Dresden, Germany, 1999.
To be published in:
"Lecture Notes in Computer Science"

Performance Improvement of Spread Spectrum Based Image Watermarking Schemes Through M-ary Modulation

Martin Kutter

Swiss Federal Institute of Technology (EPFL)
Ecublens, 1014 Lausanne, Switzerland
martin.kutter@epfl.ch,
WWW home page: http://ltswww.epfl.ch:1248/kutter

Abstract. In spread spectrum communication the use of M-ary modulation is a popular extension to binary signaling usually resulting in a significant performance improvement. Furthermore it was shown that for increasing M and for certain schemes M-ary modulation works at the channel capacity. In this work we investigated on how to use M-ary modulation in the context of spatial spread spectrum based watermarking schemes. The performance of M-ary watermarking schemes is theoretically analyzed and the results verified through empirical test. We show that in general M-ary modulation based on the biorthogonal extension of the Hadamard matrix results in a significant performance improvement for values of $M > 4$. Furthermore we show that the performance improvement does not decrease under noise like distortion such as lossy JPEG compression.

1 Introduction

Digital watermarking, the art of hiding information in multimedia data, has become a very popular research area for the past 7 years. The reason for its fast growing interest is mainly due to potential applications addressing issues such as copyright protection, data monitoring, and data authentication. For a in-depth introduction to the topic and an overview of current state of the art techniques for a variety of multimedia data, the reader is referred to [1].

Most digital image watermarking methods employ some kind of spread spectrum modulation [8, 10]. This approaches have proven to be efficient, robust and cryptographically secure. From communication theory, we know that the performance and reliability of modulation schemes may be increased through channel coding. However, this concept is not very efficient if applied to watermarking techniques because in watermarking the message is usually of limited size (such as 64 or 128 bits) and not continuous. Furthermore the channel distortion is very variable because it depends on the image size, the image content, as well

as processing applied to the watermarked image. It is therefore very difficult to identify appropriate code lengths [7].

Another approach to increase the performance of communication schemes is based on M-ary modulation [10]. Often in communication, a binary information is transmitted on a bit by bit basis, usually referred to as *binary signaling*. In such schemes, the smallest information entity is a bit which can take on two values. In communication the states of the smallest piece of information transmitted are called *symbols*. For binary signaling, we have two symbols. The idea of M-ary modulation is to increase the size of the transmitted information, resulting in an increase of the number of different symbols. The M in M-ary refers to the number of symbols used in the communication scheme. Therefor, binary signaling is a special case of M-ary modulation with $M = 2$. In general for large values of M this schemes result in a significant performance improvement. For small values of M, the performance does not necessarily increase, and even decreases in some cases.

In this paper we investigate the performance of M-ary modulation in the context of digital image watermarking. We start by reviewing the common spatial spread spectrum watermarking scheme in Sec. 2. In Sec. 3 we look at the extension of this scheme for M-ary modulation based on biorthogonal modulation functions generated using the Hadamard matrix. Results will be shown in Sec. 4 and conclusions are drawn in Sec. 5.

2 Spatial Spread Spectrum Watermarking

We start the technical description of the proposed method, by first introducing the underlying digital watermarking scheme for binary signaling, that is $M = 2$. This scheme is based on a previous work [6, 5] and can be viewed as some sort of spread spectrum watermarking with the difference that, in the watermark recovery process, an additional step which predicts the embedded watermark is introduced to increase detector performance.

2.1 Watermark Embedding

The generic watermark embedding system is depicted in Fig. 1. Our goal is to embed an N bit long information $B = \{b_1, b_2, \ldots, b_N\}$ in an image I. The image I defines a set of pixels used to represent the visual information. Each pixel is a sets of three tristimulus values:

$$I(x, y) = \{R(x, y), G(x, y), B(x, y)\} \tag{1}$$

where $(x, y) \in \mathbb{Z}^2$ is the spatial location in the Cartesian coordinate system, and $R(x, y)$, $G(x, y)$, $B(x, y)$, are the three tristimulus values in a color coordinate system such as the *red-green-blue* color coordinate system [3].

