On the Duality between Distributed Source Coding and Data Hiding *

Jim Chou S. Sandeep Pradhan Kannan Ramchandran

Department of Electrical Engineering and Computer Science
269 Cory Hall
University of California - Berkeley, Berkeley, CA 94708
jimchou, pradhan5, kannanr@eecs.berkeley.edu

Abstract

Inspired by a recently proposed constructive framework for the distributed source coding problem [10], we propose a powerful constructive approach to the watermarking problem, emphasizing the dual roles of "source codes" and "channel codes." In our framework, we explore various source and channel codes to achieve watermarks that are robust to attackers in terms of maximizing the distortion between the corrupted coded-source signal and the original signal while holding the distortion between the coded-source signal and the original signal constant. Simulations show that performance gains are accredited to a careful choice of source and channel codes.

1. Introduction

Digital watermarking (data hiding) is an emerging research area that has received a considerable amount of attention in recent years. The basic idea behind digital watermarking is to embed information into a signal such that the amount of distortion incurred on the signal is minimal. The motivation behind embedding information into a signal, is if the embedded information can be reliably recovered, then this information can specify the affiliation between the signal and its original owner; thus the information must be embedded in a manner that will preclude others from destroying it easily. Methods for embedding watermarks are wide and varied; popular methods range from modulating the information onto the least significant bits [1] to using the information as a key for indexing pseudo-random noise sequences which are additively combined with the signal [6]. Irrespective of the method, the main goal that each method has in common is to embed the maximum amount of information possible given a fixed distortion constraint between the signal and the watermarked signal, while allowing for reliable recovery of the embedded information subject to a fixed-distortion attack. Recent research has focused on the achievable capacity of various watermarking systems [3, 2, 9]. In particular, [3] made the connection between the capacity of the watermarking problem and the capacity of the problems considered in [7, 4, 8]. In doing so, a practical system was devised in [3] that achieves a higher capacity than previous systems based on low-bits-modulation and spread spectrum techniques.

A dual problem is source coding with side information (about the source) at the receiver. In this problem, encoding of a source X is considered given that the decoder has access to some side information Y. Many engineering problems such as sensor networks, distributed data base systems and interactive communication systems can be addressed using the distributed source coding paradigm. Information theoretic analysis of this problem has been considered in [11, 12]. Recently, practical and structured methods of developing codes for this problem was proposed [10] as Distributed source coding using syndromes (DISCUS). In this paper we analyze the duality of watermarking and distributed source coding problems. We then construct a solution to the watermarking problem based on DISCUS principles and the insightful theoretical results presented in [4, 7]. We will provide experimental results based on gaussian sources and real images with additive white gaussian noise (AWGN) attacks to demonstrate the efficacy of this approach.

2 Duality of DISCUS and watermarking

In this section we explore the duality between the distributed source coding problem and the watermarking problem.

*Sponsored in part by: NSF MIP 97-03181 (CAREER). ONR under Grant No. N00014-97-1-0864

2.1 Distributed source coding

Source coding with side information is shown schematically in Fig. 1. In this problem, the source X is to be encoded and transmitted to a receiver which has access to some side information Y. Even though the encoder does not have access to Y, using joint statistics of X and Y, the encoder can perform as well as the case when both the encoder and decoder have access to Y [11]. Let us consider the following illustrative example.

Example 1: Discrete Case: Consider X and Y to be equiprobable 3-bit data sets which are correlated in the following way: $d_H(X,Y) \leq 1$, where $d_H(.,.)$ denotes Hamming distance. When Y is known both at the encoder and decoder, we can compress X to 2 bits, conveying the information about the uncertainty of X given Y (i.e., the modulo-two sum of X and Y given by: (000),(001),(010) and (100)). Now if Y is known only at the decoder, we can surprisingly still compress X to 2 bits. The method of construction stems from the following argument: if the decoder knows that X=000 or X=111, then it is wasteful to spend any bits to differentiate between the two. In fact, we can group X=000 and X=111 into one coset (it is exactly the principal coset of the length-3 repetition code). In a similar fashion, we can partition the remaining space of 3-bit binary codewords into 3 different cosets with each coset containing the original codewords offset by a unique and correctable error pattern. Since there are 4 cosets, we only need to spend 2 bits to specify the coset in which X belongs. The four cosets are given as

coset-1 $= (000, 111)$, coset-2 $= (001, 110)$,
coset-3 $= (010, 101)$, coset-4 $= (011, 110)$

The decoder can recover X perfectly by decoding Y to the closest (in hamming distance) codeword in the coset specified by the encoder. Thus the encoder does not need to know Y for optimum encoding.

