Modulation & Multiplexing Techniques for Multimedia Data Hiding

Min Wu Bede Liu

Dept. of Electrical Engineering, Princeton University, Princeton, NJ 08544

ABSTRACT

Many multimedia data hiding systems demand either multiple bits or multiple sets of data to be embedded. This paper examines the modulation and multiplexing techniques for accomplishing the task of extending the basic single-bit embedding to multiple-bit embedding. Amplitude modulo modulation, orthogonal/bi-orthogonal modulation, and TDMA or CDMA type modulation/multiplexing are discussed and compared. Several examples are included to demonstrate the use of such techniques in practical designs.

Keywords: multimedia data hiding, digital watermarking, modulation, multiplexing.

1. INTRODUCTION

Data hiding has been proposed for a variety of applications involving digital media, such as ownership protection, copy control, annotation, and authentication. From a theoretical point of view, data hiding can be considered as a communication problem where the embedded data is the signal to be transmitted, and communication theory has been found to be useful in the study of data hiding.

Drawing the analogy between classic communication and data hiding, we can arrive at a conceptual layered architecture [18], as illustrated in Fig. 1. In communications, the gap between the theoretical Shannon channel capacity and the practical limitations are bridged by studies on system and implementation issues such as modulation/demodulation, coding/decoding, and equalization [1]. In this paper, we focused on the modulation and multiplexing techniques used to embed multiple bits and to hide multiple sets of data for different purposes. This is an issue that needs to be addressed in most designs, yet has not received much attention in literature regarding the pros and cons for each candidate strategy.

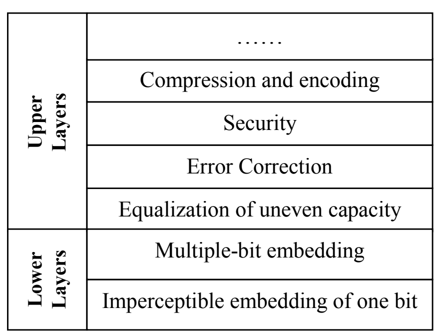

Figure 1. *Layered algorithmatic architecture of data hiding system*

In Section 2, we shall review a few concepts and principles of data hiding, including a data hiding framework, two basic embedding mechanisms, and the comparison of their embedding capacity. We then discuss the candidate

---

This work was supported by a New Jersey State R&D Excellence Award and Intel Technology for Education 2000 Grant. Min Wu is now with the ECE Dept., Univ. of Maryland, College Park. The authors can be reached at minwu@eng.umd.edu and liu@ee.princeton.edu .

modulation and multiplexing techniques in Section 3 to be used to extend the single-bit embedding to multiple-bit embedding. We shall cover amplitude modulo modulation, orthogonal/bi-orthogonal modulation, and TDMA/CDMA type modulation/multiplexing. These techniques are quantitatively compared in Section 4. Finally in Section 5, we present two examples, namely, data hiding in binary image and in video, to demonstrate the use of various techniques in practical designs.

2. PRELIMINARIES

In this section, we review a few concepts and principles of data hiding that will be used throughout the discussion in this paper.

2.1. A Data Hiding Framework

A typical data hiding framework is illustrated in Fig. 2. Starting with an original digital media (I_0), which is also commonly referred as the *host media* or *cover media*, the embedding module inserts in it a set of secondary data ($\underline{b}$), which is referred as *embedded data* or *watermark*, to obtain the *marked media* (I_1). The insertion or embedding is done such that I_1 is perceptually identical to the original I_0. The difference between I_1 and I_0 is the distortion introduced by the embedding process and is referred to as *embedding distortion*.

The embedded data b can be extracted from the marked media I_1 by a detector. However, I_1 may be subjected to various processing and attacks before detection. The input media to the detector is referred to as *test media* (I_2), and the difference between I_2 and I_1 is called *noise*. The *extracted data* from I_2 is denoted by $\hat{\underline{b}}$. In such applications as ownership protection, fingerprinting [1] and access control, accurate decoding of hidden data from distorted test media is preferred. They are commonly referred as *robust data hiding / watermarking*. In other applications such as authentication and annotation, robustness against processing and attacks are not a principal requirement in general.

