Implications for High Capacity Data Hiding in the Presence of Lossy Compression

Deepa Kundur
10 King's College Road
Department of Electrical and Computer Engineering
University of Toronto
Toronto, Ontario, Canada M5S 3G4
deepa@comm.toronto.edu

Abstract

We derive capacity bounds for watermarking and data hiding in the presence of JND perceptual coding for a class of techniques that do not suffer from host signal interference. By modeling the lossy compression distortions on the hidden data using non-Gaussian statistics, we demonstrate that binary antipodal channel codes achieve capacity. It is shown that the data hiding capacity is at most equal to the loss in storage efficiency bit rate if watermarking and quantization for lossy compression occur in the same domain.

1 Introduction

Data hiding is the general process by which a discrete information stream is hidden within a multimedia signal by imposing imperceptible changes on the host signal. The hidden information is often publicly retrievable, binary and contains data helpful for information interpretation. The problem has gained interest in applications involving captioning, maintenance of audit trails, and embedding hyperlink information in "hyper-media." In contrast to the related research problem of robust digital watermarking, the threat of intentional removal of the hidden data is small as such attacks are only for the purpose of vandalism. For data hiding the main source of concern is the effects of lossy compression on the hidden information. Such manipulation is necessary to facilitate efficient information transfer and storage. The main design challenge is to maximize the capacity of the embedded data in the face of lossy compression.

1.1 Previous Work

Initial research into data hiding in multimedia signals concentrated on the design of sophisticated embedding strategies to improve robustness against typical signal distortions. More recent work has focused in particular on assessing the effects of perceptual coding on the embedded data [3, 8, 9, 23]. *Perceptual coding* refers to the *lossy* compression of multimedia signal data using human perceptual models.

A duality exists between the problems of perceptual coding and data hiding; the former problem attempts to remove irrelevant and redundant information from a signal, while the latter uses the irrelevant information to mask the presence of the hidden data. Thus, the objectives of the two mechanisms are somewhat at odds. As a result, several papers have dealt with integrating perceptual coding with data hiding [2, 5, 14, 15, 16, 22]. Most research has dealt with the design of new algorithms to incorporate or combat lossy compression. Newer work has also analytically studied data hiding in the presence of compression to derive new insights [13, 19].

There is, however, a need for more comprehensive work which investigates not only the robustness, but the capacity of different data hiding techniques in the presence of lossy compression. This paper attempts to provide some new insights and implications along these lines.

1.2 Contributions of this Paper

The objectives of this work is to evaluate and assess the potential of high capacity data hiding in the presence of perceptual coding. In particular,

1. we derive capacity bounds for data hiding in the presence of perceptual coding using the just noticeable difference (JND) perceptual paradigm. Unlike previous work [1, 7, 18], we model the distortions on the hidden data using non-Gaussian statistics.

2. we demonstrate that antipodal signals opposed to ran-

dom Gaussian sequences achieve capacity.

3. we relate the capacity of the embedded information to the sacrifice in compression efficiency.

4. we draw new insights and implications for data hiding for aggressive and mild compression.

The next section summarizes the JND perceptual paradigm used for the data embedding and perceptual coding techniques and discusses the communications analogy for the problem. Section 3 describes the general analysis framework and class of systems encompassed by our theoretical work. Analytic results are presented in Section 4 followed by a discussion of the implications. Final remarks conclude the paper.

2 Paradigms and Principles

2.1 JND Perceptual Model

Many models exist to describe the masking characteristics of the human perceptual system [20]. Of these, one of the most popular is based on a JND paradigm [10]. A set of JNDs is associated with a particular invertible transform T. Given that a multimedia signal is transformed using T, the JNDs provide an upper bound on the extent that each of the coefficients can be perturbed without causing perceptual changes to the signal quality. The set of signal and transform dependent JNDs can be derived using complex analytic models or through experimentation.

