold quantizer selected accordingly to the results of the rate-distortion optimization that is followed by an optimal adaptive arithmetic coding.

While at the high-rate regime the approximation of the local variance by its quantized version is valid, it is not the case at low rates. The reason for that is the quantization to zero most of the data samples that makes local variance estimation extremely inaccurate. The simple solution proposed

in [21] consists in the placement of all the coefficients that fall into the quantizer deadzone in the so-called unpredictable class, and the rest in the so-called predictable class. The samples of the first one are considered to be distributed globally as an i.i.d. generalized Gaussian distribution, while the infinite Gaussian mixture model is used to capture the statistics of the samples in the second one. This separation is performed using a simple rate-dependent thresholding operation. The parameters of the unpredictable class are exploited in the rate-distortion optimization and are sent to the decoder as side information. The experimental results presented in [12,19] allow to conclude about the state-of-the-art performance of this technique in the image compression application.

In this paper we used the following parameters for the EQ algorithm: 4 level DWT based on $\frac{9}{7}$ biorthogonal wavelet filter pair with local variance ML estimate in a causal window of size 3×3 without considering interscale dependencies).

The performed experiments for all mentioned encoders were targeting the following compression rates: 0.20, 0.25, 0.50 and 1.00 bpp. The obtained results are presented in Figs. 8 and 9. Moreover, for comparison reasons, we indicate in Tables 1 and 2 the results of direct compression by the above methods assuming that one can directly