The watermark embedding process takes place in the watermarking space $\mathcal{E}$. To project an image into the watermarking space the transformation χ is

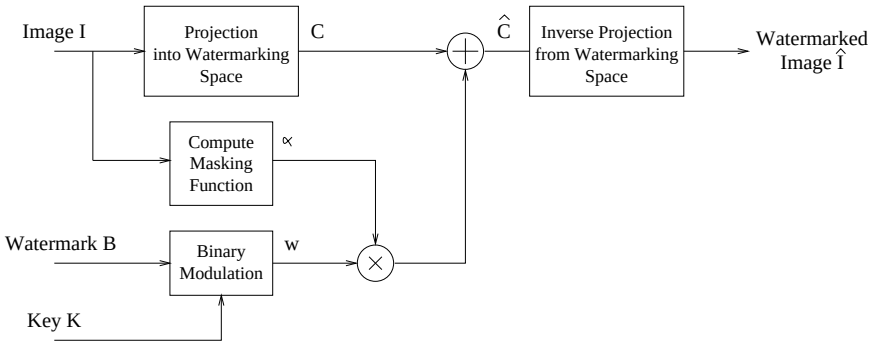


Fig. 1. Watermark embedding system for binary signaling.

applied to the image, i.e. $\chi : (I) \rightarrow (C)$, where C is the projected image. Once the watermark is embedded in the watermarking space, the inverse projection is applied to compute the watermarked image, that is $\chi^{-1} : (\hat{C}) \rightarrow (\hat{I})$. For the moment the only constraint on the projection is that it has to maintain the spatial dimensionality of the input data. That is, the image representation in the watermarking space has to be two dimensional.

In binary signaling we have two symbols Y^0 and Y^1, corresponding to the bit values of 0 and 1, respectively. It is well known, that for binary signaling, the optimal modulation functions are antipodal signal pairs [10, 4], that is $s^0 = -s^1$, where s^0, and s^1 are the signals corresponding to the symbols Y^0 and Y^1, respectively. For the embedding of the N bit watermark we use a set S^k of N two dimensional orthogonal functions $s_i, i = \{1, \ldots, N\}$, where k defines the secret key used as initializing seed to generate the set. Each function s_i in the set is used to represent one bit value of the watermark. In our case the functions can be considered as uniformly distributed random sets of random variables. The random variables are all independent and have a zero mean bilevel distribution with unit variance. In order not to introduce initial *inter symbol interference (ISI)*, the orthogonal functions are designed such that they are non overlapping. That is, if $S_i = \{(x, y), \forall s_i(x, y) \neq 0\}$ is the set of all locations for which the function s_i has non-zero values, then the intersection of all sets of non zero locations is the empty set, $S_i \cap S_j = \emptyset, \forall i \neq j$.

The watermark w is defined as the superposition of all modulated and weighted functions s_i:

$$w(x, y) = \sum_{i=1}^{N} b_i' \alpha(x, y) s_i(x, y) \tag{2}$$

where $\alpha(x, y)$ is a local weighting factor with the purpose to adapt the watermark to the human visual system, and b_i' represents the bit value mapped from $\{0, 1\}$ to $\{-1, 1\}$. The watermarked image is now given by adding the watermark w

to the image representation in the embedding space $\mathcal{E}$ and applying the inverse transformation:

$$\hat{I} = \chi^{-1}(C + w). \tag{3}$$

The weighting function α depends on the visual characteristics of the original image I and the watermark embedding space. Finding the optimal weighting function is a very difficult and delicate issue because it depends on both, the image and the metric used to measure the distortion. Because in this work we focus on the performance improvements through M-ary modulation and not optimal weighting of the watermark, we use a simple function for α in order to facilitate tractability of the problem. This approach is a valid basis because using more sophisticated functions increases the watermark energy by maintaining the visual distortion and hence results in an increase of the overall system performance.

In order to remain as general as possible, we do not watermark all locations in the watermarking domain. To describe the occupancy in the watermarking domain we define the density D as:

$$D = \frac{|\{\bigcup_{i=1}^{N} S_i\}|}{|\{S\}|} \tag{4}$$

where $|\{.\}|$ refers to the set cardinality and S represents the universal set. The probability for a location to be assigned to any set is equiprobable and hence given by D/N. It is obvious that the density has a direct impact on the watermark robustness and should therefore be set to 1 all the time. Nevertheless, the density gives us additional control over the watermark visibility and there is a trade-off between the watermark visibility and robustness.

2.2 Watermark Detection

The introduced watermarking scheme can be seen as a modulation system in which the image acts as additive noise. If the statistics of the image were Gaussian, common models could be used to design efficient detectors. However, it is well known that the statistics of images are in general very difficult to model. Nevertheless, it is a common method in digital watermarking to use a linear correlator as detection statistics. In our case we slightly modify this scheme by introducing a pre-processing step to compute a prediction of the embedded watermark and hence improve detector performance.