Consider the discrete signal $f(i)$ transformed with T to produce the set of coefficients $F(u)$. By this paradigm, each $F(u)$ will have an associated JND, $J^*(u)$, such that we may form $F'(u)$ as follows:

$$F'(u) = F(u) + \beta(u)J^*(u) \quad (1)$$

where $\beta(u)$ is any signal with coefficients between the values -1 and 1. Taking the inverse transform T^{-1} of $F'(u)$ produces the signal $f'(i)$ which is guaranteed to be perceptually identical to $f(i)$. The challenge is to make the JND values as large as possible to fully exploit the masking characteristics of a broad class of signals.

For lossy compression, the JND values are used to determine the quantization step size or, equivalently, determine perceptually based bit allocation [10]. For data hiding in raw multimedia, they are used to compute the maximum level of signal energy embedded in specific signal coefficients. The maximization of this energy improves the robustness of the discreet data [18]. Even techniques which do not explicitly use the JND models such as [7] may be considered to fall within this class if we consider the JNDs to be conservative and trivially constant over all u.

When both data embedding and perceptual coding are applied to a signal, the combined effects of the processes should not result in a change to any host signal coefficient $F(u)$ which exceeds $J^*(u)$. Thus, we assume the individual perceptual models used for data hiding and compression are conservative. Specifically, if the data hiding algorithm is restricted to making changes to the coefficient $F(u)$ below or equal in magnitude to $\alpha(u)$, then the compression algorithm must have an effective JND for quantization of $J(u) = J^*(u) - \alpha(u)$ to be both efficient yet cause no visual distortions.

2.2 Communications Analogy for Data Hiding

One popular analogy for data hiding in the presence of distortion such as lossy compression is digital communications. Communicating the hidden signal information is likened to transmission of the signal through an associated communication channel as shown in Figure 1. Embedding the signal is equivalent to channel coding and extraction of the hidden information serves the same purpose as a communications receiver. As discussed in the introduction of the paper, for most data hiding applications the only potential source of manipulation after embedding is perceptual coding. For this situation, the process of lossy compression characterizes the associated communication channel for the hidden data.

It follows that many of the same figures of merit used in communications systems may be used to assess the quality of data hiding approaches. The particular measure we are concerned with in this paper is that of transmission capacity. We consider the relationship between capacity and the relative efficiency of both perceptual models used for hiding and compression. Employing the structured JND paradigm described in the previous section, we can treat the problem as an information theoretic one to derive new mathematical bounds and insights.

2.3 Models

The overall communication channel is considered to be comprised of smaller sub-channels denoted c_i, for $i = 1, 2, \ldots, M$. We assume that the coefficients $F(u)$ are grouped into disjoint sets G_i such that if $F(v) \in G_i$, then $\frac{\alpha(v)}{J(v)} = \mathcal{E}_i$ for some positive value $\mathcal{E}_i$ which we call the *relative perceptual efficiency*. The values of $\mathcal{E}_i$ do not necessarily have to be distinct for each i. As discussed in Section 2.1, $\alpha(u)$ is the maximum magnitude by which $F(u)$ can be perturbed to embed the hidden data, and $J(u)$ is the quantization step size for perceptual coding. The number of elements N in G_i is sufficiently large that a practical length channel code may be used to transmit watermark information. Embedding data into the coefficients in G_i effectively

