

Image Adaptive High Volume Data Hiding Based on Scalar Quantization

N. Jacobsen, K. Solanki, U. Madhow, B. S. Manjunath, and S. Chandrasekaran
 Dept. of Electrical and Computer Engineering
 University of California at Santa Barbara
 Santa Barbara, CA 93106

Email: {jacobsen, solanki, madhow, manj, shiv}@ece.ucsb.edu

Abstract—

Information-theoretic analysis for data hiding prescribe embedding the hidden data in the choice of quantizer for the host data. In this paper, we consider a suboptimal implementation of this prescription, with a view to hiding high volumes of data in images with low perceptual degradation. The three main findings are as follows:

- (i) **Scalar quantization based data hiding schemes incur about 2 dB penalty from the optimal embedding strategy, which involves vector quantization of the host.**
- (ii) **In order to limit perceivable distortion while hiding large amounts of data, hiding schemes must use local perceptual criteria in addition to information-theoretic guidelines.**
- (iii) **Powerful erasure and error correcting codes provide a flexible framework that allows the data-hider freedom of choice of where to embed without requiring synchronization between encoder and decoder.**

I. INTRODUCTION

Our objective is to obtain techniques for embedding high volumes of data in images, in a manner that causes minimal perceptual degradation, and is robust to “benign” JPEG compression as well as more malicious Additive White Gaussian Noise (AWGN) attacks. Information-theoretic treatments [1], [2] of the data hiding problem typically focus on hiding in independent and identically distributed (i.i.d.) Gaussian host samples. The hider is allowed to induce a mean squared error of at most D_1 , while an attacker operating on the host with the hidden data is allowed to induce a mean squared error of at most D_2 . Information-theoretic prescriptions in this context translate, roughly speaking, to hiding data by means of the choice of vector quantizer for the host data, with the AWGN attack being the worst-case under certain assumptions.

The main contributions of the present paper are as follows. First, we motivate the use of scalar quantization based methods by providing an information-theoretic analysis that shows that scalar quantization based strategies come within about 2 dB (in terms of the ratio of D_1/D_2 required to attain a given hiding capacity) of the optimal vector quantization based strategy for AWGN attacks. Such strategies, in which each host symbol is quantized using a scalar quantizer, whose choice is dictated by the message symbol, were first proposed by Chen and Wornell [3], but the capacity penalty due to the suboptimal embedding was not previously known. Next, we propose practical strate-

gies for embedding high volumes of data in images with low perceptual degradation. As in many recent papers, the information is embedded in the low and mid frequency bands of the Discrete Cosine Transform (DCT) coefficients of the image, which have higher variance, and can therefore better tolerate the additional distortion due to the embedding. The novelty of our approach lies, however, in that it allows the flexibility of embedding data in designated areas of the image where it would be less noticeable, without requiring any additional side information to be sent to the decoder regarding the location of the embedded message.

Our embedding approach is as follows. Both the encoder and decoder know the set of DCT coefficients in the low and mid frequency bands that are possible candidates for embedding. In addition, the encoder uses local criteria to decide which subset of the coefficients it will actually embed data in. The decoder employs the same criteria to guess where the encoder has embedded data. The distortion due to attacks may now lead to insertion errors (the decoder guessing that a coefficient has embedded data, when it actually does not) and deletion errors (the decoder guessing that a coefficient does not have embedded data, when it actually does). However, unlike the insertion-deletion channel in which the length of the overall symbol sequence can vary [4], in our situation, the set of candidate coefficients for embedding is the same, and is known to both encoder and decoder. This enables an elegant solution to the problem of synchronizing the encoder and decoder. Use a code on the hidden data that spans the entire set of candidate embedding coefficients, and that can correct both errors and erasures. The subset of these coefficients in which the encoder does not embed can be treated as *erasures at the encoder*. The decoder guesses the erasure locations using the same criteria as the encoder. Insertions now become errors, and deletions become additional erasures. Thus, the use of a suitable error correction code simultaneously yields robustness to attack and solves the synchronization problem between encoder and decoder.