The linear detector is shown in Fig. 2. Let $\hat{F}$ be the preprocessed watermarked image given by:

$$\hat{F} = \hat{I} * H \tag{5}$$

where the operator $*$ expresses convolution, and H is a space variant linear filter defined by its impulse response $h_{m,n}(x, y)$. The coordinates (x, y) define the spatial position and (m, n) the relative shift to the local position. We assume

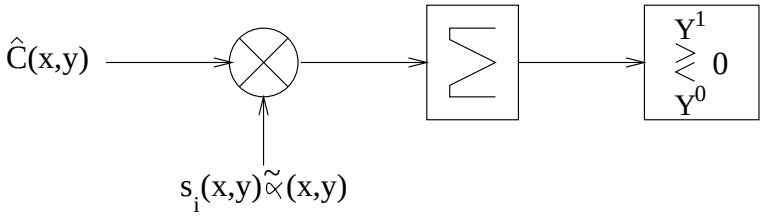


Fig. 2. Linear correlator detector with pre-processing step to predict the embedded watermark.

that the projection into the watermarking space is a linear operation with the property that we can inverse the operators:

$$\hat{C}_F = \chi(\hat{I} * H) = \chi(\hat{I}) * H \tag{6}$$

where $\hat{C}_F$ is the projection of the watermarked and filtered image $\hat{F}$ into the watermarking space. This property means that, convolving the projection of the image in the watermarking space is the same as projecting the filtered image into the watermarking space. Now we continue our analysis by expanding the watermarked and preprocessed image into its components:

$$\begin{aligned} \hat{C}_F &= (C + w) * H \\ &= C * H + w * H \end{aligned} \tag{7}$$

We use this expression to derive the detector statistic r_i for bit i conditioned on the modulation function s_i, that is $r_i|s_i$:

$$\begin{aligned} r_i|s_i &= < \hat{F}, s_i\tilde{\alpha} > \\ &= < C * H + w * H, s_i\tilde{\alpha} > \end{aligned} \tag{8}$$

$$\begin{aligned} &= \sum_{(x,y)} \left[\left(\sum_{(m,n)} w(x-m, y-n) h_{m,n}(x,y) \right) s_i(x,y)\tilde{\alpha}(x,y) \right] \\ &\quad + \sum_{(x,y)} \left[\left(\sum_{(m,n)} C(x-m, y-n) h_{m,n}(x,y) \right) s_i(x,y)\tilde{\alpha}(x,y) \right] \end{aligned} \tag{9}$$

where $< . >$ denotes the inner product operator. Because we do not have the original weighting function in the detection process, an estimate $\tilde{\alpha}$ of it is used.

Using this expression as a starting point we may now proceed to compute the expectation and variance of the detector statistic. After some algebra we find the expectation as:

$$E[r_i] = b'_i \frac{D}{N} \sum_{(x,y)} E[\alpha(x,y)\tilde{\alpha}(x,y)]\, h_{0,0}(x,y) \tag{10}$$

and the variance:

$$\begin{aligned} Var[r_i] = {} & \frac{D}{N} \sum_{(x,y)} E\left[C_F^2(x,y)\tilde{\alpha}^2(x,y)\right] \qquad (11) \\ & + \frac{ND - D^2}{N^2} \sum_{(x,y)} E\left[\alpha^2(x,y)\tilde{\alpha}^2(x,y)\right] h_{0,0}^2(x,y) \\ & + \frac{D^2}{N} \sum_{(x,y)} \sum_{(m,n)\neq 0} E\left[\tilde{\alpha}^2(x,y)\alpha^2(x-m,y-n)\right] h_{m,n}^2(x,y) \\ & + \frac{D^2}{N^2} \sum_{(x,y)} \sum_{(m,n)\neq(x,y)} E[\alpha(x,y)\tilde{\alpha}(x,y)\alpha(m,n)\tilde{\alpha}(m,n)] \qquad (12) \\ & \qquad h_{x-m,y-n}(m,n) h_{m-x,n-y}(x,y) \end{aligned}$$

Looking at these results we can make several observations. The detector performance increases with increasing pulse size, that is the size of the image, and the expectation is inversely proportional to the watermark length and proportional to the embedding density. The variance has a major contribution from the filtered image. The variance terms due to the correlation of the watermark can be decreases by decreasing the embedding density.

