

# Host-aware spread spectrum watermark embedding techniques

Hugh Brunk  
Digimarc, 19801 SW 72<sup>nd</sup> Ave., Tualatin, OR 97062

## ABSTRACT

This paper explores techniques that involve the use of the embedder's knowledge of the cover work to help determine the watermark signal to be added to it. While the receiver always seeks to maximize a detection statistic which is a function of an a priori known pseudorandom sequence, the signal added to the cover work by the embedder is allowed to vary, on a per-chip basis, based upon the characteristics of the cover work. Although adaptation of an added watermark signal can be aimed at minimization of visual artifacts, this paper focuses on adaptation of the watermark signal to improve the readability of the signal outside of any human visual system constraints. This idea can be applied in various scenarios. Two specific examples are discussed. When source models are available and maximum-likelihood detection is used, the added watermark signal can be allowed to adapt to host signal variations in order to maximize the likelihood ratio detection statistic used at the receiver. Another instance where per-chip variation can be put to use is when a pre-filter is used to suppress the cover work prior to reading the watermark signal. In this case, the watermark signal is varied in such a way as to maximize the signal at the output of the pre-filter.

Keywords: Watermarking, adaptive watermarking, informed embedding, host-aware embedding

## 1. INTRODUCTION

Spread-spectrum watermarking is one of the most widely used watermarking methods; an early example was proposed by Cox in [1]. Spread-spectrum watermarking's good performance stems from the relative immunity of the underlying spread-spectrum modulation to noise and jamming. Many improvements have been made to spread-spectrum watermarking, both in embedding and in watermark detection and reading. One important development has been the application of perceptual models, such as human visual system models, to the watermark embedding process. These perceptual models allow watermark signals to be added to a source with maximum efficiency. Another area of improvement has been in the development of filters for use prior to watermark detection and reading. In image watermarking, such filters are designed to suppress the host image and thereby boost the watermark to host signal ratio. Equivalently, this amounts to estimating the host image and subtracting it from the watermarked image under study. Examples of these types of filters can be found in [2][3][4].

Spread spectrum watermarking is akin to spread spectrum CDMA technology used for wireless communication. There are, however, important practical differences in how spread spectrum technology functions in these two applications. In spread spectrum watermarking, the host signal is analogous to the channel noise in CDMA wireless communication. A key difference between the practical application of spread spectrum technology to watermarking and wireless communication becomes apparent: in watermarking the embedder has perfect knowledge of the host signal, however, in wireless communication, the transmitter does not have similar knowledge of the wireless channel noise. In [5], Malvar presents an improvement to spread-spectrum watermark embedding that takes advantage of this difference. This improvement allows the embedded signal power to vary for each bit embedded. The embedded signal is split between embedding the normal spread-spectrum signal and compensating for the host component interference, which is the inner product of the host and the spread-spectrum PN sequence. This compensation technique is similar to a process of pre-canceling developed in [6].

This paper explores a pair of watermark embedding techniques similar to that of [5]. However, instead of fixing the watermark variance for all host samples used to embed a bit, these techniques allow the watermark strength to be varied on a per-chip basis. This per-chip variation is informed by knowledge of the host signal characteristics, and by knowledge of the design of the watermark decoder. This idea of varying the watermark signal on a per-chip basis is applied to systems with two different watermark decoder designs. The first system considered uses a model of

the host signal distribution to perform maximum-likelihood decoding. In this case the embedded signal is varied in order to maximize the likelihood ratio statistic at the decoder. The second system's decoder uses a pre-filter to suppress the host signal prior to performing correlation-based detection. For this system, the embedded system is varied in a way that maximizes the watermark signal present at the pre-filter output.

Section 2 introduces the spread-spectrum notation to be used and Section 3 develops the host aware technique that is the topic of the paper. Section 4 describes the application of this technique to cases of spread-spectrum watermarking where MAP decoding is used and a source model is available to model the host signal. The next section presents use of the technique in the case where a pre-filter is used by the watermark detector/reader prior to detection. Section 6 draws some conclusions and presents areas for future work.