Figure 1. Distributed source coding: only decoder has access to the side information Y.

The above concepts can be generalized to other cases including the encoding/decoding of continuous-valued random variables, where the decoder uses Y to reconstruct the source based on a minimum fidelity criterion. The minimum rate of encoding [12] for a given fidelity criterion D, is

$$R(D) = min_{\hat{X}=f(U,Y),p(U|X)}[I(U;X) - I(U;Y)] \quad (1)$$

where U is the set of codewords representing X, $I(U;X)$ is the Shannon mutual information [5], and the minimization is carried out over all conditional probability density functions $p(U|X)$ and a function $f(U;Y)$ such that $E(X-\hat{X})^2 \leq D$. The main idea of encoding is as follows: (1) build a source code to represent X using nearly $2^{nI(U;X)}$ codewords, (2) partition this set into $2^{nR(D)}$ cosets with each coset containing $2^{nI(U;Y)}$ codewords (the set of codewords act as a channel code for the fictitious channel between U and Y), and (3) find the optimal reconstruction, which is obtained as a function $f(U,Y)$.

To elucidate the encoding construction, consider the case where X is an independent identically distributed ($i.i.d.$) gaussian random variable with side information Y given by $Y = X + N$ (N is an $i.i.d.$ Gaussian random variable independent of X). The side information is therefore a noisy version of X and as in the discrete case, it has been shown [12] that an encoder can be designed to represent X as well as the case where the encoder also has access to Y. The practical method proposed in [10] is to design a source code and partition it into a bank of cosets of channel codes. The source code is designed for the optimal representation of X, and is partitioned into cosets which have good distance properties. The source X is quantized to a codeword (referred to as the active codeword) and the index of the coset containing this codeword is sent to the decoder. With the help of Y, the decoder finds the active codeword in the coset whose index is given by the encoder. A practical example is as follows:

Example 2: Continuous Case: Consider an 8-level fixed-length scalar quantizer as the source coder (see Fig. 2), with the rate of transmission fixed at 1 bit/source sample. To partition the source code, we partition the reconstruction levels into two cosets. The reconstruction levels in coset 1 are denoted as (r_0, r_2, r_4, r_6) and the reconstruction levels in coset 2 are (r_1, r_3, r_5, r_7) (see Fig. 2). The partition is constructed in such a way that the cosets are symmetric and

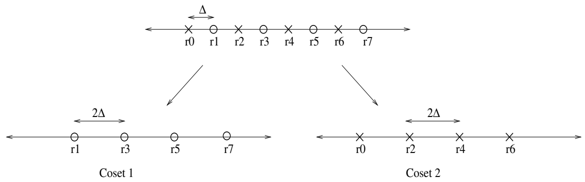

Figure 2. An 8-level quantizer partitioned into 2 cosets containing 4 levels each.

the minimum distance between any two elements within a coset is kept as large as possible. Upon encoding, X is quantized to a codeword in the composite 8-level quantizer and the index of the coset containing this codeword is sent to the decoder. The decoder finds this codeword in the coset whose index is sent by the encoder, as the one which is closest (in the appropriate distance measure) to Y. In doing so, the decoder can reconstruct the value of X to a codeword which is "close" in terms of squared-error distortion.

In the above example, a simple 8-level scalar quantizer was considered for the sake of simplicity. Encoding constructions, however, can be generalized to more sophisticated source coders. For example, the partition can be done in higher-dimensional spaces using trellis codes. This was proposed in [10] where the n-dimensional product space of scalar quantized reconstruction levels is partitioned into cosets of trellis coded quantizer codebooks.