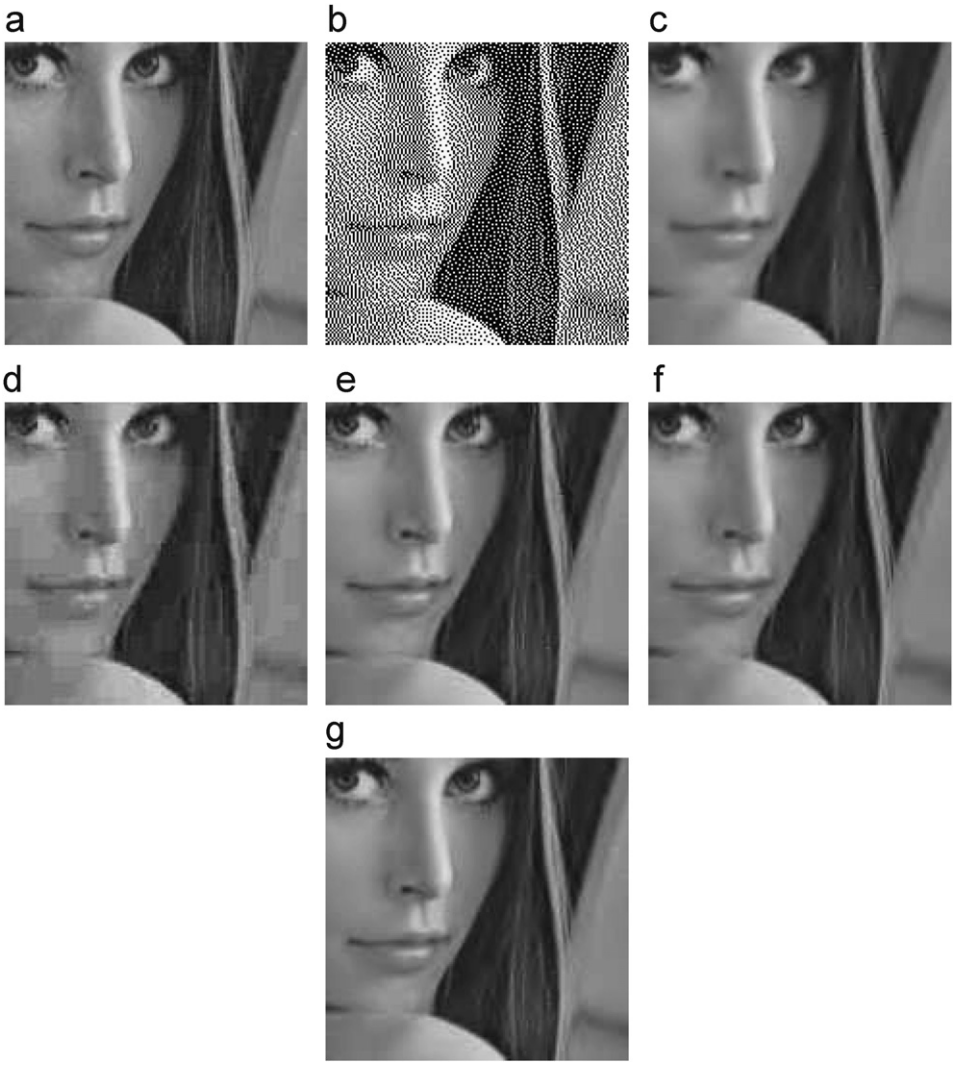

Fig. 8. Results of computer modeling for the frame of Lena image: (a) original image Lena; (b) halftone image; (c) WInHD inverse halftone image; (d) directly JPEG compressed image (rate 0.2 bpp); (e) JPEG compressed image with side information (rate 0.2 bpp); (f) directly EQ compressed image (rate 0.2 bpp) and (g) EQ compressed image with side information (rate 0.2 bpp).

communicate the compressed image via 2-D bar code to the decoder without usage of side information for final image reconstruction. A possible gain due to the decoding with the side information in the form of a halftone image is also shown in the table. The PSNR of the inverse halftone images (with respect to the original images) was 31.95 dB for the case of image Lena and 23.73 dB for image Baboon, respectively. The maximum gain is obtained for the EQ coder, and is in the range of 0.26–2.87 dB for different rates with respect to the directly compressed image. Obviously, the capacity of the auxiliary channel will be a factor that determines a particular operational rate depending on the application constraints. The enhancement of "analog" (halftone) image quality is also varying for different rates between 2–8 dB, which indicates a considerable potential of such an "upgrade" for many practical applications where the image quality should be relatively high. Using advanced methods of inverse halftoning and more accurate stochastic image models one can expect a further enhancement of image quality for low-bit rates.

One important question that concerns the proposed bar code-based solution to the problem of quality enhancement of printed-and-scanned documents consists in the physical size of a bar code conveying quality update information. Accordingly to the recently obtained results [30], the storage capacity of modern such storage modules is more than