## 2. NOTATION

Figure 1 illustrates a basic spread spectrum watermark embedding process for images. An  $N$  bit message  $b = (b_1, b_2, b_3, \dots, b_N)$ ,  $b_i \in \{-1, 1\}$  is embedded into a host image  $x = (x_1, x_2, x_3, \dots, x_M)$  containing  $M$  sample values. These sample values may be taken from the spatial domain or from a transform domain such as the DCT or DWT. The image  $x$  is partitioned into  $N$  disjoint ordered sets  $S_1, S_2, S_3, \dots, S_N$  of  $\lfloor \frac{M}{N} \rfloor$  samples, each of which is referred to as a *chip*. A random key  $K$  is used to initialize a pseudorandom generator which produces a random sequence  $u = (u_1, u_2, u_3, \dots, u_n)$  of  $n = \lfloor \frac{M}{N} \rfloor$  samples with  $u_i \in \{-1, 1\}$ . Bit  $b_i$  is embedded in the image samples of  $S_i$  by calculating  $y = x + \sigma_w b_i u_i$ , where here  $x$  denotes the ordered samples of  $x$  in  $S_i$ .

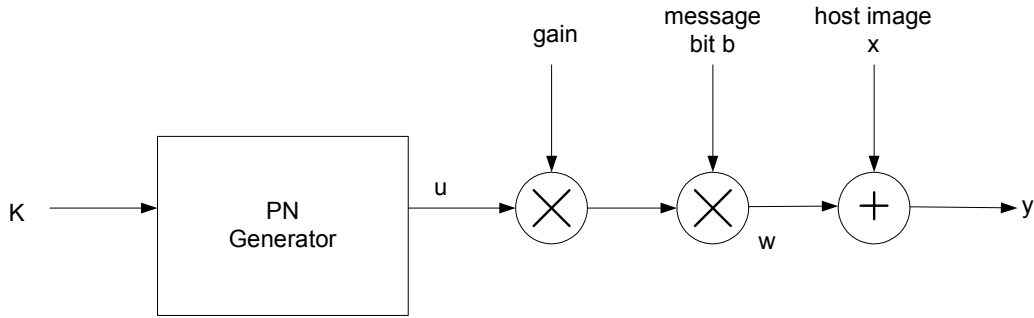


Figure 1: Spread spectrum embedding process.

The simplest method of decoding a spread spectrum watermark is the correlation detector, which is optimal for uncorrelated Gaussian host and image noise. Figure 2 shows a diagram of a simple correlation detector. The random sequence  $u$  is reconstructed from key  $K$ , and an estimate of message bit  $b$  is calculated from the inner product of  $u$  and  $y$ :

$$\hat{b} = \text{sign}\left(\frac{1}{n} \sum_{j=1}^n y_j u_j\right). \quad (1)$$

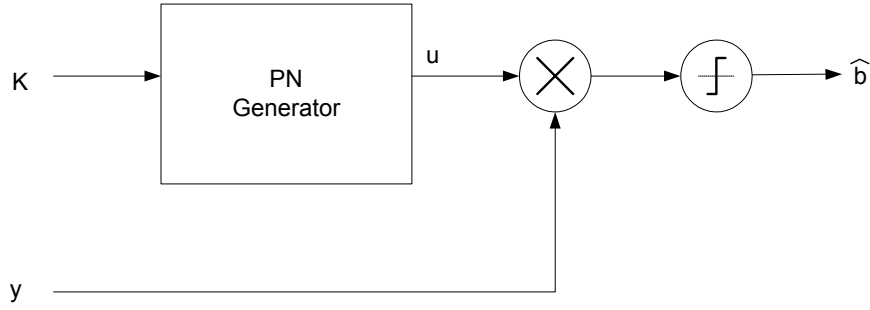


Figure 2: Correlation detection of spread spectrum watermark.

In some cases where a model of the host distribution is available, a correlation detector may not be the optimal detection method for decoding spread spectrum watermarks. Such an example is when the host samples are taken from an image transform domain such as the DCT or DWT. In this case it has been found that the transform domain coefficients can be effectively modeled as coming from a generalized Gaussian distribution [7][8].

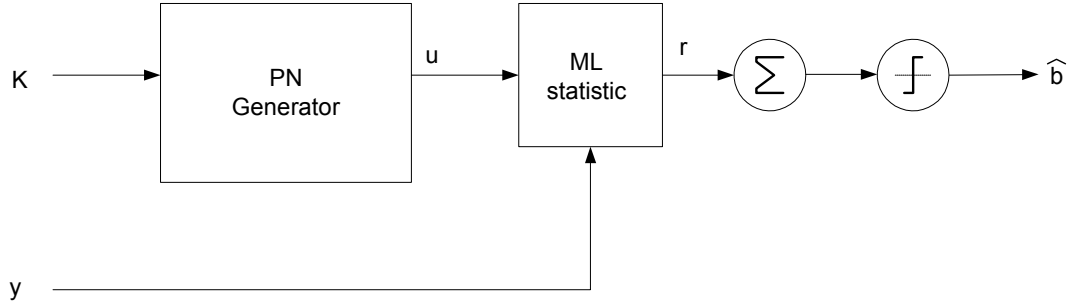


Figure 3: Maximum Likelihood detection of spread spectrum watermark.