2.2 Watermarking problem

The digital watermarking problem can be formulated as follows. The encoder has access to two signals; the information (an index set), M, to be embedded, and the signal that the information is to be embedded in. The output of the encoder is the watermarked signal. The attacker will attempt to degrade the signal so that the decoder (who wants to authenticate the watermarked signal) will fail to decode the watermark. The attacker has a distortion constraint on the amount of degradation that he can inflict on the signal. The encoder has to be designed such that the attacker needs to inflict an amount of distortion (to destroy the watermark) that will render the signal to be useless. Mathematically, the goal is to solve the following constrained minimization problem:

$$min_{||W-S||^2 \leq D_1, ||Y-W||^2 \leq D_2} P_e(\hat{M}) \qquad (2)$$

where $P_e(\hat{M})$ represents the probability of decoding error.

It was shown in [3, 9] that this problem can be viewed as channel coding with side information (about the channel) at the encoder (see fig. 3). We will consider the case when the

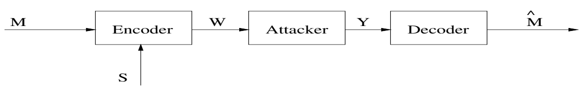

Figure 3. The watermarking problem: the encoder has access to a signal S. M represents the information to be embedded and W represents the watermarked signal.

side information is jointly gaussian with the noise source. It is obvious that this problem becomes the regular AWGN communications problem when both the encoder and decoder have access to the side information. Surprisingly, in the case when only the encoder has access to the side information, it has been shown that this system can *perform as well as when both the encoder and decoder have access to the side information [4].* In fact, this is the key insight which motivates our analysis for the digital watermarking system through exploitation of its dual nature with the distributed source coding problem. Before formulating the details for constructing an encoder/decoder for the watermarking problem, we first look at an analogous problem that will illustrate key concepts.

The problem of writing on defective memories [8]is very similar to the watermarking problem. In this problem the encoder attempts to write data onto faulty registers with one of the registers always being "stuck-at" 0 or 1. The encoder, however, has full knowledge of which register is always "stuck". If the decoder also has this knowledge, then $n-1$ bits can be stored reliably in n registers with one "stuck-at" fault. It was shown in [8] that the encoder can be designed to perform as well as the case when both the encoder and the decoder have access to the "stuck-at" position. We illustrate the method for constructing such an encoder in the following example[1].

Example3: Faulty Registers: Consider a 3-bit memory device, which has one "stuck-at" in any position with a uniform probability of the "stuck-at" being either a one or a zero. When both the encoder and the decoder know the value and the position of the "stuck-at", the encoder can write 2 bits reliably. Now consider the case when only the encoder has access to the "stuck-at". Let $\mathcal{A}$ denote the the set of binary three tuples: $\{0,1\}^3$. We partition $\mathcal{A}$ into cosets of codewords which are compatible with any type of "stuck-at". This partition is the same as the one considered in Example 1. If the encoder wants to write the first message, he chooses a codeword from the first coset which is compatible with the "stuck-at". If the first bit has a "stuck-at" fault of 1, then he chooses [1 1 1] as the codeword to be written on the memory. In this fashion, we can again reliably write two bits (corresponding to the index of the coset), even though the decoder does not have access to the "stuck-at".

The above problem can be generalized into the problem of channel coding with side information, for which the capacity can be calculated. The capacity [8, 7, 4] of such systems is given by

$$C = max_{p(U,S|X)}[I(U;Y) - I(U;S)] \qquad (3)$$

[1] We use the term "stuck-at" loosely to refer to a faulty memory register which is constrained to always be a 1 or a 0.

where the maximization is over all conditional probability density functions $p(U, X|S)$. The signal S is the side information about the channel and U represents the codeword space. The main steps for encoding should be as follows: (1) build a channel code over the space of U, with the number of codewords being nearly equal to $2^{nI(U;Y)}$ (where n is the block length of encoding), (2) partition this channel codeword space into cosets of source codes with each coset containing nearly $2^{nI(U;S)}$ codewords, and (3) choose a codeword, U, to represent S from the coset which has an index equal to the message. The size of the index set should be nearly equal to 2^{nC}, where C is given in (3). The signal, X, which is transmitted over the channel is a function $f(U, S)$.