The above coding framework is applied to two embedding strategies that we proposed in earlier work on uncoded image-adaptive embedding [5]. The first is the block-level Entropy Thresholding (ET) method, which decides whether or not to embed data in each 8×8 block of DCT coefficients, depending on the entropy, or energy, within that block. The second is the Selectively Embedding in

Coefficients (SEC) method, which decides whether or not to embed data based on the magnitude of the coefficient. Reed-Solomon codes are a natural choice for the block-based ET scheme, while turbo-like codes are employed for the SEC scheme. We are able to hide high volumes of data under both JPEG and AWGN attacks. Due to the suboptimality of our embedding strategy and the decoding algorithm, our results are not directly comparable with information-theoretic estimates of the hiding capacity of the image [6]. However, the amount of hidden data is of the same order of magnitude as these estimates.

II. QUANTIZATION BASED DATA HIDING

A. Embedding data in choice of quantizer

As in [3], data is embedded in the host medium through the choice of scalar quantizer. For example, consider a uniform quantizer of step size Δ , used on the host's coefficients in some transform domain. Let odd reconstruction points represent a hidden data bit '1'. Likewise, even multiples of Δ are used to embed '0'. Thus, depending on the bit value to be embedded, one of two uniform quantizers of step size 2Δ is chosen.

Hard decision decoding in this context is performed by quantizing the received coefficient to the nearest reconstruction point of all quantizers. An even reconstruction point decodes to a hidden '0' and an odd point to a '1'. If more information regarding the statistics of the attack is available, soft decisions can be used to further improve performance. In Section II-B, we compute the capacity of scalar quantization based hiding for the specific case of AWGN attacks. Implicit in our formulation is the use of soft decisions that account for both the quantization noise and the AWGN.

B. Capacity of scalar quantization based data hiding

While our scalar quantization based scheme is well matched to JPEG compression attacks, we now show that it incurs roughly a 2 dB penalty for the worst-case AWGN attack. Letting D_1 and D_2 denote the mean squared embedding induced distortion and mean squared attack distortion, the hiding capacity with AWGN attack is given by [1], [2] $C = \frac{1}{2} \log(1 + \frac{D_1}{D_2})$, in the small D_1, D_2 regime that typical data hiding systems operate. We compare this "vector capacity" (termed thus because the optimal strategy involves vector quantization of the host) to the mutual information attained by a scalar quantizer embedding scheme with soft decision decoding.

Consider a data hiding system where the information symbol to be embedded is taken from an alphabet \mathcal{X} . The host's original uniform quantizer is divided into M uniform sub-quantizers (each with quantization interval $M\Delta$), where $M = |\mathcal{X}|$, a power of two. Thus, we have that $\log_2 M$ bits are hidden per host symbol.

We consider the distortion-compensated quantization embedding scheme of [3] with soft decision decoding. Here, the uniform quantizer is scaled by $\alpha \in (0, 1)$, increasing the quantization interval to Δ/α . As such, the robustness is in-

creased by a factor $1/\alpha^2$ (in the squared minimum distance sense), and embedding induced distortion is increased by the same factor. Encoding the information symbol as a linear combination of the host symbol and its quantized value, as in the following, compensates for the additional distortion. Denoting the host symbol by H , and the hidden message symbol by X , the decoder receives the following symbol:

$$\begin{aligned} Y &= \alpha q_X(H) + (1 - \alpha)H + W \\ &= q_X(H) + (1 - \alpha)(H - q_X(H)) + W \end{aligned} \quad (1)$$

where $q_x(\cdot)$ the scaled uniform quantizer used to embed the information symbol x (with quantization interval $M\Delta/\alpha$), and W the AWGN attack (i.e. $W \sim \mathcal{N}(0, D_2)$). The parameter α achieves a tradeoff between uniform quantization noise and AWGN. The optimal value for α for maximizing the signal-to-noise ratio (SNR) at the decoder, which we have found numerically also to maximize the mutual information $I(X; Y)$, is [3]

$$\alpha_{opt} = \frac{D_1}{D_1 + D_2}$$

The probability density function of the combined additive interferers, $N = (1 - \alpha)Z + W$, where $Z \equiv H - q_X(H)$ is the uniform quantization noise, is given by convolving the uniform and Gaussian densities:

$$f_N(x) = \frac{\alpha(2\pi D_2)^{-\frac{1}{2}}}{(1 - \alpha)M\Delta} \int_{-\frac{(1-\alpha)M\Delta}{2\alpha}}^{\frac{(1-\alpha)M\Delta}{2\alpha}} \exp\left(-\frac{(x - \tau)^2}{2D_2}\right) d\tau \quad (2)$$

We compute the mutual information $I(X; Y) = H(X) - H(X|Y)$ for X uniform over its M -ary alphabet as an estimate of the capacity with scalar quantization based embedding. Thus, $H(X) = \log_2 M$. To find, $H(X|Y)$, we now compute $p_{X|Y}$, the conditional probability mass function of X given Y , and f_Y , the probability density function of Y .