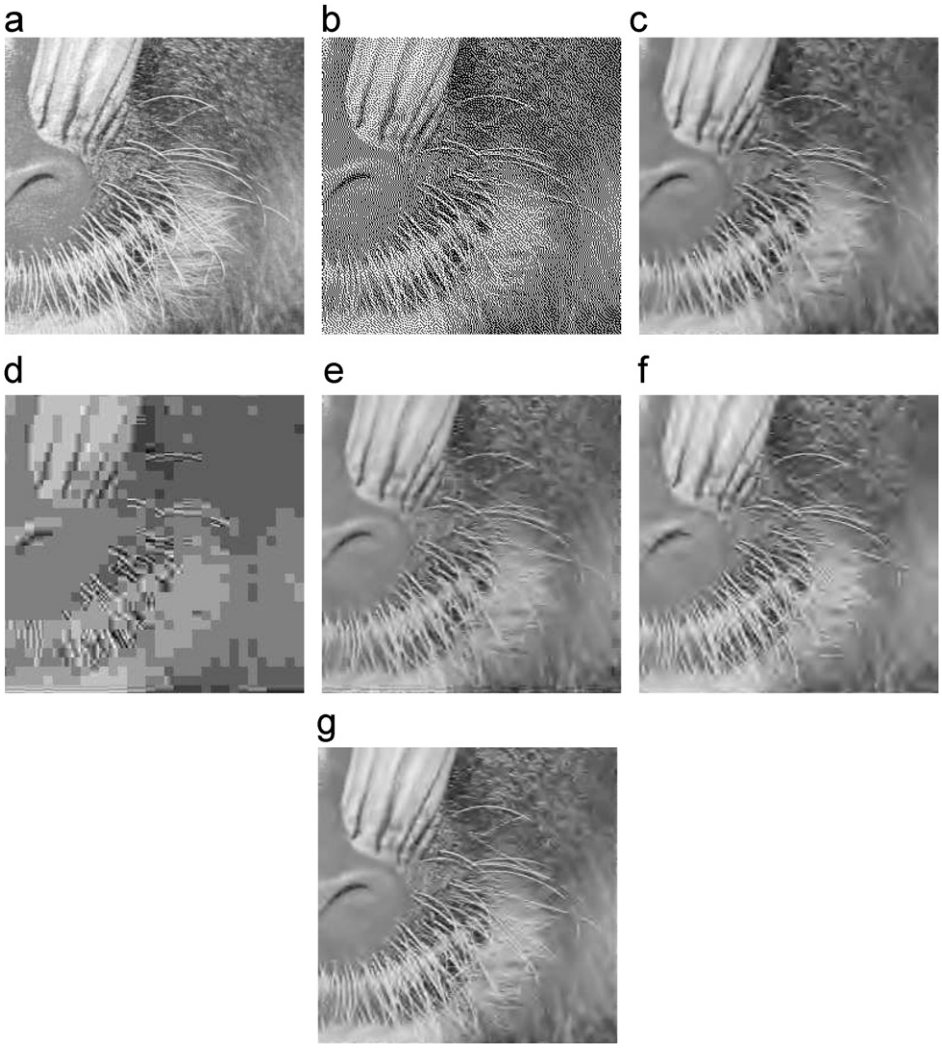

Fig. 9. Results of computer modeling for the frame of Baboon image: (a) original image Baboon; (b) halftone image; (c) WInHD inverse halftone image; (d) directly JPEG compressed image (rate 0.2 bpp); (e) JPEG compressed image with side information (rate 0.2 bpp); (f) directly EQ compressed image (rate 0.2 bpp) and (g) EQ compressed image with side information (rate 0.2 bpp).

Table 1
Comparison of resulting PSNR [dB] for several compression methods for image "Lena"

Compression method	Rate (bpp)			
	0.20	0.25	0.50	1.00
DCT-based JPEG coder				
Direct compression	28.89	30.41	34.64	37.83
Compression with SI	33.55	34.17	36.19	38.61
Gain over direct compression	4.66	3.76	1.55	0.78
Gain over direct halftoning WInHD	1.60	2.22	4.24	6.66
WaveConvert coder				
Direct compression	32.47	33.48	36.59	39.66
Compression with SI	35.11	35.39	37.77	40.03
Gain over direct compression	2.64	1.91	1.18	0.37
Gain over direct halftoning WInHD	3.16	3.44	5.82	8.08
EQ-based coder				
Direct compression	32.87	34.11	37.20	40.40
Compression with SI	35.42	35.90	38.24	40.76
Gain over direct compression	2.55	1.79	1.04	0.36
Gain over direct halftoning WInHD	3.47	3.95	6.29	8.81

Table 2
Comparison of resulting PSNR [dB] for several compression methods for test image "Baboon"

Compression method	Rate (bpp)			
	0.20	0.25	0.50	1.00
DCT-based JPEG coder				
Direct compression	19.42	21.51	23.67	26.44
Compression with SI	23.61	24.39	25.59	27.66
Gain over direct compression	4.19	2.88	1.92	1.22
Gain over direct halftoning WInHD	−0.12	0.66	1.86	3.93
WaveConvert coder				
Direct compression	22.19	22.70	24.97	28.19
Compression with SI	25.35	25.71	27.25	30.32
Gain over direct compression	3.16	3.01	2.28	2.13
Gain over direct halftoning WInHD	1.62	1.98	3.52	6.59
EQ-based coder				
Direct compression	22.90	23.32	25.84	29.60
Compression with SI	25.77	26.19	28.00	31.04
Gain over direct compression	2.87	2.87	2.16	1.44
Gain over direct halftoning WInHD	2.04	2.46	4.27	7.31