Returning to the watermarking problem, we consider a related problem for which previous work has already been done. In particular, we consider the case where the channel is AWGN (see Fig. 3) and the signal (side information, S) is $i.i.d.$ gaussian. In this case, it was shown by [4] that the capacity (3) is given as

$$C = \frac{1}{2} log \left\{ \frac{P}{N} + 1 \right\} \tag{4}$$

where P and N represent the transmitter power constraint and the variance of the channel noise respectively. To achieve capacity, [4] showed that the space of codewords must be of the form $U = X + \alpha S$, where $\alpha = \frac{P}{P+N}$ and N is the variance of the attacker's noise. The codeword U should be chosen in a way that will ensure that the projection of U along the direction orthogonal to S has power P.

The watermarking problem can be posed as a version of the above problem, where the signal (i.e., image, audio, etc.) is given by S, the attack is AWGN[2] (with a distortion constraint of N) and the distortion constraint between the signal and the watermarked signal is given as P. In doing so, the above procedure is reinterpreted as follows: the codeword U is chosen to be a quantized representation of some scaled version of S such that the quantization noise is orthogonal to U (see Fig. 4). This scaling factor can be obtained by the geometrical visualization of the encoding process suggested by [4]. Using geometry, the scaling factor is solved to be

$$\beta S = U + V \tag{5}$$

where V denotes the quantization noise and $\beta > 1$ is such that U is the "ideal" rate-distortion-quantized version of βS with $\beta = \frac{P+\alpha^2 Q}{\alpha Q}$. As a result, we have a constructive approach for encoding watermarks into a given signal. The encoding process is summarized as follows: (1) pick a channel code over the space of U, (2) partition the channel code into cosets of source codewords with each codeword representing a quantized version of βS and (3) choose the index of the coset in which the signal is to be quantized based on the watermark that is to be encoded into the signal.

[2]It has been shown [9] that AWGN is the optimal attack given that the signal S is gaussian

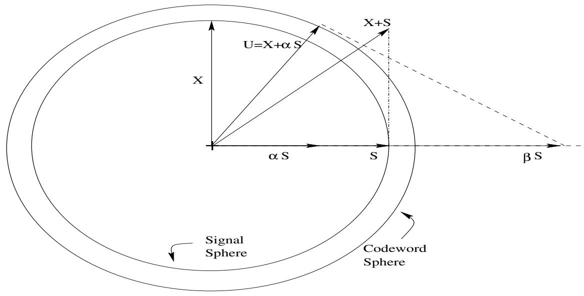

Figure 4. Geometric interpretation of the optimal encoding: U and S denote the codeword chosen for encoding and the signal to be watermarked respectively. This can also be interpreted as choosing that U which is the quantized version of βS as shown. The watermark added onto the signal S is X.

Upon decoding, the decoder finds a codeword in the composite channel code (containing nearly $2^{nI(U;Y)}$) which is closest to the received vector (in some sense). The coset containing the decoded codeword is declared as the decoded watermark message. It can be shown [4] that the scaling which is done at the encoder provides the required distance property to tolerate the attacker's noise at the decoder.

3. Design and construction: Simulation results

To test the effectiveness of our above encoding and decoding approach, we have devised a couple of practical solutions based on different source and channel codes. In one approach we use trellis-coded-modulation (TCM) as our composite channel code, and partition the channel code into source codewords given by an 8-level Lloyd-Max quantizer; we will refer to this method as TCM-SQ. In the other approach we again use TCM as the composite channel code, but partition the channel code into source codewords given by a trellis-coded-quantizer (TCQ); we refer to this method as TCM-TCQ.