Next we have to look at the error performance of the proposed scheme. We consider a symbol-by-symbol hard decoder and assume that there is no inter symbol interference. This approach is justified by the fact that ISI is by an order of magnitude smaller than the interference from the image data. Fig. 3 illustrates the error region for the binary detection process by showing the conditional probability density function $f(r|Y^k)$ of the detector statistic r. As illustrated, in our case the two conditional probabilities are symmetric about the origin $r = 0$. The symbol error probability P_S, which is equivalent to the bit error for binary signaling, is given by integrating the overlapping Gaussian tails:

$$P_S = \frac{1}{2} P(\epsilon|Y^0) + \frac{1}{2} P(\epsilon|Y^1) \tag{13}$$

where ϵ stands for the error. As we have seen, the probability function of the detector statistic, conditioned on a symbol, is the sum of non-Gaussian random variables. If the number of terms is large enough we can make use of the *central limit theorem* [2] and approximate the sum by a Gaussian distribution with mean and variance as derived in Eq. (10) and Eq. (11), respectively. Hence in our case, the conditional distribution of the detector statistic $f(r_i|s_i)$ is given by:

$$f(r_i|s_i) \simeq \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\frac{-(r_i-\mu_i)^2}{2\sigma_i^2}} \tag{14}$$

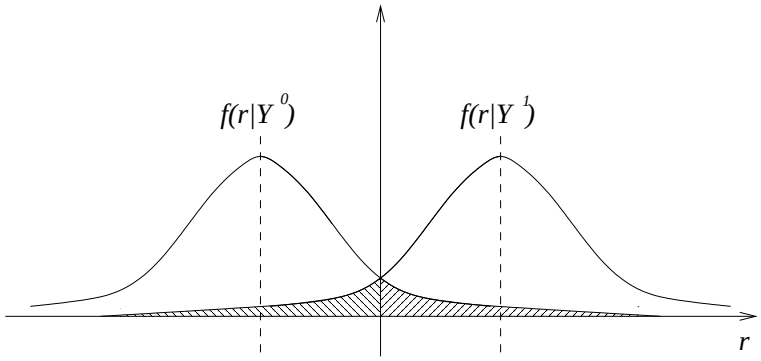


Fig. 3. Conditional probability functions for binary detection and error regions.

whereμ_i is the expectation as defined in Eq. (10) and σ_i^2 the variance as defined in Eq. (11);

The resulting symbol error probability P_S is given by:

$$\begin{aligned} P_S &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mu}^{\infty} e^{\frac{-x^2}{2\sigma^2}} dx \\ &= Q\left(\frac{\mu}{\sigma}\right) \end{aligned} \tag{15}$$

where $Q(.)$ is the so-called *Q-function*[1] defined as the tail integral of a Gaussian function with zero mean and unit variance:

$$Q(x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz \tag{16}$$

The symbol error probability is a very useful measure in digital communication because it represents the probability of detecting a symbol wrong. In digital watermarking this information is not that interesting because we do not have continuous data transmission and hence limited possibilities for efficient coding and error correction. We are mainly interested in the probability of detecting the embedded watermark without error, or the error probability of at least not detecting one bit. We denote the error probability of not detecting the right watermark by P_w. It is defined by 1 minus the probability of correctly detecting all bits:

$$P_w = 1 - (1 - P_S)^N \tag{17}$$

where P_S is again the symbol error probability.

[1] The Q-function is the complementary of the *error function* defined as: $erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-z^2} dz$.

3 M-ary modulation

Until now we have worked with antipodal modulation signals because they provide best performance for binary signal modulation in a noisy environment. From modulation theory we know that increasing the number of symbols may for certain modulation schemes result in a decrease of the symbol error probability. Furthermore, in the limit, that is for $M \to \infty$, we can design schemes working at the channel capacity. We will now investigate the concept of M-ary signaling in the context of digital watermarking. The idea for an improvement of the performance is given by the fact that the fewer symbols we have to hide in an image, the more locations we can use per symbol. In the previous sections we have seen, that the expected value of the detector statistic increases with the number of locations, often referred to as *pulse size* in spread spectrum modulation. Therefore, the larger the pulse size, the higher the probability of detecting the right symbol. Modulation and demodulation systems for M-ary signaling are shown in Fig. 4 and Fig. 5, respectively. The N bit long message, which can be considered as a two symbol signal, that is $M = 2$, is first mapped to M symbols required for the M-ary modulation. This is usually done by grouping $\log_2(M)$ bits of the original message and then taking the resulting decimal value as an index to select the appropriate function from a set of basis functions. In the decoding process the received signal is correlated with all modulation functions representing the different symbols. The index of the largest correlation determines the transmitted symbol.