When a host model is available, maximum likelihood (ML) detection may provide performance better than that of a correlation detector. A ML detector chooses as its estimate of the embedded bit the value which is most likely, given the received image samples  $y$ :

$$\hat{b} = \arg \max(f_x(y + bu\sigma_w)), \quad (2)$$

where  $f_x$  is the host distribution pdf. The log-likelihood ratio is commonly used for ML detection; the sign of the log-likelihood ratio determines the estimate of the embedded bit:

$$\hat{b} = \text{sign}\left(\ln \frac{f_x(y|b=1)}{f_x(y|b=-1)}\right). \quad (3)$$

### 3. HOST-AWARE ADAPTATION

The spread spectrum embedding process described in the previous section adds a watermark of fixed magnitude  $\sigma_w$  to each host image sample. In some situations, allowing the strength of the watermark to vary based upon host characteristics can provide superior performance. In [5], Malvar fixes the magnitude of the watermark over each set of image coefficients  $S$  (within the samples common to an embedded bit), while allowing the watermark magnitude to vary between embedded bits. In this paper, the opposite is considered: the energy available for

embedding a given bit is assumed fixed, and the magnitude of the watermark signal is allowed to vary over each chip, constrained by the energy available for embedding the bit:

$$y_i = x_i + \beta_i b u_i, \quad i = 1, 2, \dots, n. \quad (4)$$

$$\beta_i = f(x, E, b, u), \quad \text{where } E = n \sigma_w.$$

The function  $f$  serves to allocate the watermark signal over the set of image samples in a way that maximizes performance. One possible goal is for  $f$  to choose the watermark allocation to maximize the detection statistic (or its expected value) at the detector. The detection statistic for both the correlation detector and the ML detector are sums of terms, with one term for each chip of the embedded bit. In both of these cases, the goal in embedding a bit can be described as choosing the  $\beta_i$  to maximize

$$F = b \sum_{i=1}^n f_D(x_i + \beta_i b u_i),$$

$$\text{constrained by: } \sum_{i=1}^n \beta_i^2 \leq E, \quad (5)$$

where  $f_D(\cdot)$  represents the detection statistic function for a single chip; this function could represent either a correlation detector or a ML detector. Let  $\varepsilon_i$  represent the energy of the watermark in the  $i^{\text{th}}$  chip. Then one method of choosing the  $\beta_i$  is to require that

$$\frac{\partial F(\cdot)}{\partial \varepsilon_i} = \frac{\partial F(\cdot)}{\partial \varepsilon_j}, \quad 1 \leq i \neq j \leq n, \quad (6)$$

at the solution point.

#### 4. APPLICATION TO MAXIMUM LIKELIHOOD DETECTOR

Maximum likelihood detection may be used if a model of the host signal distribution is available. Both DCT and DWT coefficients of images have been shown to be closely approximated by members of the family of generalized Gaussian distributions. This class of distributions is characterized by  $\alpha$ , an exponential decay rate and  $\sigma$ , the standard deviation, where the p.d.f.  $p_X(x)$  is described by

$$p_X(x) = \left[ \frac{\alpha \eta(\alpha, \sigma)}{2\Gamma(1/\alpha)} \right] \exp\left\{-[\eta(\alpha, \sigma)|x|]^\alpha\right\} \quad -\infty < x < \infty$$

$$\text{and } \eta(\alpha, \sigma) \equiv \sigma^{-1} \left[ \frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)} \right]^{1/2}. \quad (7)$$

If the host has a generalized Gaussian distribution, the log-likelihood ratio test can be calculated as:

$$\hat{b} = \text{sign} \left( \sum_{j=1}^n \frac{|y_j + u_j \sigma_w|^\alpha - |y_j - u_j \sigma_w|^\alpha}{\sigma^\alpha} \right). \quad (8)$$

For  $\alpha=1$  (a laplacian distribution), the function  $f_D$  from Eq. 5 can be written as:

$$f_D(x) = \begin{cases} \frac{2\sigma_w}{\sigma} : x > \sigma_w \\ \frac{2x}{\sigma} : -\sigma_w \leq x \leq \sigma_w \\ \frac{-2\sigma_w}{\sigma} : x < -\sigma_w \end{cases} \quad (9)$$

In this case, it is relatively simple to construct an algorithm that calculates the optimization given in Eq. 5.