The first set of simulations that we conducted, assumed that the signal was gaussian and that the attack was AWGN. We fixed the probability of decoding error to be less than 10^{-5} and determined the amount of distortion that was necessary to ensure this probability of error given a fixed-distortion attack. The performance curves representing the signal-to-distortion (i.e., signal power vs. watermark power) ratio vs. signal-to-noise (i.e., signal power vs. noise power) at $P_e(\hat{M}) = 10^{-5}$ and a rate of 1 embedded-bit/sample is given in Fig. 5. It can be seen from the figure

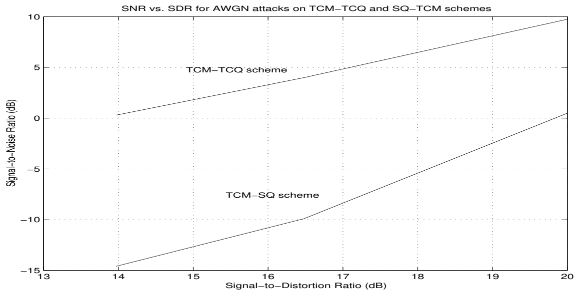

Figure 5. Performance curves of TCM-SQ and TCM-TCQ given fixed distortion attacks.

that TCM-TCQ outperforms TCM-SQ significantly. The reason for this performance discrepancy is that TCM-TCQ uses a better source codebook than TCM-SQ.

The next simulation that we conducted is of a more practical use. We consider the case where S is the 'Lena' image and the attack is confined to JPEG compression. We again tested both TCM-SQ and TCM-TCQ. At an embedding rate of $\frac{1}{64}$ bits/sample, TCM-SQ incurred a Peak-Signal-to-Distortion-Ratio (PSDR) of 41.76dB in order to withstand a JPEG compression factor of 40%. On the other hand, TCM-TCQ incurred a PSDR of 42.22dB in order to withstand a JPEG compression quality factor of 25%. Again TCM-TCQ outperforms TCM-SQ. In Fig. 6 we show the degraded JPEG (Q=25%) image. The probability of decoding error was less than 10^{-5}.

Figure 6. Degraded Lena image (JPEG compressed, Q=25%). With an embedding rate of 1/64 bps, all bits were recovered successfully.

4. Conclusion

In conclusion, we have shown that the watermarking problem and the distributed source coding problem can be viewed as the duals of each other, where the former is channel coding with side information at the transmitter and the latter is source coding with side information at the receiver. In the first case, the channel code is partitioned into a bank of cosets of source codes, while in the latter the source code is partitioned into a bank of cosets of channel codes. The 3-bit binary data example clearly demonstrates the dual nature of the two problems. Through simulation results, it can be seen that this method performs fairly well and is easily amenable to codes built for higher-dimensions [10].

Acknowledgements
We thank Pierre Moulin for insightful discussions.

References

[1] J. M. Barton. Method and apparatus for embedding authentication information within digital data. *United States Patent #5,646,997*, Issued July 8 1997.

[2] C. Cachin. An information-theoretic model for steganography. *Proc 1998 Workshop on Information Hiding*, 1998.

[3] B. Chen and G. W. Wornell. An information-theoretic approach to the design of robust digital watermarking systems. *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, March 1999.

[4] M. Costa. Writing on dirty paper. *IEEE Trans. on Information Theory*, 29:439–441, May 1983.

[5] T. M. Cover and J. A. Thomas. *Elements of Information theory*. Wiley, New York, 1991.

[6] I. Cox, J. Killian, T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Trans. on Image Processing*, 6(12):1673–1687, December 1997.

[7] S. Gel'fand and M. Pinsker. Coding for channel with random parameters. *Problems of Control and Information Theory*, 9:19–31, 1980.

[8] C. Heegard and A. El Gamal. On the capacity of computer memory with defects. *IEEE Trans. on Information Theory*, 29:731–739, September 1983.

[9] P. Moulin and J. O'Sullivan. Information-theoretic analysis of information hiding. *Preprint*.

[10] S. S. Pradhan and K. Ramchandran. Distributed source coding using syndromes: Design and construction. *Proceedings of the Data Comression Conference (DCC)*, March 1999.

[11] D. Slepian and J. K. Wolf. Noiseless encoding of correlated information sources. *IEEE Trans. on Inform. Theory*, IT-19:471–480, July 1973.

[12] A. D. Wyner and J. Ziv. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. on Inform. Theory*, IT-22:1–10, January 1976.