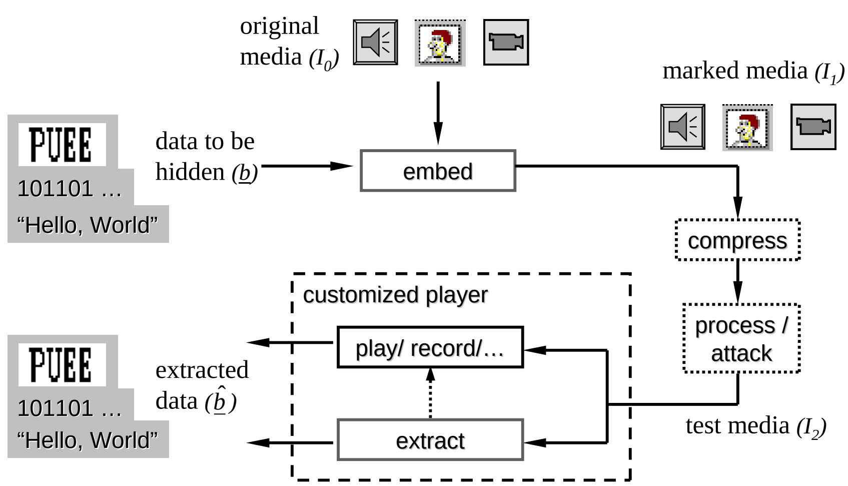

Figure 2. *General framework of data hiding systems*

2.2. Basic Embedding Mechanisms

The embedding of one bit in original media is basic to every data hiding system. Many embedding approaches have been reported in the literature. Some work with the multimedia signal samples while others with the transformed

[1] The *fingerprinting* here refers to the application where different labels are embedded in copies of the same media content before distributing to multiple recipients and the hidden labels are used for tracing each recipient.

data. It is helpful to study the existing embedding approaches under noise-free conditions (i.e., the test media is the same as the marked media) and to examine whether knowledge of the original host media will enhance the detection performance, regardless of whether a detector uses such knowledge or not [21]. Many existing embedding approaches would then fall in one of the following two categories.

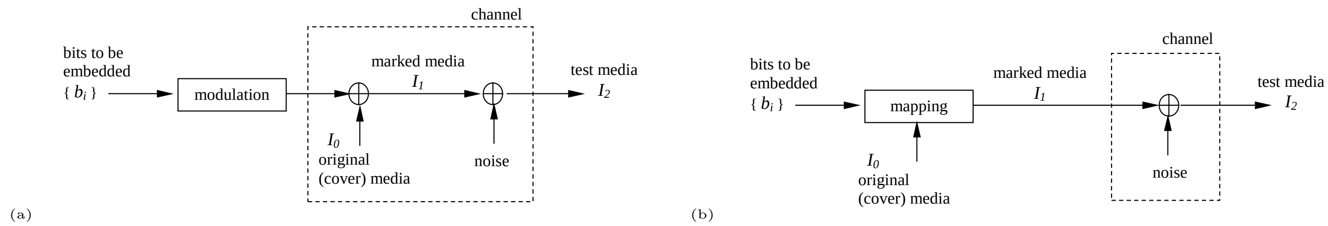

Figure 3. *Channel models for Type-I (a) and Type-II (b) embedding.*

In the first category, which we shall call *Type-I*, the secondary data, possibly encoded, modulated, and/or scaled, is added to the host signal, as illustrated in Fig. 3(a). The addition can be performed in a specific domain or on specific features. Considering the embedding of only one bit, the difference between marked signal I_1 and the original host signal I_0 is a function of b, the bit to be embedded, i.e., $I_1 - I_0 = f(b)$. I_0 can be regarded as a major noise source in such detection. Although it is possible to detect b directly from I_1 [7], the knowledge of I_0 will enhance detection performance by eliminating the interference. Additive spread spectrum watermarking is a representative of this category [5, 6].

In the second category (referred as *Type-II*), the signal space is partitioned into subsets which are mapped by a function $g(\cdot)$ to the set of values taken by the secondary data (e.g., $\{0, 1\}$ for binary hidden data), as illustrated in Fig. 3(b). The marked value I_1 is then chosen from the subset which maps to b, so that the relationship of $b = g(I_1)$ is deterministically enforced. To minimize perceptual distortion, I_1 should be as close to I_0 as possible. That is,

$$I_1 = \arg_I \min_{\text{s.t.} g(I)=b} D(I_0, I). \tag{1}$$

where the distance measure $D(\cdot, \cdot)$ is chosen using perceptual models. Unlike the first category, the detectors for this type of schemes do not need the knowledge of original value I_0 because the information regarding b is solely carried in I_1. Note that there may be other constraints imposed on I_1 for robustness considerations, for example, the enforcement may be done in a quantized domain with uniform quantization step size Q [18].

A simple example of Type-II is the *odd-even embedding*, whereby a closest even number is used as I_1 to embed a "0" and a closest odd number is used to embed a "1". The embedded data is detected by checking the odd-even parity (or equivalently, the least significant bit – LSB [2]) from a pixel or a coefficient of a test image. Data hiding can also be done by enforcing a global relationship. For example, one may change the sum of several source components to a nearby even number to encode a "0", and to an odd number to encode a "1". This is equivalent to reducing the bits allocated for representing the original vectors and to re-allocate them for conveying side information. When keeping the total distortion fixed and moving from 1-D space to a space of higher dimension, the magnitude of the introduced distortion per dimension is reduced. Also, there are more choices to select a new signal vector with desired bits embedded in, which allows embedding to be performed in such a way that the human-visual-model-weighted distortion is minimized. On the other hand, the embedding bit rate is reduced. This is a tradeoff between embedding rate and invisibility [3] .