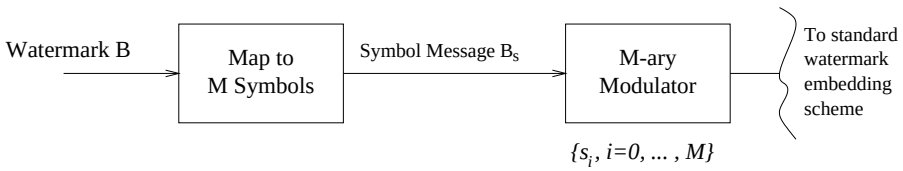


Fig. 4. M-ary modulation. The input message is fist mapped to M symbols by grouping $\log_2 M$ bits of the original message.

As mentioned, in M-ary signaling, every symbol conveys m bits, that is $M = 2^m$. Each symbol is represented by a bilevel spread spectrum modulation function from a basis set containing M functions. In general, the functions in a set are either *orthogonal*, *biorthogonal*, or *transorthogonal* [10]. The orthogonal case is the most obvious selection because it inherently results in orthogonal symbols, and therefore watermarks orthogonal to each other. One way to generate orthogonal function sets is based on the *Hadamard-Walsh functions*[9]. Given any spread spectrum function of length M, we may generate a set of M orthogonal functions by multiplying the initial spread spectrum function with the Hadamard matrix of order M. The smallest possible Hadamard matrix is of order 2 and is defined as:

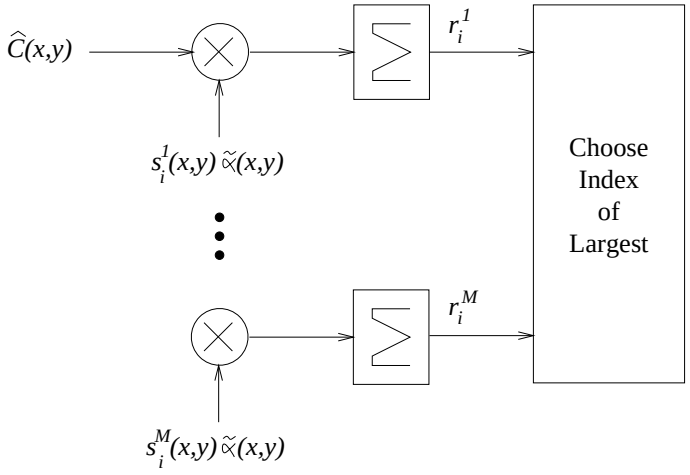


Fig. 5. M-ary demodulation scheme for orthogonal signals. The received signal is projected onto all symbol functions and the index of the largest is used to decode the message.

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{18}$$

Any other Hadamard matrix with an order of a power of 2 may then be iteratively constructed from the $\mathbf{H}_2$ matrix. The construction is based on the *Kronecker* product recursion and given by:

$$\mathbf{H}_N = H_2 \otimes H_{N/2} = \begin{bmatrix} \mathbf{H}_{N/2} & \mathbf{H}_{N/2} \\ \mathbf{H}_{N/2} & -\mathbf{H}_{N/2} \end{bmatrix} \tag{19}$$

As an example, this construction leads us to the following Hadamard matrix of order 4:

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{20}$$

Inspecting this matrix, we can first verify that all rows are orthogonal. Furthermore, it is interesting to note that the first value in each row is 1. This means that for the purpose of modulation this information is redundant and may therefore be dropped without any loss in performance. The detector statistics for a correlator, assuming the right symbol is given, follows the expression derived in the previous sections, with the difference that it does not depend on the bit value anymore. We have seen in Sec. 2 that for antipodal sequences the distance between the two conditional expectations for bit values of 0 and 1 is

twice the expected value of one conditional expectation. This is due to the fact that the antipodal signal design provides largest distance for fixed energy. For an orthogonal signal design this is not the case because the expectation for a wrong symbol is 0, which means that the distance between the symbols is only once the mean of the conditional expectation. This fact has been realized long ago and it was proposed to use biorthogonal basis sets for M-ary modulation.

Sets of biorthogonal basis functions can be generated from orthogonal sets by adding the sign reversed versions of the functions. This approach is very similar to the design of the antipodal functions. In our case, if we want do design an M-ary modulation scheme, where M is a power of 2. We start by computing an $M/2$ order Hadamard matrix $\mathbf{H}_{M/2}$. This matrix is then completed by appending its sign-reversed version to generate the M biorthogonal matrix entries:

$$\mathbf{B}_M = \begin{bmatrix} \mathbf{H}_{M/2} \\ -\mathbf{H}_{M/2} \end{bmatrix} \tag{21}$$

where $\mathbf{B}_M$ is the final matrix used to generate the set of M biorthogonal functions from one initial function. Besides the superior performance of biorthogonal function sets, there are also other advantages over orthogonal function sets. For example the number of correlators in the detection process is only half the number required for orthogonal functions. This issue is especially important for large values of M. The detector for biorthogonal M-ary signaling is shown in Fig. 6