To investigate the performance of the proposed method, a synthetic laplacian source was watermarked at a rate of 1/128 bits per source sample. Three different methods were used. First, a standard spread spectrum embedding process was followed (with equal watermark energy given to each source sample) and a correlation detector was used to read the watermark. Second, the same method of embedding was used, but a maximum-likelihood detector was used. Finally, the proposed host aware method was used to embed the watermark, and the same maximum-likelihood detector from the previous method was used for reading. Results are shown in Figure 4. Various watermark strengths were tested; the source to watermark distortion ratio (SDR) was varied from 6 to 30 dB. It can be seen that the host aware embedding provides superior results. While the host aware method was tested at 6 and 12 dB SDR, no errors were observed. The tests were repeated, this time with AWGN noise added to the watermarked source so that the source to added noise ratio was 4 dB. Results of these tests are shown in Figure 5. When noise is added to the watermarked images, the performance improvement over the first system provided by the host aware embedding is reduced. This is also true for the system using ML decoding. Note that both of these systems could be modified to explicitly take the added noise into account.

To investigate performance in a more realistic application, a set of 12 512x512 images was watermarked in the wavelet domain using the Daubechies wavelet of length 8. Only the second level of detail subbands were used; i.e. three 128x128 blocks of subband coefficients were available for marking. These coefficients were modeled as having a Laplacian distribution, and the three watermarking cases explored above for symmetric sources were repeated for the image set. The results are shown in Figure 6, for a number of different payload sizes. In all cases the watermark signal variance was set to 8. Note that the improvement provided by the host aware embedding technique is much smaller than for the previous synthetic source example. This is most likely because the Laplacian model used was not a close match to the host distribution. It was noted empirically that the subband distributions were more strongly peaked than a Laplacian distribution.

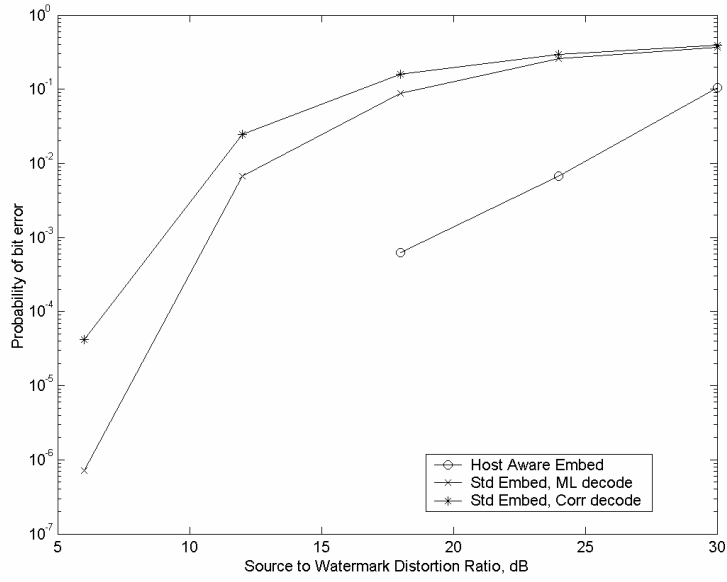


Figure 4: Watermarking of synthetic Laplacian source.