Consider the quantization interval in which the received symbol Y appears, and define its midpoint as the origin. Letting y denote the abscissa, the nearest quantizers appear at $y = \pm \frac{\Delta}{2\alpha}$. Conditioned on the input $X = x$ and host coefficient $H = h$, the distribution of Y is given by $f_{Y|X, H}(y|x, h) = f_N(y - m_x \frac{\Delta}{2\alpha} - k_h \frac{M\Delta}{\alpha})$, with f_N as in (2). Here, $m_x \in \mathcal{M} = \{\pm 1, \pm 3, \dots, \pm 2M - 1\}$ is uniquely determined by the information symbol x , $k_h \in \mathbb{Z}$ by the host coefficient h , and the hidden quantized host coefficient $q_x(h)$ by the pair (m_x, k_h) . Thus we have

$$\begin{aligned} f_{Y|X}(y|x) &= \int_{\mathcal{H}} f_{Y|X, H}(y|x, h) f_H(h) dh \\ &\propto \sum_{k \in \mathbb{Z}} f_N\left(y - m_x \frac{\Delta}{2\alpha} - k \frac{M\Delta}{\alpha}\right) \end{aligned} \quad (3)$$

$$\begin{aligned} f_Y(y) &= \sum_{x \in \mathcal{X}} f_{Y|X}(y|x) p_X(x) \\ &\propto \sum_{m \in \mathcal{M}} \sum_{k \in \mathbb{Z}} f_N\left(y - m \frac{\Delta}{2\alpha} - k \frac{M\Delta}{\alpha}\right) \end{aligned} \quad (4)$$

where we have assumed that the host H and message X are statistically independent, and that the host's density f_H is roughly constant on an interval around Y , an assumption that is reasonable in the low distortion regime, where the quantization interval is small with respect to variations in the host's distribution.

Since the density of Y is $\frac{\Delta}{\alpha}$ -periodic, it suffices to restrict attention to the interval $[-\frac{\Delta}{2\alpha}, \frac{\Delta}{2\alpha}]$, with f_Y normalized accordingly. Applying Bayes' theorem, the distribution of X given Y is

$$p_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)p_X(x)}{f_Y(y)} \quad (5)$$

so that we can now compute

$$H(X|Y) = \int_Y \sum_{x \in \mathcal{X}} p_{X|Y}(x|y) \log p_{X|Y}(x|y) f_Y(y) dy$$

and hence $I(X; Y)$.

Due to the exponential decay of the Gaussian density, the summation in (3) is well approximated with only the $k = 0$ term (α near one), i.e. the nearest quantization point explaining a transmitted x . Figure 1 plots the mutual information obtained with 2, 4 and 8-ary signaling, as well as vector capacity. We observe roughly a 2 dB loss due to the suboptimal scalar quantization encoding strategy.

We conclude our analysis by noting that the soft decision statistic, used by an iterative decoder, is the log likelihood ratio (LLR), given in the following for the case of binary signaling.

$$\Lambda(y) = \log \frac{p_{X|Y}(0|y)}{p_{X|Y}(1|y)} = \log \frac{f_{Y|X}(y|0)}{f_{Y|X}(y|1)} \quad (6)$$

When $\alpha = 1$ and (3) approximated as above, the LLR reduces to

$$\Lambda(y) = \log \frac{f_W(y - \frac{\Delta}{2})}{f_W(y + \frac{\Delta}{2})} = \frac{y\Delta}{D_2} \quad (7)$$

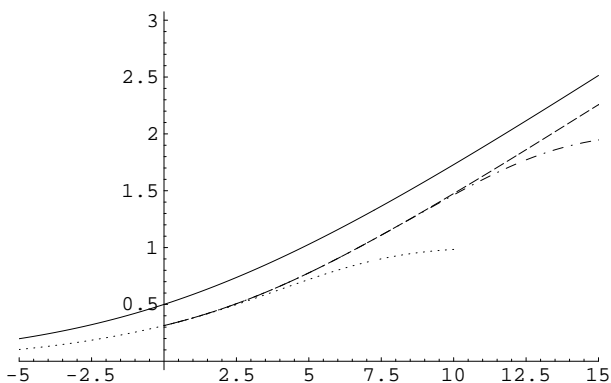


Fig. 1. Gap between scalar [binary: dot, 4-ary: dash-dot, 8-ary: dash] and vector [solid] quantizer data hiding systems. Capacity is in bits per source symbol, DNR in dB.