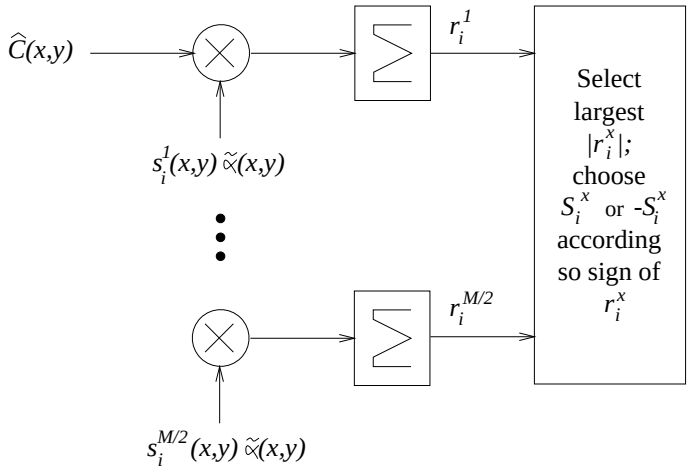


Fig. 6. M-ary demodulation scheme for biorthogonal signals. The received signal is projected onto the non sign reversed modulation functions. The symbol is found by locating the correlation with the largest absolute value and then choosing the symbol corresponding to the sign of the correlation.

To analyze the performance of M-ary biorthogonal signaling, we start by computing the probability of detecting a symbol right. Let us assume we transmitted symbol Y^0. The probability of correct detection P_C is defined as the probability that the absolute value of the detector correlation for Y^0 is larger than zero, and all other correlations are smaller than the correlation for Y^0;

$$P_C = p(abs(r_i^1) < r_i^0, \ldots, abs(r_i^{M/2-1}) < r_i^0 | r_i^0 > 0) \tag{22}$$

where r_i^n is the detector statistic for the n_{th} of the i^{th} M-tuple of bits from the encoded message. Under the assumption of large pulse sizes we may again use the central limit theorem and conclude that the distribution of r_i^0 approaches a Gaussian distribution with mean and variance as derived in the previous sections. The distribution of the random variable $r_i^n, n = 1, \ldots, M/2-1$, is as well Gaussian with the same variance as for r_i^0, but with zero mean. Assuming that there is no inter symbol interference the random variables r_i^n are independent, which means that their joint distribution is the product of the individual probability density functions and because they all have the same distribution we can raise is to the power of $M/2-1$. The resulting probability of correct detection is then given by:

$$P_C = \int_0^\infty \frac{e^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}}}{\sqrt{2\pi\sigma_0^2}} \left[1 - 2Q\left(\frac{x}{\sigma_0}\right)\right]^{\frac{M}{2}-1} dx \tag{23}$$

The 2 in front of the $Q(.)$ function is due to the absolute value of r_i^0 which means that to compute the probability $p(abs(r_i^1) < r_i^0)$ we have to integrate the Gaussian distribution from $-x$ to x and this is equivalent to twice the integral from 0 to x.

The symbol error probability P_s is then given by:

$$P_s = 1 - P_C \tag{24}$$

It should be clear to the reader that the symbol error probability is different from the bit error probability. For biorthogonal signaling two distinct symbol errors can occur. Either $-s_i$ is selected instead of s_i or one of the $M-2$ signals orthogonal to s_i. Several procedures are possible for bit-labeling, that is the assignment of grouped bits to the corresponding symbol. In general, and in order to decrease bit error probability the assignment for antipodal biorthogonal signaling assigns the antipodal signal pair with complementary label. For example in 8-ary signaling an antipodal signal pair would be assigned to the bit triple "010" and its complement "101". Assuming the error occurs of selecting $-s_i$ instead of s_i, the complementary bit-labeling results in decoding all bits wrong. However, this case is relatively rare because the distance between the antipodal signals is twice the distance to the orthogonal signals. For the second and more probable error, there are $M-2$ equally probable decision errors and $\frac{M-2}{2}$ have bit discrepancies with Y_i^0 in any given position. The bit-error rate P_b is therefore bound by:

$$\frac{P_s}{2} < P_b \leq P_s \tag{25}$$

where P_s is the symbol error probability.