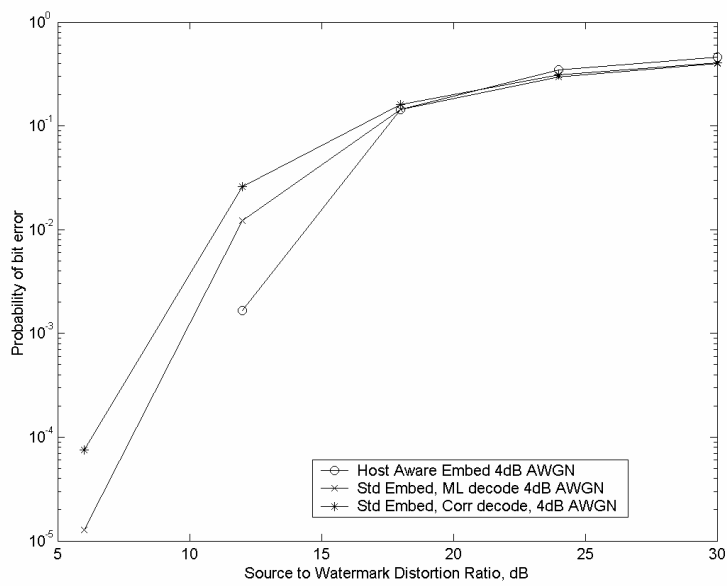


Figure 5: Watermarking of synthetic Laplacian source, followed by addition of  $-4$ dB AWGN.

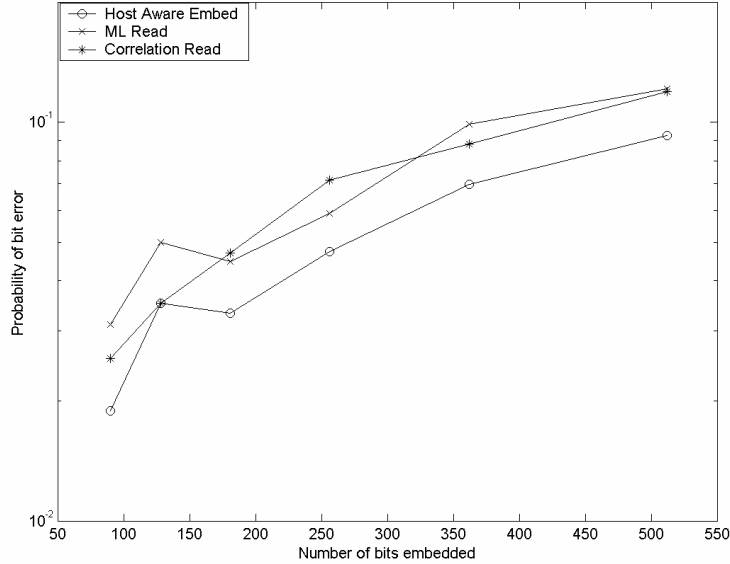


Figure 6: DWT Watermarking of 512x512 image set.

## 5. APPLICATION TO PRE-FILTERING

A simple linear prefilter was used to estimate the watermark signal in [3]. This filter calculated the difference of each image pixel with the average of a cross-shaped spatial neighborhood of that pixel. If this difference is clipped near the magnitude of the watermark signal, it is possible to suppress large spatial variations in the host image. If  $y(i, j)$  is a pixel of the received image, such a clipped filter can be specified as

$$\hat{w}(i, j) = f_c \left( y(i, j) - \frac{1}{4} [y(i-1, j) + y(i, j-1) + y(i+1, j) + y(i, j+1)] \right),$$

$$m : x > m$$

$$\text{where } f_c(x) = x : -m \leq x \leq m.$$

$$-m : x < -m$$

This prefilter can be used to estimate the watermark signal in the marked image. Host aware embedding can take the filter design into account and vary the embedded watermark signal to maximize the watermark signal's correlation with the PN sequence at the output of the prefilter. This can be done simply by taking the function  $f_D$  from Eq. 5 equal to the filter used in the decoder.

Spatial spread spectrum systems using the above nonlinear filter were tested on the same set of 12 images used in the previous section. Three similar variations were examined. First, a simple spread spectrum embedder and a correlation detector with no prefilter was tested to provide a base for comparison. Second, this system was tested with the nonlinear filter added to the decoder prior to the correlator. Finally, a system using a host aware embedder to maximize the watermark signal at the output of the nonlinear filter was tested. Results for a variety of payload sizes are shown in Figure 7. In each case, the watermark signal variance was set to 1. These results indicate that the system using the nonlinear filter provides significantly lower probability of error than the system with no filter. Adding the host aware embedder to the system provides further significant reduction in probability of error. The host aware system is able to support payloads from two to four times the payload of a system using only the prefilter with comparable probability of error.