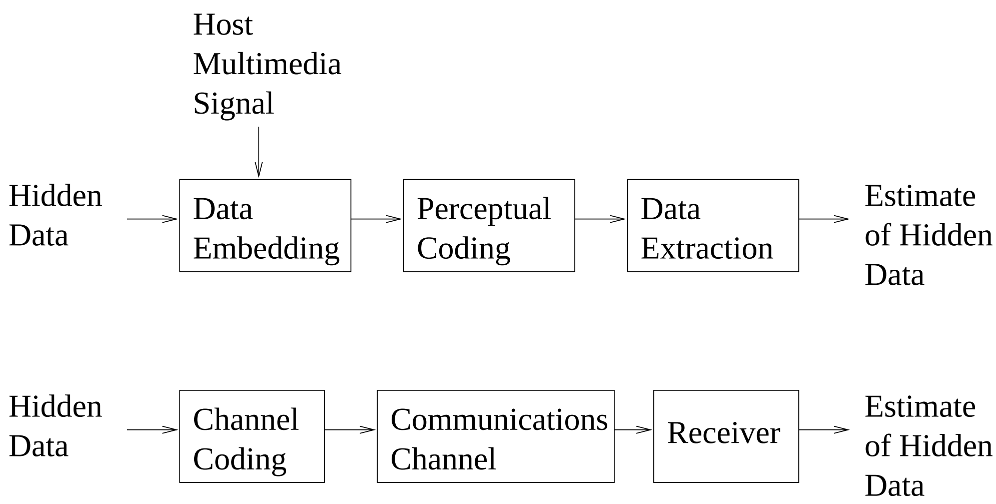

Figure 1. The Communications Analogy for Data Hiding.

represents one "use" of the sub-channel c_i. These assumptions are not restrictive for video marking applications in which a great deal of host data is available.

Using Bennett's Theorem [17], the quantization noise on $F(u)$ is modeled as a uniformly distributed random variable between $-J(u)$ and $J(u)$. Although Bennett's Theorem is valid for a narrow set of conditions [17], it is commonly applied for the analysis of A/D converters and provides some useful insights into their behaviour. In the same way, we believe that this assumption will also provide better understanding into watermarking in the face of JND perceptual coding. We assume that during watermark extraction or detection there is no interference from the host signal; all the watermark degradation is a result of the lossy compression process. The information is embedded by modifying the amplitude of $F(u)$ (i.e., effectively adding a signal to it), we then model the "noise" experienced by the hidden data in sub-channel c_i also as a uniformly distributed random variable.

3 Framework

3.1 Scope

Due to the convenience of using similar structures and transforms for data hiding and compression, the general trend in current research is to embed the watermark in the same domain as for performing perceptual coding [3, 5, 8, 22]. We adopt this framework in order for our results to be applicable to this broad class of data hiding and watermarking techniques. The capacity analysis in this work is applicable to the following situation:

- Data hiding of the information signal w into the digital multimedia signal f occurs in the transform T domain. Specifically, the data w is hidden in the discrete coefficients $F(u)$ produced by applying the invertible T on f. The new coefficients from the embedding process $\hat{F}(u)$ are transformed using the inverse of T to produce the output of the data hiding process denoted $\hat{f}$.

- For each coefficient $F(u)$, the signal change due to watermark embedding to produce $\hat{F}(u)$ does not exceed $\alpha(u)$ which is below the JND threshold $J^*(u)$ for that coefficient.

- Perceptual coding of a signal $\hat{f}$ occurs in the same domain as data hiding, after signal embedding. Specifically, the marked signal $\hat{f}$ is transformed with T to produce coefficients which are then quantized to reduce the signal storage requirements.

- The perceptual paradigm for lossy compression is based on the JND model. However, to keep the combined data hiding and lossy compression operations below perceptual detection, $\hat{F}(u)$ is quantized to degree $J(u) = J^*(u) - \alpha(u)$. Thus, if there was no data hiding, $\alpha(u) = 0$ for all u, and the lossy compression would be equivalent to standard JND perceptual coding using $J^*(u)$.

- The only source of error on the extracted information is due to lossy compression. The host signal does not provide any interference to the hidden data.

4 Analysis and Insights

4.1 Formulation

Let $W(u)$ represent the signal change in $F(u)$ to embed the hidden data. After compression, assuming no interference from the host signal, the received signal is

$$\hat{W}(u) = W(u) + Q(u) \tag{2}$$

where $Q(u)$ is additive uniformly distributed noise and is assumed to be independent of $W(u)$. For the remainder of the analysis, we drop the argument u.