As we have argued earlier, in digital watermarking applications we are more interested in the probability P_w of not correctly detecting the entire embedded watermark. Taking into account that the number of symbols decreases with increasing M and following the same strategy used to derive Eq. (17) we can derive the following expression for P_w:

$$\begin{aligned} P_w &= 1 - P_C^{N/\log_2(M)} \\ &= 1 - \left\{ \int_0^\infty \frac{e^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}}}{\sqrt{2\pi\sigma_0^2}} \left[1 - 2Q\left(\frac{x}{\sigma_0}\right) \right]^{\frac{M}{2}-1} dx \right\}^{N/\log_2(M)} \end{aligned} \tag{26}$$

To evaluate this expression we have to resort again to numerical methods.

4 Results

In order to illustrate the behavior of the detection error probability, we show the detection error as a function of the embedding density for different values of M. Usually in modulation the symbol error is shown as a function of the signal to noise ratio for different values of M. However, this approach is not suitable for watermarking systems because both, the signal and noise depend on the embedding density. The watermarking space is defined as the gray scale representation of the image. The weighting function $\alpha(x,y)$ is defined as a fraction of the luminance and a minimum offset, that is $\alpha(x,y) = C(x,y)/128 * \beta + \delta$,where β defines the embedding strength and δ the offset. The term 128 was only introduced to normalize the luminance and has no functional impact. The tests were performed using the following parameter settings: $\beta = 4$, $\delta = 1$, $N = 240$. Each test was repeated 100 times and using a different key for each test. In the detection process we use a cross-shaped prediction filter [6, 5] as motivated below with its coefficients defined as:

$$H_+^{\mathcal{W}} : h_+(x,y) = \begin{cases} 1 & x = 0, y = 0 \\ -\frac{1}{2\mathcal{W}-2} & x = 0, 0 < |y| \leq \frac{\mathcal{W}-1}{2} \\ -\frac{1}{2\mathcal{W}-2} & y = 0, 0 < |x| \leq \frac{\mathcal{W}-1}{2} \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

where $\mathcal{W}$ defines the filter size. For the tests the size was set to $\mathcal{W} = 7$. Other filters are also possible, For example filters based on the adaptive Wiener filter result in an impressive performance increase. However, the goal of this work is to investigate the impact of M-ary modulation and for this purpose the

cross-shaped prediction filter provides a good trade-off between computational complexity and performance.

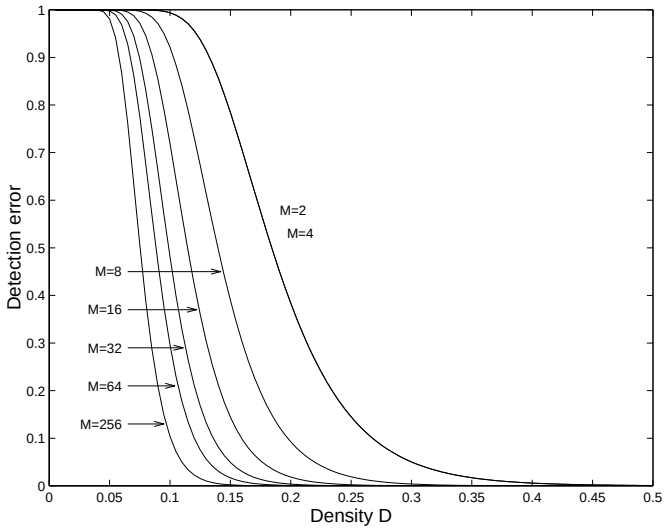
Fig. 7(a) shows the theoretical detection error probability as a function of the density D. Changing the density may actually be considered as some sort of changing the signal to noise ratio. The curves clearly show the significant performance improvement for large values of M. For small values of $M \leq 4$ the performance is approximately equivalent. This results agree with the theoretical performance results known from standard M-ary modulation [10] in that small values of M do not necessarily result in a performance increase.

Empirical detector results are shown if Fig. 7. The curves show an interpolation of the data points. We can clearly see the relative correspondence of the theoretical results shown in Fig. 7(a) and the empirical results in Fig. 7(b). Again, for small values of M, the performance does not increase, or decreases even for some densities. However, as soon as M is larger than 4, the detector performance increase significantly.