[2] Please note that odd-even embedding is not equivalent to putting the bit to be embedded in the LSB because the embedding pursues the smallest necessary change, which may involve the bits beyond LSB. This can also be seen from the calculation of embedding distortion. The MSE of odd-even embedding is $Q^2/3$ while the embedding by replacing LSB is $7Q^2/12$.

[3] Equivalently, if the embedding distortion per dimension is fixed, the total distortion that can be introduced increases when moving to higher dimensions. This aggregated energy enables embedding more reliably via quantization, as discussed in [18].

The odd-even embedding can be viewed as a special case of the table-lookup embedding [11, 20], which uses a lookup table to determine the mapping between the possible values of a media component and the data to be embedded. There are many other possible ways to partition the space and to enforce a desired relationship. One can enforce the ordering of a pair of samples or coefficients v_1 and v_2. For example, we generate marked coefficients v'_1 and v'_2 close to v_1 and v_2 such that $v'_1 > v'_2$ to embed a "1" and $v'_1 \leq v'_2$ to embed a "0" [9]. One can also enforce signs to embed a "1" or "0", as used in [10, 12]. Extending the basic ways of enforcement, more sophisticated schemes can be designed and/or analyzed. Many proposed schemes in the literature that claimed to have the ability of non-coherent detection [4] belong to this category. It is the deterministically enforced relationship on I_1 that removes the need of using original signal I_0. For the convenience of discussion, we shall refer the collection of image pixels or coefficients on which the relation is enforced as an *embedding unit*. If the enforcement is performed on a quantity derived from the embedding unit (e.g., the sum of a few coefficients, the signs of a coefficient, etc.), we shall refer the quantity as a *feature*.

2.3. Capacity Comparison for Type-I & Type-II

By fixing the mean squared error introduced by the embedding process as E^2, we can compare the capacity of Type-I and Type-II schemes under AWGN noise with the following simplification. For Type-I, we consider a Continuous-Input-Continuous-Output (CICO) channel model and assume that the AWGN noise consists of gaussian processing noise (with variance σ^2) and host interference (with standard deviation 10 times as much as the amplitude of the watermark signal, i.e., $\sigma_I = 10E$). For Type-II, we consider a Discrete-Input-Discrete-Output (DIDO) Binary-Symmetric-Channel (BSC) for odd-even embedding with such quantization step Q that the embedding MSE distortion equals to E^2, i.e., $Q = \sqrt{3}E$. The capacity is thus obtained as:

$$C_I = \frac{1}{2}\log_2\left(1 + \frac{E^2}{(10E)^2 + \sigma^2}\right) \tag{2}$$

$$C_{II} = 1 - h_{\min\{1/2,\, 2\cdot\sum_{k=0}^{+\infty} \mathcal{Q}(\frac{(4k+1)Q}{2\sigma}) - \mathcal{Q}(\frac{(4k+3)Q}{2\sigma})\}} \tag{3}$$

The capacity vs. watermark-to-noise ratio E^2/σ^2 for the two types are plotted in Fig. 4. It shows that the capacity of Type-II is much higher than that of Type-I until the watermark-to-noise ratio (WNR) falls negative. The comparison suggests that Type-II is useful under low noise condition while Type-I is suitable for severe noise. The capacity of both Type-I and Type-II can be approached via channel coding, such as RS / BCH codes used in [8, 21]. More details about the capacity comparison can be found in [18].

Recently, motivated by Costa's techniques in proving the channel capacity [15], Chen *et al.* proposed to incorporate multiplicative scaling into quantization-based enforcement embedding. The enforcement is then linearly combined with the host signal to form a watermarked signal. The scaling factor is a function of watermark-to-noise ratio and has the capability of enhancing the number of bits that can be embedded. This can be viewed as a combination of Type-I and Type-II embedding. Interested readers may refer to [16, 17] for details.

3. MODULATION AND MULTIPLEXING TECHNIQUES

In this section, we discuss the techniques that may possibly be used to extend the single-bit embedding to multiple-bit embedding. They have evolved from the classic communication [1]. The applicability of a particular modulation/multiplexing technique also depends on the type of multimedia sources and the embedding mechanism being used, which will be seen in later discussions.

[4] Non-coherent detection in data hiding refers to being able to detect the embedded data without the use of the original unwatermarked copy. It is also called "blind detection".