2000 bytes/in^2. In particular, our implementation of the 2-D bar codes based on 15 level non-equidistant 8 pulse amplitude modulation MLC/MSD schemes using the LDPC codes allows to reliably store and extract 2400 bytes/in^2 of information. Thus, it is easy to compute, for instance, that in the case of compression rate equal to 0.2 bits/element a bar code will occupy approximately 17.6 cm^2. Supposing that a printed document should be reproduced on the paper of A4 format, a possible solution might consist in the placement of a stripe containing compressed information $1 \times 17.6\,\text{cm}^2$ in the top/bottom of the page.

6. Conclusions

We have considered the problem of visual communications with side information via distributed printing channels. In particular, we have analyzed the possibility of side information-assisted image processing in printing applications. We have also provided a mathematical modeling of the printing channel in the scope of digital communications with side information. The related problem of source coding based on Wyner–Ziv and Gray setups has been considered. Corresponding practical scenario has been analyzed and finally the theoretical bounds on systems performance are given. The results of computer simulation show a range of possible applications where the expected gain can considerably increase the possibilities and functionalities of modern multimedia and security systems.

Acknowledgments

This paper was partially supported by the Swiss National Science Foundation Professorship Grant No. PP002-68653/1, by the European Commission through the IST Programme under contract IST-2002-507932-ECRYPT, European project FP6-507609-SIMILAR and Swiss IM2 projects. The authors are thankful to Fernando Perez-Gonzalez for many helpful and interesting discussions on different subjects of lattice coding in personal communications and during his visit to University of Geneva. The authors are also thankful to the members of Stochastic Image Processing group (University of Geneva) for many stimulating discussions during seminars.

The information in this document reflects only the authors views, is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

References

[1] I.J. Cox, M.L. Miller, J.A. Bloom, Digital Watermarking, Morgan Kaufmann Publishers, Academic Press, Los Altos, CA, New York, 2002.