III. IMAGE ADAPTIVE DATA HIDING

Most prior image-based data hiding schemes use global criteria regarding where to hide the data, such as statistical criteria independent of the image (e.g. embedding in mid-frequency bands), or criteria matched to a particular image (e.g. embedding in high-variance bands). These are consistent with information theoretic guidelines [6], which call for hiding in “channels” in which the host coefficients have high variance. However, when hiding large volume of data, hiding based on such global statistical criteria can lead to significant perceptual degradation. We present examples of such effects at [7]. This motivates the novel hiding framework presented here, which affords the flexibility of adapting to local characteristics within a particular image, thus significantly reducing the perceptual degradation due to the hidden data. While we use specific criteria for such local adaptation, our coding framework, which obviates the need for synchronization between encoder and decoder, applies for any other scheme for locally adaptive hiding.

We now describe and extend two image-adaptive hiding techniques, which we had first proposed for uncoded hidden data in [5].

A. Entropy Thresholding (ET) scheme

The entropy thresholding scheme uses the energy (or 2-norm entropy) of an 8×8 block to decide whether to embed in the block or not. The entropy is computed as follows.

$$E = \sum_{i,j} \|x_{ij}\|^2, \quad (i,j) \neq (0,0) \quad (8)$$

Only those blocks whose entropy exceeds a predetermined threshold are used to hide data. Likewise, the decoder checks the entropy of each 8×8 block to decide whether data has been hidden.

The embedding procedure is based on JPEG. The image is divided into 8×8 non-overlapping blocks, the energy of the blocks is evaluated, and an 8×8 DCT of the qualifying blocks is taken. The coefficient blocks are divided by the JPEG quantization matrix according to the design quality factor, and finally, bits are hidden using choice of scalar quantizer. Low frequency coefficients are used to embed in qualifying blocks. Embedding in these coefficients induces minimal distortion due to JPEG's finer quantization in this range. Thus, this scheme employs a statistical criterion by hiding in the frequency subbands of large variance, while satisfying a local perceptual criterion via the block entropy threshold.

B. Selectively Embedding in Coefficients (SEC) scheme

Now, instead of deciding where to embed at the block level, we do a coefficient-by-coefficient selection, with the goal of embedding in those coefficients that cause minimal perceptual distortion.

Here too, an 8×8 DCT of non-overlapping blocks is taken and the coefficients are divided by the JPEG quantization matrix at design quality factor. These coefficient values

are quantized to the nearest integer for further processing. We embed in a given coefficient only if the magnitude exceeds a positive integer threshold. As this threshold increases, fewer coefficients qualify for embedding, and hence less data can be hidden, which provides a tradeoff between hiding rate and perceptual quality. For thresholds beyond 2, it gets difficult for a human observer to distinguish between the original and composite image, while achieving embedding rates of ~ 0.014 bits/pixel (about 3500 bits in 512×512 image) with perfect decoding.

The simplest SEC scheme is the zero-threshold SEC scheme, where the coefficients that are not quantized to zero, are used to embed information. In the uncoded version of the scheme [5], quantization embedding is performed as usual when embedding a ‘1’. If a ‘0’ is to be embedded, the coefficients are quantized to even reconstruction values. However, if this results in quantization to zero, we leave it as such and the same ‘0’ is embedded in the next coefficient satisfying the non-zero criterion. The decoder simply disregards all coefficients that quantize to zero. Otherwise, decoding is performed as usual. Such a procedure (of embedding same ‘0’ in next coefficient) becomes unnecessary when SEC scheme is used with a coding framework, wherein a symbol is erased at the encoder if it does not meet the threshold criteria (Section IV).

High embedding rates are achieved using this zero-threshold SEC scheme with very low perceptual degradation. To understand this intuitively, it should be noted that there are many image coefficients that are very close to zero once divided by the JPEG quantization matrix, and would be quantized to zero upon JPEG compression. Embedding ‘1’ in such coefficients introduces a large amount of distortion relative to the original coefficient size, a factor that seems to be perceptually important. This is avoided by choosing not to use zeros for embedding.