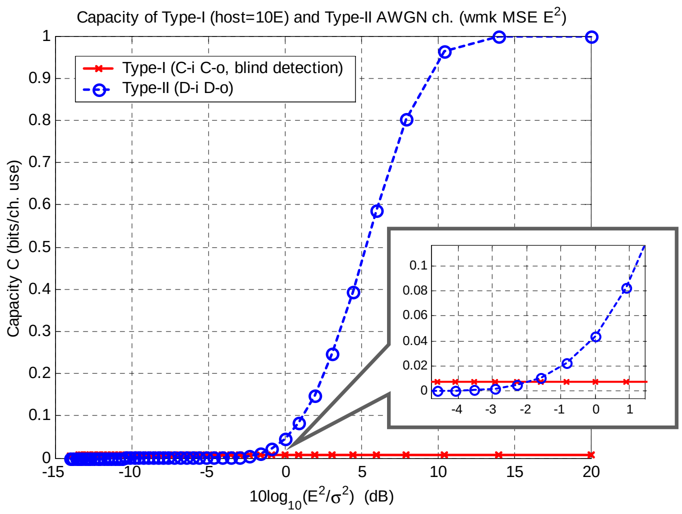

Figure 4. *Capacity of Type-I (CICO channel) and Type-II (DIDO channel) embedding under AWGN noise.*

3.1. Amplitude Modulo Modulation

The Type-I additive embedding formulated as an antipodal modulation or as an on-off modulation can be viewed as amplitude modulation. For blind detection of additive embedding that is subject to host interference, using amplitude modulation to convey more than two constellation points is rare in practice. We therefore focus here on the amplitude modulo modulation that is applicable to Type-II embedding.

In general, B bits can be embedded in each embedding unit by enforcing a feature derived from this unit into one of M subsets, where $B = \log_2 M$. A straightforward example that extends from odd-even embedding is to enforce the relation via modulo-M operation to hide B bit per element. That is,

$$I_1 = \arg_I \min_{\text{s.t. } I = kQ, k\in\mathbf{Z}, \text{mod}(k,M)=m} |I - I_0| \tag{4}$$

where $m \in \{0, 1, ..., M-1\}$ represents the B-bit information to be embedded, I_0 is the original image feature, I_1 is the watermarked feature, and Q is the quantization step size for obtaining robustness. Assuming I_0 follows uniform distribution in each quantization interval $(kQ - \frac{Q}{2}, kQ + \frac{Q}{2})$ where k is an integer, the MSE distortion introduced by embedding is $Q^2M^2/12$. This indicates that with the minimal separation Q between the M subsets being fixed, larger embedding distortion will be introduced by a larger M. For fixed MSE embedding distortion, the enforced relation with a larger M has smaller separation hence can tolerate less distortion. The idea is easily extensible to table lookup embedding or other enforcement scheme and the analysis is similar.

3.2. Orthogonal & Biorthogonal Modulation

The *orthogonal modulation* is mainly used for Type-I additive embedding. M orthogonal signals are used to represent $B = \log_2 M$ bits by embedding one of the M signals into the host media. A detector computes the correlation of the test signal with all M signals. The signal that gives the largest correlation and exceeds some threshold is selected as the signal embedded by the sender and the corresponding B-bit value determined accordingly. A variation, called *biorthogonal modulation*, encodes $\log_2 2M = (B+1)$ bits by adding or subtracting one of M signals. These two approaches are inefficient except for small M, because the computational complexity of detection grows exponentially with the number of bits being conveyed.

It should be noted that there are many degrees of freedom in selecting the M orthogonal signals, because each of such signal generally contains many elements to realize Type-I embedding. The difficulty is the bookkeeping of all

M such signals and for both the embedder and the decoder to agree on which ones are being used. In practice, M random signals having certain statistics (such as different spread spectrum signals) are used, as they are approximately orthogonal for large M. To ensure this, an orthogonality check may be done.

3.3. TDMA-Type Modulation and Multiplexing

This type of modulation/multiplexing partitions an image or audio source into non-overlapped regions or time segments and hides one or several bits in each region or segment. A video can be partitioned into regions within each frame and into time segments across frames. TDMA type modulation is a simple way to realize orthogonal embedding for both Type-I and Type-II, as the bits embedded in different regions or segments do not interfere with one other. However, different regions/segments can tolerate different amount of changes without causing perceptible artifacts. For example, very few bits can be embedded into a smooth area of an image, whereas more bits can be embedded into areas with significant amount of details. The difficulty arising from this uneven embedding capacity can be handled with random shuffling before embedding, as explained in [18].

3.4. CDMA-Type Modulation and Multiplexing

For Type-I additive embedding, B bits are encoded into to a watermark signal $\underline{w}$ via

$$\underline{w} = \sum_{k=1}^{B} b_k \cdot \underline{u}_k, \tag{5}$$

where $b_k \in \{\pm 1\}$, and the vectors $\{\underline{u}_k\}$ are chosen to be orthogonal to each other. Similar to the situation for orthogonal/biorthogonal modulation, there is considerable freedom in selecting the B orthogonal vectors. The orthogonality of $\{\underline{u}_k\}$ implies that the total signal energy is the sum of the energy allocated for each bit. If a fixed, total amount of energy is uniformly allocated to each bit, the energy per bit will be reduced as B increases, implying a decrease in detection reliability and more generally, a limit on the total number of bits that can be hidden for low error rate extraction.

For Type-II, the embedding of multiple bits can be done by enforcing relations deterministically along different directions that are orthogonal to each other. For images, relations on the projections of a feature vector along several orthogonal directions can be enforced in an image block [13, 14]. The total modification introduced by embedding is the sum of the change along each direction, implying a tradeoff among capacity, robustness, and imperceptibility.