[2] F. Deguillaume, S. Voloshynovskiy, T. Pun, Hybrid robust watermarking resistant against copy attack, in: Eleventh European Signal Processing Conference (EUSIPCO'2002), Toulouse, France, September 3–6, 2002.
[3] P. Moulin, M.K. Mihcak, A framework for evaluating the data-hiding capacity of image sources, IEEE Trans. Image Process. 11 (9) (September 2002) 1029–1042.
[4] F. Perez-González, F. Balado, J.R. Hernández, Performance analysis of existing and new methods for data hiding with known-host information in additive channels, IEEE Trans. Signal Process. (Special Issue on Signal Processing for Data Hiding in Digital Media and Secure Content Delivery) 51(4) (April 2003).
[5] E. Martinian, G. Wornell, B. Chen, Authentication with distortion criteria, IEEE Trans. Inform. Theory (July 2005) 2523–2542.
[6] T. Kalker, F.M. Willems, Coding theorems for reversible embedding, in: DIMACS Workshop on Network Information Theory, vol. 66, Rutgers University, Piscataway, NJ, March 2003.
[7] A. Sutivong, M. Chiang, T. Cover, Y.-H. Kim, Channel capacity and state estimation for state-dependent gaussian channels, IEEE Trans. Inform. Theory 51 (4) (April 2005).
[8] T. Cover, J. Thomas, Elements of Information Theory, Wiley, New York, 1991.
[9] T. Cover, M. Chiang, Duality of channel capacity and rate distortion with two sided state information, IEEE Trans. Inform. Theory (2002) 48(6).
[10] R. Gray, A new class of lower bounds on information rates of stationary sources via conditional rate-distortion function, Inform. Control 19 (7) (July 1973) 480–489.
[11] T. Berger, The information theory approach to communications, in: G. Longo (Ed.), Multiterminal Source Coding, Springer, Berlin, 1978.
[12] T.J. Flynn, R.M. Gray, Encoding of correlated observations, IEEE Trans. Inform. Theory 33 (11) (November 1987) 773–787.
[13] A. Wyner, J. Ziv, The rate-distortion function for source coding with side information at the decoder, IEEE Trans. Inform. Theory 22 (1) (January 1976) 1–10.
[14] R. Zamir, The rate loss in the Wyner–Ziv problem, IEEE Trans. Inform. Theory 19 (November 1996) 2073–2084.
[15] A. Wyner, The rate-distortion function for source coding with side information at the decoder-II: general sources, Inform. Control 38 (1978) 60–80.
[16] D. Slepian, J.K. Wolf, Noiseless encoding of correlated information sources, IEEE Trans. Inform. Theory 19 (July 1973) 471–480.
[17] T.D. Kite, B.L. Evans, A.C. Bovik, T.L. Sculley, Modeling and quality assessment of halftoning by error diffusion, IEEE Trans. Image Process. 9 (6) (May 2000) 909–922.
[18] R. Neelamani, R. Nowak, R.G. Baraniuk, WInHD: wavelet-based inverse halftoning via deconvolution, submitted to IEEE Trans. Image Process. (October 2002), submitted for publication.
[19] D. Donoho, Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition, Appl. Comput. Harmonic Anal. 2 (1995) 101–126.
[20] A. Hjorungnes, J. Lervik, T. Ramstad, Entropy coding of composite sources modeled by infinite gaussian mixture distributions, in: IEEE Digital Signal Processing Workshop, 20–24 January, 1996, pp. 235–238.
[21] S. LoPresto, K. Ramchandran, M. Orhard, Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework, in: Data Compression Conference 97, Snowbird, Utah, USA, 1997, pp. 221–230.
[22] M.K. Mihcak, I. Kozintsev, K. Ramchandran, P. Moulin, Low-complexity image denoising based on statistical modeling of wavelet coefficients, IEEE Signal Process. Lett. 6 (12) (December 1999) 300–303.
[23] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, T. Pun, A stochastic approach to content adaptive digital image watermarking, in: Third International Workshop on Information Hiding, vol. 1768, Dresden, Germany, September 29–October 1, 1999, pp. 212–236.
[24] S. Voloshynovskiy, O. Koval, T. Pun, Wavelet-based image denoising using non-stationary stochastic geometrical image priors, in: IS&T/SPIE's Annual Symposium, Electronic Imaging 2003: Image and Video Communications and Processing V, Santa Clara, CA, USA, 20–24 January 2003.
[25] N. Degara-Quintela, F. Perez-Gonzalez, Visible encryption: using paper as a secure channel, in: Security and Watermarking of Multimedia Contents V, vol. 5020, Santa Clara, CA, USA, 20–24 January 2003.
[26] J. Huber, U. Wachsmann, R. Fischer, Coded modulation by multilevelcodes: overview and state of the art, in: ITG–Fachbericht: Codierung fur Quelle, Kanal und Ubertragung, Aachen, Germany, March 1998, pp. 255–266.
[27] H. Imai, S. Hirakawa, A new multilevel coding method using error-correcting codes, IEEE Trans. Inform. Theory 3 (23) (1995) 371–377.
[28] T. Strutz, in: WaveConvert, 2001, ⟨http://www-nt.e-technik.uni-rostock.de/~ts/software/download.html⟩.
[29] E. Müller, T. Strutz, Scalable wavelet-based coding of color images, in: Fourth International Conference on Actual Problems of Electronic Instrument Engineering (APEIE'98), Nowosibirsk, Russia, September 1998, pp. 29–35.
[30] R. Villan, S. Voloshynovskiy, O. Koval, T. Pun, Multilevel 2d bar codes: Towards high capacity storage modules for multimedia security and management, IEEE Trans. Inform. Forensics Secur. (2006), accepted for publication.