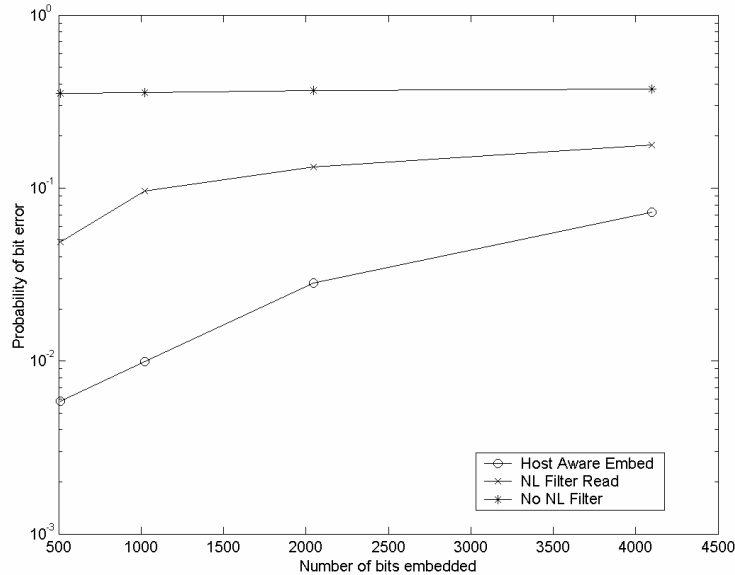


Figure 7: Spatial watermarking of 512x512 image set.

## 6. CONCLUSIONS

A new variant of spread spectrum watermarking has been proposed, in which the embedder is “host-aware”. By using knowledge of the host signal to vary the embedded watermark strength at each host sample, the host aware embedder is able to maximize the detection statistic used for decoding the watermark. This was demonstrated for two different watermarking systems, one in the DWT domain and using maximum likelihood detection, and the other in the spatial domain and using a nonlinear prefilter. In the case of the system using maximum likelihood detection, the host aware embedder was shown to provide significant improved performance on a synthetic source. The performance improvement decreased, however, when used with realistic image sources. This was likely due to inadequate modeling of the image source. Another factor that lowered the performance advantage of the host aware embedder was the addition of Gaussian image noise to the watermarked images.

One area for possible future investigation of host aware embedders is explicit inclusion of channel noise in the design of the embedder. Another approach is to design the host aware embedder to maximize with respect to an estimate of watermark robustness, rather than with respect to a detection statistic[9].

## REFERENCES

1. J. Cox, J. Killian, F. T. Leighton and T. Shanon, “Secure Spread Spectrum Watermarking for Multimedia”, *IEEE Transactions on Image Processing*, **6**(12):1673-1687, Dec. 1997.
2. G. Depovere, T. Kalker and J.-P. Linnartz, “Improved Watermark Detection Reliability Using Filtering Before Correlation”, *Proceedings of ICIP 98*, pp. 430-434.
3. U. Uludag, B. Gunesel, A. M. Tekalp, “Robust Watermarking of Busy Images”, *Proceedings of SPIE: Security and Watermarking of Multimedia Contents III*, Jan. 2001, vol. 4314, pp. 18-25.
4. M. Kutter, S. Voloshynovskiy, and A. Herrigel, “The Watermark Copy Attack”, *Proceedings of SPIE: Security and Watermarking of Multimedia Contents III*, 2000, vol. 3971, pp. 371-380.
5. H. S. Malvar and D. A. Florencio, “An Improved Spread Spectrum Technique for Robust Watermarking”, *Proceedings of ICASSP 2002*, vol. IV, pp. 3301-3304.
6. B. Chen and C.-E. Sundberg. “Digital Audio Broadcasting in the FM Band by Means of Contiguous-band Insertion and Pre-canceling Techniques,” *IEEE Transactions on Communications*, **48**(10):1634-1637, 2000.

7. N. Tanabe and N. Farvardin, "Subband image coding using entropy-coded quantization over noisy channels," *IEEE Journal on Selected Areas in Communications*, vol. **10**, no. 5, 926-943, 1992.
8. R. J. Clarke, *Transform Coding of Images*, Academic Press, New York, 1985.
9. I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*, Morgan Kaufmann Series in Multimedia Information and Systems, 2002.
10. S. Glisic and B. Vucetic, *Spread Spectrum CDMA for Wireless Communications*, Artech House, Norwood, MA, 1997.
11. A. Viterbi, *CDMA Principles of Spread Spectrum Communication*, Addison-Wesley, Reading, MA, 1995.
12. J. R. Hernandez and F. Perez-Gonzalez, "Statistical Analysis of Watermarking Schemes for Copyright Protection of Images", *Proceedings of the IEEE*, Vol. **87**, No. 7, Jul 1999, 1142-1165.