4. COMPARISON OF MODULATION/MULTIPLEXING TECHNIQUES

4.1. Applicable Media Types

Amplitude Modulo Modulation is applicable to most medias including audio, image, and video, as long as the features participating in the embedding are properly chosen. TDMA can be used in the temporal domain for audio and video, as well as in spatial domain for image and video. For both generalized CDMA [5] and orthogonal/biorthogonal modulation, one needs mutually orthogonal directions in the embedding domain, which can be a significant number at times hence can become a non-trivial task. For example, it is rather difficult to find in a binary image many orthogonal feature directions that are manipulable within the just-perceptible-range. Audio offers another example in that a large window of samples would be needed in order to get many overlapped but orthogonal directions. This will lead to significant processing delay, causing difficulties in situations requiring real-time considerations.

[5] Here, by "generalized" we mean to exclude the special case of TDMA.

4.2. TDMA vs. CDMA Approaches

TDMA and CDMA approaches are equivalent from the energy allocation point of view. TDMA is a special case with the supports of $\underline{u}_k$ being non-overlapping with each other in the sample domain, i.e., the pixel domain for image and the time domain for audio. Alternatively one can choose orthogonal but overlapped $\{\underline{u}_k\}$, similar to CDMA in communication [2]. Uneven embedding capacity is no longer a concern as the $\{\underline{u}_k\}$ can be chosen so that each bit is spread over all the media data. But B orthogonal sequences have to be generated and shared with the detector(s), which may require additional effort for large B. The TDMA and CDMA approaches can be combined to encode multiple bits.

4.3. TDMA/CDMA vs. Orthogonal Modulation

The orthogonal modulation and TDMA/CDMA-type modulation can be compared by studying the distance between signal constellation points that represent the hidden data (Fig. 5). This distance, in many cases, are directly related to the likelihood of detection errors. Considering the case of conveying B bits using total energy $\mathcal{E}$. The minimum distance between signal points is $\sqrt{2\mathcal{E}}$ for orthogonal modulation, and is $2\sqrt{\mathcal{E}/B}$ for TDMA/CDMA. When $B > 2$, orthogonal modulation gives smaller probability of detection error at a cost of detection complexity.

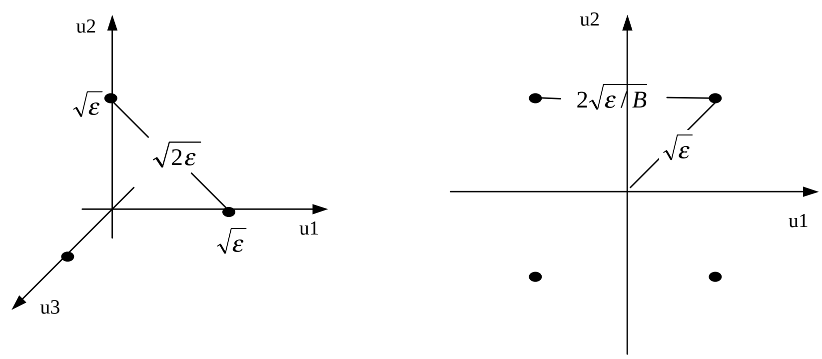

Figure 5. *Comparison of distance between signal constellation points for orthogonal modulation (left) vs. TDMA/CDMA-type modulation (right) with total signal energy being fixed at $\mathcal{E}$.*

TDMA or CDMA can be combined with orthogonal or biorthogonal modulation to increase the embedding rate while balancing the detection complexity. For example, a watermark conveying $2B$ bits is formed by

$$\underline{w} = \sum_{k=1}^{B} b_k \cdot [I(b_{B+k} = 1) \cdot \underline{u}_k^{(1)} + I(b_{B+k} \neq 1) \cdot \underline{u}_k^{(2)}] \tag{6}$$

where $I(\cdot)$ is an indicator function, and all vectors in the two sets $\{\underline{u}_k^{(1)}\}$ and $\{\underline{u}_k^{(2)}\}$ are orthogonal. Here we have used the TDMA/CDMA type modulation to convey B bits and used the orthogonal modulation to double the number of bits conveyed. The resulting total watermark energy is the same as using TDMA or CDMA alone.

4.4. Energy Efficiency

A comparison of the modulation/multiplexing techniques discussed above [6] is summarized in Table 1. The modulation/multiplexing is applied to one embedding unit of S elements. The quantity $\mathcal{W} = \frac{\mathcal{Y}}{\mathcal{X}\cdot\mathcal{Z}^2}$ measures the energy efficiency of embedding, where $\mathcal{X}$ is the number of embedded bits per element, $\mathcal{Y}$ is the MSE distortion per element introduced by embedding, and $\mathcal{Z}$ the minimum separation between the enforced constellation points hence reflects

[6] The modulo-M modulation extended from odd-even embedding is taken as a representative of amplitude modulation.

the robustness against noise. Because $\mathcal{W}$ describes the MSE embedding distortion per bit per unit squared separation distance, a smaller value is more preferable. It can be seen that, except for very small S and M, biorthogonal techniques has the smallest $\mathcal{W}$ values, while the amplitude modulo technique gives large $\mathcal{W}$ values as M grows larger – it equals to $\frac{1}{3}$ for $M = 2$, and to $\frac{2}{3}$ for $M = 4$. This suggests that to embed multiple bits with limited watermark energy, orthogonal and biorthogonal modulation should be used at a cost of computation. On the other hand, TDMA and CDMA techniques, being applicable to both Type-I and Type-II embedding under blind detection as well as having a constant $\mathcal{W}$ value of $\frac{1}{4}$ and linear complexity, show broad applicability and a good balance between energy efficiency and detection complexity.