IV. CODING FOR INSERTIONS AND DELETIONS

In the previous section, it was noted that use of image adaptive criteria is necessary when hiding high volumes of data into images. As discussed in Section I, any embedding scheme that employs local perceptual criterion could suffer from insertion/deletion problems, causing the catastrophic loss of synchronization between encoder and decoder. In the ET scheme, insertions and deletions are observed when the attack quality factor is mismatched with the design quality factor for JPEG attack. In the SEC scheme, very few images have insertions/deletions under matched or mismatched JPEG attacks. However, both the schemes have insertions/deletions under AWGN attacks.

A. Coding Framework

The bit stream to be hidden is coded, using a low rate code, assuming that all host coefficients that meet the global criterion will actually be employed for hiding. A code symbol is erased at the encoder if the local perceptual criterion for the block or coefficient is not met. Since we code over entire space, long codewords can be constructed to achieve very good correction ability. A min-

imum distance separable (MDS) code, like Reed Solomon (RS) code, does not incur any penalty for erasures at the encoder. Turbo-like codes, which operate very close to capacity, incur only a minor overhead due to erasures at the encoder. It should be noted that a deletion, which causes an erasure, is half as costly as an insertion, which causes an error. Hence, it is desirable that the data-hiding scheme be adjusted in such a manner that there are very few (or no) insertions.

Thus, using a good erasure correcting code, one can deal with insertions/deletions without significant decline in original embedding rate. Reed Solomon codes [8] have been used for ET scheme and Repeat Accumulate codes [9] have been used for the SEC scheme as described in following sections.

B. Reed-Solomon (RS) coding for ET scheme

Reed Solomon codes [8] are MDS codes, such that any k coordinates of an (n, k) RS code can be treated as message positions. In other words, the k message symbols can be decoded even if any k symbols out of total n symbols sent are received at the decoder. The code can correct $(n-k)$ erasures, or half as many errors. The block length n of a Reed-Solomon code must be smaller than the symbol alphabet. Thus, the Reed-Solomon code operates over large nonbinary alphabets (each symbol can be interpreted as a block of bits), which matches the block-based ET scheme, where an entire block gets inserted or deleted. The code symbols should be interleaved in such a way that at least k of the n code symbols of a codeword are received at the decoder with high probability, under the erasures-only assumption. For example, if one whole codeword were placed in a smooth area of the image, all or most of the symbols would be erased, and it would be impossible to decode this particular codeword at the receiver.

Reed-Solomon codes are not well matched to AWGN channels, but are ideal for the purpose of illustrating how to deal with the erasures caused by application of local criteria at the encoder and decoder. We now turn to the SEC scheme, where we consider powerful binary codes that are well-matched to AWGN channels, as well as close to optimal for dealing with erasures.

C. Repeat-accumulate (RA) coding for SEC scheme

Any turbo-like code that operates close to Shannon limit could be used with SEC scheme. We used RA codes [9] in our experiments because of their simplicity of decoding and near-capacity performance for erasure channels [10]. A rate $1/q$ RA encoder involves q -fold repetition, pseudorandom interleaving and accumulation of the resultant bit-stream. Decoding is performed iteratively using the sum-product algorithm [11].

The zero-threshold SEC scheme was used for embedding where data was hidden in the coefficients lying within a designated low frequency band that meet the threshold criteria. The size of this band determines the length of RA code. Typically, it consists of 14 to 35 coefficients (out of a total of 64, scanned in zig-zag fashion as in JPEG) de-

pending on the design quality factor used (e.g. 20 for QF = 25 and 35 for QF = 75).

V. RESULTS

The use of perceptual criterion enables us to hide large volume of data without significant perceptual degradation. Due to the poor reproduction of images on paper, the reader is referred to our home page [7] for various examples of images with hidden data. We also compare images with same amount of data hidden using local and statistical criteria in [7] so as to demonstrate that hiding techniques must use local criteria in order to hide large volume of data in a robust and transparent manner.