Consider $|W| \leq \alpha$ and that the probability density function (pdf) of Q is given by

$$p_Q(q) = \begin{cases} \frac{1}{2J} & |q| \leq J \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Because of the independence of W and Q, it follows that

$$p_{\hat{W}}(\xi) = p_W(\xi) * p_Q(\xi) \tag{4}$$

where we use the notation that p_X is the pdf of random variable $X \in \{\hat{W}, W, Q\}$.

4.2 Results for Aggressive Compression

Aggressive compression is defined as a situation for which $\alpha < J$, or equivalently $\mathcal{E}_i < 1$. For this case, Equation 4 reduces to

$$p_{\hat{W}}(\xi) = \begin{cases} 0 & \xi \leq -\alpha - J \\ \frac{1}{2J} P_W(\xi + J) & -\alpha - J < \xi \leq \alpha - J \\ \frac{1}{2J} & \alpha - J < \xi \leq -\alpha + J \\ \frac{1}{2J}(1 - P_W(\xi + J)) & -\alpha + K < \xi \leq \alpha + J \\ 0 & \xi > \alpha + J \end{cases} \tag{5}$$

where $P_W(\xi)$ is the cumulative distribution function (cdf) of W (i.e., $P_W(\xi) = \int_{-\infty}^{\xi} p_W(\eta) d\eta$).

The *data hiding capacity* of the associated sub-channel c_i is defined as

$$C_i = \max_{p_W(\xi)} I(\hat{W}, W) \tag{6}$$

where $p_W(\xi)$ is the pdf of W, and $I(\cdot, \cdot)$ is the mutual information between the two argument distributions [6].

Taking the derivative of $I(\hat{W}, W)$ with respect to $p_W(\xi)$ and equating to zero, we find that the capacity is achieved for

$$P_W(\xi) = \begin{cases} 0 & \text{for } \xi \leq -\alpha \\ \frac{1}{2} & \text{for } -\alpha < \xi \leq \alpha \\ 1 & \text{for } \xi > \alpha \end{cases} \tag{7}$$

or equivalently,

$$p_W(\xi) = \frac{1}{2}\delta(\xi + \alpha) + \frac{1}{2}\delta(\xi - \alpha) \tag{8}$$

where $\delta(\xi)$ is the Dirac delta function. This means that to achieve capacity, W has a discrete binary uniform distribution. This is *not* a Gaussian distribution as many researchers have assumed because in the formulation of the problem we limit both the noise and water mark to be bounded in amplitude. The capacity of the sub-channel is then given by

$$\begin{aligned} C_i &= \mathcal{H}(\hat{W}) - \mathcal{H}(Q) && (9) \\ &= \left[\frac{\alpha}{J} + \log(2J)\right] - \log(2J) && (10) \\ &= \frac{\alpha}{J} && (11) \\ &= \mathcal{E}_i && (12) \end{aligned}$$

where $\mathcal{H}(\cdot)$ is the entropy of the argument random variable. Assuming that the signal and noise are independent for each sub-channel c_i, the overall data hiding channel capacity is given by

$$\mathcal{C} = \sum_{i=1}^{M} C_i = \sum_{i=1}^{M} \mathcal{E}_i \tag{13}$$

The "compression sacrifice" $\mathcal{CS}$ due to data hiding is defined as the number of additional bits required for storage of the signal because some of the perceptual masking properties are used for data hiding. For the uth coefficient, it is given by $\log_2(\frac{\alpha+J}{J})$. Thus, for sub-channel c_i, the bit rate sacrifice is $N \log_2(\mathcal{E}_i + 1)$. Relating this to the sub-channel capacities,

$$\mathcal{CS} = N \sum_{i=1}^{M} \log_2(C_i + 1). \tag{14}$$

where N is the number of coefficients comprising each sub-channel.