Table 1. *Comparison of Modulation/Multiplexing Techniques*

(S elements per embedding unit, with $B \leq S$ and $M \leq S$)

	Amplitude Modulo	*TDMA / CDMA*	Orthogonal	Biorthogonal
Type-I embed.		Applicable	Applicable	Applicable
Type-II embed.	Applicable	Applicable		
$\mathcal{X}$ # embedded bits per element	$\frac{\log_2 M}{S}$	$\frac{B}{S}$	$\frac{\log_2 M}{S}$	$\frac{\log_2 2M}{S}$
$\mathcal{Y}$ MSE distortion per element	$\frac{Q^2 M^2}{12S}$	$\frac{\mathcal{E}}{S}$	$\frac{\mathcal{E}}{S}$	$\frac{\mathcal{E}}{S}$
$\mathcal{Z}$ minimum separation	Q	$2\sqrt{\frac{\mathcal{E}}{B}}$	$\sqrt{2\mathcal{E}}$	$\sqrt{2\mathcal{E}}$
$\mathcal{W} = \frac{\mathcal{Y}}{\mathcal{X} \cdot \mathcal{Z}^2}$ energy efficiency of embedding	$\frac{M^2}{12 \log_2 M}$	$\frac{1}{4}$	$\frac{1}{2 \log_2 M}$	$\frac{1}{2(1+\log_2 M)}$
computational complexity for detecting B bits	*const*	$O(B)$	$O(2^B)$	$O(2^{B-1}) \sim O(2^B)$

5. APPLICATIONS AND EXPERIMENTAL RESULTS

In this section, we use two specific designs and some experimental results to demonstrate the use of various modulation/multiplexing techniques.

5.1. Hiding Multiple Bits in Binary Image

Hiding data in binary image is considered a challenging problem because of the limited room to make invisible changes. A block-based pixel-domain approach was proposed in [18][19] to hide a non-trivial number of bits in such binary images as signatures, documents, and line drawings. Shown in Fig. 6 is an example of hiding 70-bit date information of “01/01/2000” in a 120×150-pixel line drawing.

Because of the difficulty in obtaining many orthogonal directions to hide data imperceptibly in binary image, a Type-II embedding mechanism was used to enforce block-based relationship to embed one bit, and a spatial TDMA approach was used to hide multiple bits by embedding one bit in each block. Shuffling was used to equalize the uneven embedding capacity, and one bit is embedded in each shuffled block of 256-pixel.

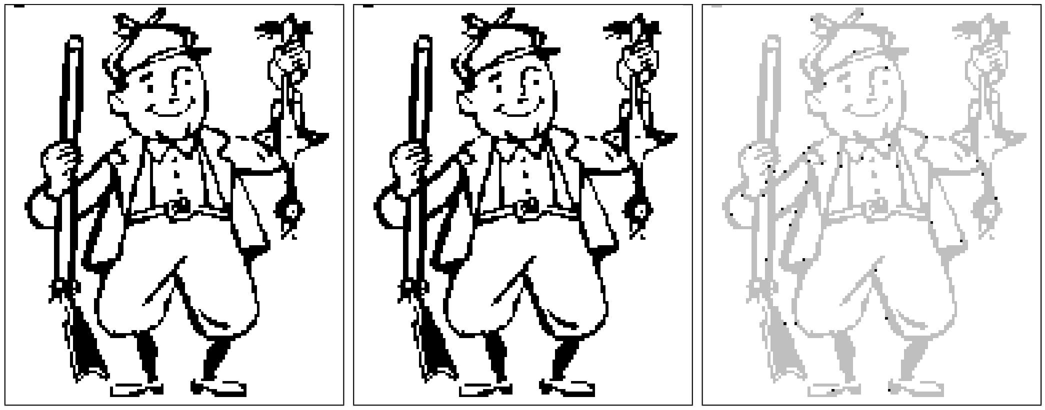

Figure 6. *An example of data hiding in binary image: (left) the original image, (middle) a watermarked copy with 10-letter date information "01/01/2000" embedded in, (right) the difference between the original and the watermarked (shown in black).*

5.2. Hiding Multiple Bits in Video

Hiding a nontrivial amount of data robustly in digital video is a challenging task. A multilevel embedding was proposed in [22] that allows the number of extractable bits to be adaptive with the actual noise conditions. The basic idea is to use both spread spectrum additive technique (Type-I) and odd-even enforcement in quantized domain (Type-II) to hide two sets of data payload, each with a different tradeoff between embedding rate and robustness. Data are embedded progressively from frame to frame with repetition in time over several consecutive frames to combat frame jitter. Each frame is essentially treated as a separate image for embedding. An example is shown in Fig. 7, where a total of 772 payload bits are embedded at two levels of robustness in the first 60 frames of the flower garden sequence.