A. JPEG attacks

The decoder estimates the location of the embedded data, and uses hard decisions on the embedded bits in these locations. The bits in the remaining locations (out of the set of candidate frequencies) are set to erasures. Since the embedding procedure of both ET and SEC scheme is tuned to JPEG, the decoding of embedded data is perfect for all the attacks lesser than or equal to the design quality factor (QF). Table I shows the number of bits embedded (with perfect recovery) in uncoded and coded ET and SEC schemes at various design QFs, under JPEG attacks for 512×512 Lena image.

QF	attack compr. (bpp)	ET scheme # of bits		SEC scheme # of bits	
		uncoded	coded	uncoded	coded
25	0.42	6,240	4,608	11,117	7,168
50	0.66	15,652	12,096	19,346	13,824
75	1.04	34,880	30,560	32,117	23,893

TABLE I

PERFORMANCE OF CODED AND UNCODED ET AND SEC SCHEMES UNDER JPEG ATTACKS AT VARIOUS QUALITY FACTORS

B. AWGN attacks

Table II summarizes the results for the ET scheme with RS coding and SEC scheme with RA coding against AWGN attack. The number of bits embedded is listed for the 512×512 Lena image and noise is added to the image in spatial domain. Although the RS code is not the best choice for AWGN, it is adequate for mild attacks. RA-coded SEC scheme uses soft decision statistic of the AWGN for decoding (as in (7) in Section II-B), and hence performs better at higher attack powers.

VI. CONCLUSIONS

The information theoretic analysis of scalar quantization based hiding reveals only about a 2 dB penalty from the optimal embedding scheme. Furthermore, Figure 1 indicates that binary signaling is sufficient to achieve capacity in the low DNR region our embedding schemes operate in. A by-product of the information-theoretic analysis is

Attack power (dB)	# of bits	
	ET Scheme	SEC Scheme
10.0	7,040	7,447
12.5	6,528	6,826
15.0	3,584	6,301

TABLE II

PERFORMANCE OF ET SCHEME WITH RS CODING AND SEC SCHEME WITH RA CODING UNDER AWGN ATTACK, DESIGN QF=25

the log likelihood ratio for AWGN attacks, which can be employed to generate soft information to be passed to the outer decoder. In previous work on iterative decoding for data hiding in a Gaussian host, the soft decisions were generated by approximating the sum of the quantization noise and AWGN as Gaussian [12].

A significant practical contribution of the paper is a powerful coding framework that allows flexibility in terms of the application of local criteria for determining where to embed data. The use of errors and erasures codes enables an elegant solution to the synchronization problem between the encoder and decoder which arises due to locally adaptive embedding. We are able to hide large volumes of data, of the same order as the capacity estimates in [6]. However, a careful comparison with such information-theoretic estimates is not the focus of the present paper, and will be undertaken in future work after further refinement of the encoding and decoding strategies.

REFERENCES

- [1] M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [2] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," Preprint, Jan. 2001.
- [3] B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Transaction on Information Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.
- [4] M. C. Davey and D. J. C. Mackay, "Reliable communication over channels with insertions, deletions, and substitutions," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 687–698, Feb. 2001.
- [5] K. Solanki, N. Jacobsen, S. Chandrasekaran, U. Madhow, and B. S. Manjunath, "High-volume data hiding in images: Introducing perceptual criteria into quantization based embedding," in *Proceedings of the IEEE Intl. Conference on Acoustics Speech and Signal Processing (ICASSP)*, May 2002.
- [6] P. Moulin and M. K. Mihcak, "The data-hiding capacity of image sources," Preprint, June 2001.
- [7] "<http://vision.ece.ucsb.edu>," UCSB Vision Research Laboratory Homepage.
- [8] S. B. Wicker and V. K. Bhargava, *Reed-Solomon Codes and Their Applications*, IEEE Press, 1994.
- [9] H. Jin, D. Divsalar, and R. J. McEliece, "Coding theorems for turbo-like codes," in *36th Allerton Conference on Communications, Control, and Computing*, Sept. 1998, pp. 201–210.
- [10] H. Jin, A. Khandekar, and R. J. McEliece, "Irregular repeat-accumulate codes," in *Proceedings 2nd International Symposium on Turbo codes and Related Topics*, Sept. 2000, pp. 1–8.
- [11] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [12] M. Kesal, M. K. Mihcak, R. Koetter, and P. Moulin, "Iteratively decodable codes for watermarking applications," in *Proc. 2nd Intl. Symp. on Turbo Codes and Related Topics*, Sept. 2000.