4.3 Results for Aggressive Data Hiding

For dominant data hiding, $\mathcal{E}_i \geq 1$. Using similar analysis and reasoning, the capacity $\mathcal{C}$ of each sub-channel is bounded as follows:

$$\log_2(\lfloor \mathcal{E}_i \rfloor + 1) \leq C_i \leq \log_2(\lceil \mathcal{E}_i \rceil + 1). \tag{15}$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and ceiling operators, respectively. The probability distribution of W which achieves capacity is discrete and uniformly distributed. Therefore,

$$\sum_{i=1}^{M} \log_2(\lfloor \mathcal{E}_i \rfloor + 1) \leq \mathcal{C} \leq \sum_{i=1}^{M} \log_2(\lceil \mathcal{E}_i \rceil + 1). \tag{16}$$

For the case in which $\mathcal{E}_i$ is a natural number, it is easily shown that

$$\mathcal{CS} = N\mathcal{C}. \tag{17}$$

That is, the bit rate sacrificed for more efficient storage is equal to the data hiding capacity bit rate times the number of coefficients in each sub-channel.

4.4 Implications

The analysis in this paper demonstrates that if the same transforms are used for both data hiding and perceptual coding then, at most, you can achieve a data hiding capacity equal to the loss in storage efficiency bit rate. This occurs for $N = 1$, and $\mathcal{E}_i$ a natural number (i.e., aggressive data hiding). In this case the quantization noise is almost always a smaller amplitude than the hidden data signal, so a very weak channel code (in particular, the trivial block length of $N = 1$) can be used as there is effectively a clean channel for transmission. For aggressive compression, the capacity is smaller than the loss of bit rate.

The results for capacity are analogous to those for additive white Gaussian noise channels except that the ratio of signal to noise energy is replaced by the relative perceptual efficiency. The modeling of compression effects on the hidden data as an amplitude limited uniformly distributed random variable, provides new insights into appropriate channel codes to achieve data hiding capacity. It has been considered in most theoretical work that Gaussian signals are appropriate for high capacity data hiding; this is true if the interference experienced by the hidden information is also Gaussian. However, using our classical model for quantization noise, we find that antipodal signals are more appropriate for achieving capacity in the presence of aggressive lossy compression. In fact, it is beyond the scope of this paper, but it can also be shown that generalized K-level antipodal signalling achieves capacity for aggressive data hiding when $\mathcal{E}_i$ is a natural number. This implies that some quantization-based algorithms such as those proposed in [11, 4] or those using pn-sequences [21] (assuming that the host signal interference is small) might be more appropriate for the task of data hiding than traditional Gaussian signals.

A long powerful block code needs to be employed to achieve capacity which suggests that N needs to be large. Thus, capacity can be reached practically for high volume host signals such as digital video. In addition, there is a relationship between N, $\mathcal{C}$, and $\mathcal{CS}$ in the sense that a more powerful block code can be used to asymptotically approach capacity, yet the cost to compression sacrifice $\mathcal{CS}$ grows linearly. Depending on the application and volume of host data available, an appropriate selection of N can provide a good compromise.

This work also provides insight into hiding data in the same domain as perceptual coding. We see that the gain in hiding ability is at most what you loose in compression efficiency. This motivates future investigation into using different transforms for both tasks. The fundamental limitation in using a structured perceptual paradigm which involves a particular transformation T and JNDs is that not all of the perceptual masking characteristics are exploited. Thus, use of *complementary* domains for the hiding and compression processes may allow one to hide information without sacrificing compression efficiency. Preliminary work in the area has demonstrated the potential of this approach for achieving higher capacity data hiding without investing compression efficiency [12, 13].

5 Final Remarks

In this work we provided analysis to gain insight into the capacity of a broad class of data hiding schemes. A communication paradigm for data hiding was established for which the primary source of channel noise was due to perceptual coding. Assuming the same structure using JND models for data hiding and perceptual coding, we see that at best the data hiding capacity is equal to the increased storage requirements of the information. In addition, through more appropriate modeling of quantization effects due to compression, we see that antipodal channel codes opposed Gaussian sequences are more appropriate to achieve capacity.

References

[1] W. Bender, D. Gruhl, and N. Morimoto. Techniques for data hiding. In W. Niblack and R. C. Jain, editors, *Proc. SPIE, Storage and Retrieval for Image and Video Databases III*, volume 2420, pages 164–173, February 1995.