In the following, we shall focus on the modulation and multiplexing strategies used in this design. Those strategies are used not only to embed multiple bits for the principal payload, but also to convey additional information that facilitates the extraction of the principal payload.

Principal data payload with Type-I embedding It is seen from Section 2.3 that the embedding rate of Type-I approaches is inherently low under blind detection. Nevertheless, one can use properly selected modulation techniques to enhance the embedding rate to some extent. Considering the tradeoff between the computation complexity and the energy efficiency of embedding, we have chosen the combination of spatial TDMA (with shuffle) and orthogonal modulation, as described in Eq. 6. Compared with using TDMA alone, this combined modulation helps doubling the number of embedded bits; compared with using CDMA, spatial TDMA with shuffle helps saving the storage of a large number of non-trivial orthogonal directions.

Principal data payload with Type-II embedding Spatial TDMA with shuffle is the typical modulation/ multiplexing for Type-II embedding, and therefore are adopted here.

Control bits with Type-I embedding In addition to the principal data payload, a small amount of side information, such as the number of bits being embedded in a frame and the frame index, must also be conveyed to facilitate the accurate extraction of the principal payload. The robustness of these additional data (known as "control bits") is critical, yet one has to limit the power allocated on them to avoid too much overhead. Such requirements prompt us to choose the robust Type-I embedding with orthogonal or biorthogonal modulation that has high energy efficiency. For example, to convey the modulo-N frame index [22] that changes in a round robin fashion $\{0, 1, ..., N-1, 0, 1, ...\}$,

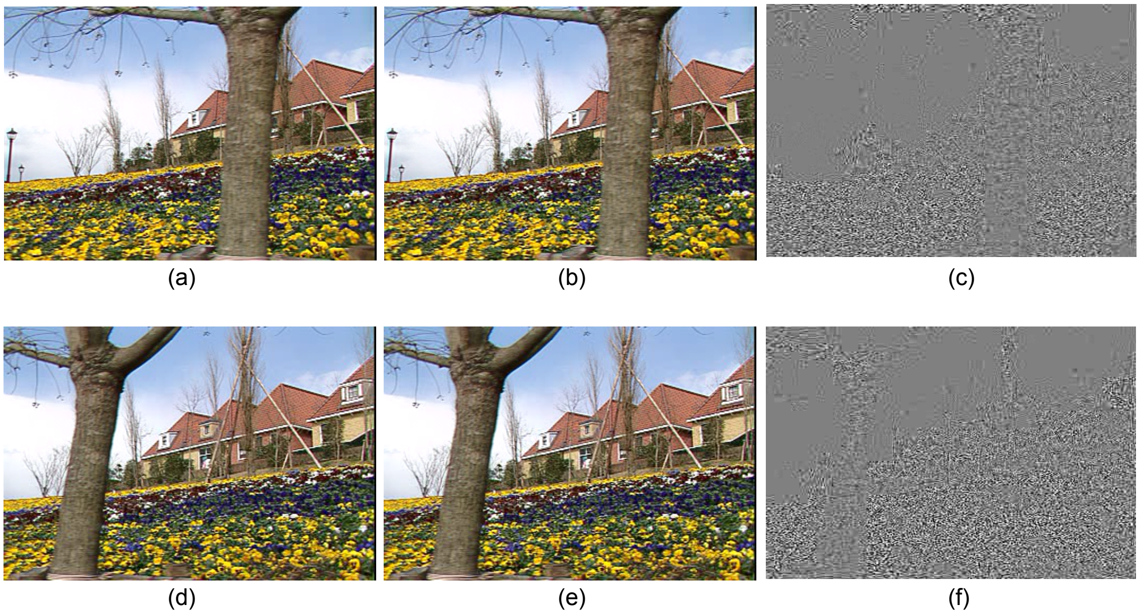

Figure 7. *Multi-level data hiding for flower garden video sequence: (a)-(c) the original 1st frame, the watermarked version, and their difference, respectively; (d)-(f) the original 30th frame, the watermarked version, and their difference, respectively. Both videos are compressed using MPEG-2 4.5Mbps, and the differences are enlarged with gray denoting zero difference and black/white denoting large difference. A total of* 772 *payload bits are embedded at two levels of robustness in the first 60 frames of the flower garden sequence.*

we select N orthogonal watermark signals labeled from 0 to $(N-1)$, and add the corresponding one in a frame. These watermark signals, properly scaled according to human visual systems, have the same number of elements as the frame size, hence are embedded all over each frame image. To avoid interference, different watermark signals used for different types of side information are mutually orthogonal, and they are also orthogonal with those used for embedding the principal payload.

6. SUMMARY AND CONCLUSION

This paper addresses the issues of extending single-bit embedding to multiple-bit embedding via modulation and multiplexing. Five approaches are investigated: amplitude modulo modulation, TDMA, CDMA, orthogonal and biorthogonal modulation. It is seen that amplitude modulo modulation, while requiring the least amount of computation, introduces the largest distortion. The orthogonal and biorthogonal modulation have the highest energy efficiency at a cost of computation. They are suitable for conveying a small amount of side information that has to be robustly embedded. The TDMA and CDMA approaches offer a good balance between energy efficiency and detection complexity, and thus suitable in a broad range of applications. The combination of TDMA/CDMA and orthogonal/biorthogonal modulation can effectively enhance the embedded data rate with the same amount of watermark energy and a small increase of computation. The two examples of data hiding in binary images and in video illustrates the use of various techniques in practical designs.

Acknowledgment

The authors would like to thank Dr. Heather Yu of Panasonic Info. & Networking Labs for the collaboration on the data hiding work for video [22], which motivated our study of modulation and multiplexing techniques in data hiding.

References

[1] J.G. Proakis: *Digital Communications*, 3rd Ed., McGraw-Hill, 1995

[2] S. Verdu: *Multiuser Detection*, Cambridge University Press, 1998

[3] T.M. Cover, J.A. Thomas: *Elements of Information Theory*, 2nd Ed., John-Wiley & Sons, 1991

[4] W. Bender, D. Gruhl, N. Morimote, "Techniques for Data Hiding", *Proc. of SPIE*, vol.2420, pp40, 1995.

[5] I. Cox, J. Kilian, T. Leighton, T. Shamoon: "Secure Spread Spectrum Watermarking for Multimedia", *IEEE Transaction on Image Processing*, vol.6, no.12, pp.1673-1687, 1997.

[6] C. Podilchuk, W. Zeng: "Image Adaptive Watermarking Using Visual Models", *IEEE Journal Selected Areas of Communications (JSAC)*, vol.16, no.4, May, 1998

[7] W. Zeng, B. Liu, "A statistical watermark detection technique without using original images for resolving rightful ownerships of digital images," *IEEE Trans. Image Processing*, vol. 8, no. 11, pp. 1534-1548, Nov. 1999.

[8] J.J.K. *Ó*Ruanaidh, T. Pun, "Rotation, Translation and Scale Invariant Spread Spectrum Digital Image Watermarking", *Signal Processing*, vol.66, no.3, 1998

[9] E. Koch, J. Zhao: "Towards Robust and Hidden Image Copyright Labeling", *IEEE Workshop on Nonlinear Signal and Image Processing*, 1995.

[10] C-T. Hsu, J-L. Wu: "Hidden Signatures in Image", *ICIP*, vol. 3, 1996

[11] M. M. Yeung, F. Mintzer: "An Invisible Watermarking Technique for Image Verification", *ICIP*, 1997.

[12] M. Ramkumar, A.N. Akansu: "A Robust Scheme for Oblivious Detection of Watermarks / Data Hiding in Still Images", *Symposium on Voice, Video, and Data Communication*, SPIE, 1998

[13] M. D. Swanson, B. Zhu, A. H. Tewfik: "Robust Data Hiding for Images", *IEEE DSP Workshop*, 1996.

[14] M. Alghoniemy, A.H. Tewfik: "Self-synchronizing Watermarking Techniques", *Symposium on Content Security and Data Hiding in Digital Media*, NJ Center for Multimedia Research and IEEE, 1999

[15] M.H.M. Costa: "Writing on Dirty Paper", *IEEE Trans. on Info. Theory*, vol. IT-29, no. 3, May 1983.

[16] B. Chen: *Design and Analysis of Digital Watermarking, Information Embedding, and Data Hiding Systems*, Ph.D. Thesis, MIT, 2000.

[17] B. Chen, G.W. Wornell: "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding", to appear in *IEEE Trans. on Info. Theory*, Sept., 2000.

[18] M. Wu: *Multimedia Data Hiding*, Ph.D. Dissertation, Princeton University, 2001.

[19] M. Wu, B. Liu: "Data Hiding in Binary Images", submitted to *IEEE Trans. on Multimedia*, 2001.

[20] M. Wu, B. Liu: "Watermarking for Image Authentication", *IEEE International Conference on Image Processing (ICIP'98)*, Chicago, IL, 1998.

[21] M. Wu, H. Yu, A. Gelman: "Multi-level Data Hiding for Digital Image and Video", *Proceedings of SPIE*, vol. 3845, *Photonics East Conference on Multimedia Systems and Applications*, Boston, MA, 1999.

[22] M. Wu, H. Yu: "Video Access Control via Multi-level Data Hiding", *IEEE International Conference on Multimedia & Expo (ICME'00)*, New York City, NY, 2000.