[2] S. Bhattacharjee and M. Kutter. Compression tolerant image authentication. In *Proc. IEEE Int. Conference in Image Processing*, volume 1, October 1998.

[3] A. G. Bors and I. Pitas. Image watermarking using block site selection and DCT domain constraints. *Optics Express*, 3(12):512–523, December 7 1998.

[4] B. Chen and G. W. Wornell. Dither modulation: A new approach to digital watermarking and information embedding. In *Proc. SPIE, Security and Watermarking of Multimedia Contents*, volume 3657, January 1999.

[5] T.-Y. Chung, M.-S. Hong, Y.-N. Oh, D.-H. Shin, and S.-H. Park. Digital watermarking for copyright protection of MPEG2 compressed video. *IEEE Transactions on Consumer Electronics*, 44(3):895–901, August 1998.

[6] T. Cover and J. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., Toronto, 1991.

[7] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for images, audio and video.

In *Proc. IEEE Int. Conference on Image Processing*, volume 3, pages 243–246, 1996.

[8] V. Darmstaedter, J.-F. Delaigle, J. J. Quisquater, and B. Macq. Low cost spatial watermarking. *Computers & Graphics*, 22(4):417–424, 1998.

[9] H. Inoue, A. Miyazaki, A. Yamamoto, and T. Katsura. A digital watermark based on the wavelet transform and its robustness on image compression. In *Proc. IEEE Int. Conference in Image Processing*, volume 2, October 1998.

[10] N. Jayant, J. Johnston, and R. Safranek. Signal compression based models of human perception. *Proceedings of the IEEE*, 81:1385–1422, October 1993.

[11] D. Kundur and D. Hatzinakos. Digital watermarking using multiresolution wavelet decomposition. In *Proc. IEEE Int. Conference on Acoustics, Speech and Signal Processing*, volume 5, pages 2969–2972, 1998.

[12] D. Kundur and D. Hatzinakos. Attack characterization for effective watermarking. In *Proc. IEEE Int. Conference in Image Processing*, October 1999.

[13] D. Kundur and D. Hatzinakos. Mismatching perceptual models for effective watermarking in the presence of compression. In *Proc. SPIE, Multimedia Systems and Applications II*, September 1999.

[14] J. Lacy, S. R. Quackenbush, A. R. Reibman, D. Shur, and J. H. Snyder. On combining watermarking with perceptual coding. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing*, volume 6, pages 3725–3728, May 1998.

[15] J. Lacy, S. R. Quackenbush, A. R. Reibman, and J. H. Snyder. Intellectual property protection systems and digital watermarking. *Optics Express*, 3(12):478–484, December 7 1998.

[16] J. Meng and S.-F. Chang. Embedding visible video watermarks in the compressed domain. In *Proc. IEEE Int. Conference in Image Processing*, volume 1, 1998.

[17] S. R. Norsworthy, R. Schreier, and G. C. Temes, editors. *Delta-Sigma Data Converters: Theory, Design and Simulation*. Wiley, New York, 1997.

[18] C. I. Podilchuk and W. Zeng. Image-adaptive watermarking using visual models. *IEEE Journal on Selected Areas in Communications*, 16(4):525–539, May 1998.

[19] M. Ramkumar and A. N. Akansu. Theoretical capacity measures for data hiding in compressed images. In *Proc. SPIE, Voice, Video and Data Communications*, November 1998.

[20] M. S. Sanders and E. J. McCormick. *Human Factors in Engineering and Design*. McGraw-Hill, New York, 7th edition, 1993.

[21] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne. A digital watermark. In *Proc. IEEE Int. Conference on Image Processing*, volume 2, pages 86–90, 1994.

[22] H.-J. Wang and C.-C. J. Kuo. An integrated progressive image coding and watermark system. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing*, volume 6, pages 3721–3724, March 1998.

[23] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp. The effect of matching watermark and compression transforms in compressed color images. In *Proc. IEEE Int. Conference in Image Processing*, volume 1, October 1998.