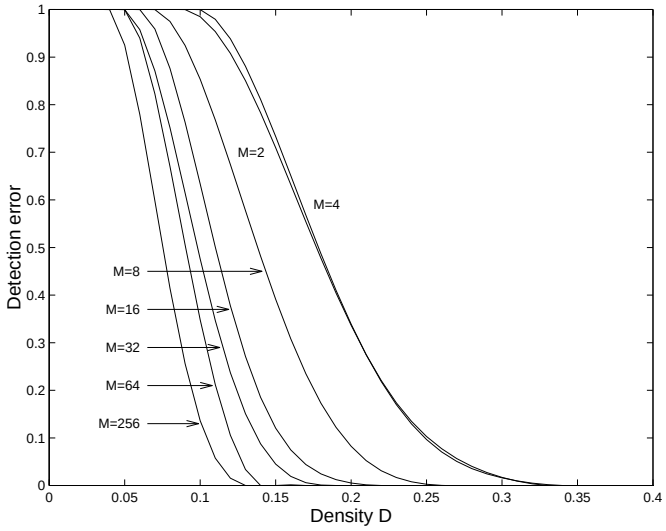
As a last test set Fig. 8 shows the watermark detection error for the *lena* and *baboon* images using M-ary modulation and lossy JPEG compression. The watermark embedding was performed using the same parameters as in the previous example. The density was fixed to $D = 0.5$. The graphs clearly illustrate the significant improvement with increasing values of M even under large JPEG compression. As an example, for a detection error probability of 0.1, the compression quality goes down to 25% for $M = 256$, and only 55% for $M = 2$.

5 Conclusions

The concept of M-ary modulation has been applied to spatial spread spectrum watermarking. The theoretical analysis using a linear correlator statistic with a preprocessing step prior to the detection have been derived. The theory was then verified with empirical results. We may conclude that in general M-ary modulation may significantly improve the robustness of digital watermarking schemes for values of M larger than 4. As we have seen, increasing values of M also increases the overall performance. However, it should be noted that large values of M result in an increased demodulation time. For real applications values of M between 8 and 64, in some cases 256 seem appropriate. Values larger than 256 would in general require to much processing time and are hence not suitable for digital watermarking applications. Comparing M-ary modulation to channel coding we may conclude that in any case M-ary modulation seems more appropriate for watermarking application because it results in any case in a performance improvement for values of $M > 4$. As opposed to M-ary modulation, in channel coding [7] the issue is much more delicate and the performance improvement is not guaranteed because it depends on the pulse size, as well as the watermark length and code length.

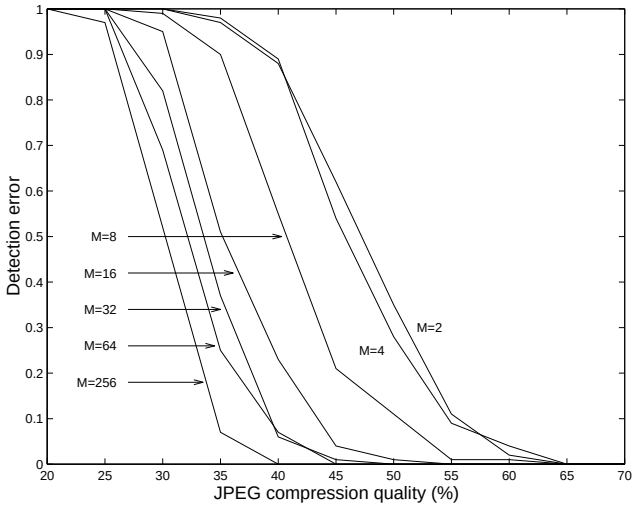


(a) Theoretical results.

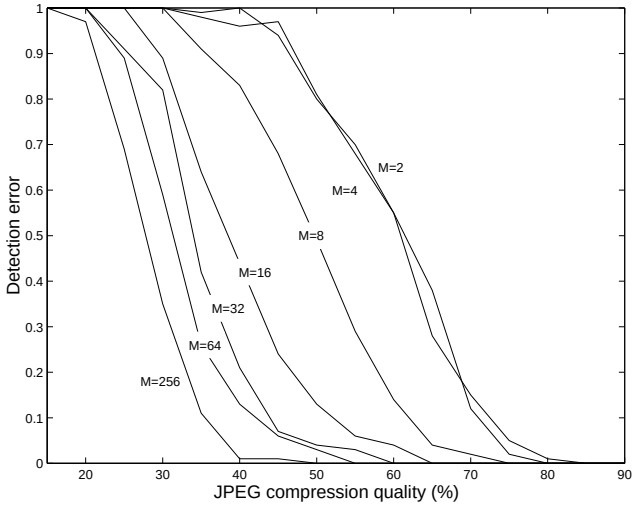


(b) Empirical results.

Fig. 7. Theoretical end empirical watermark detection error as function of the embedding density D for the *lena* image. For small values of M the detector performance is approximately the same. However, with values of M large than 4, the detector performance increases significantly with increasing M.



(a) *lena*



(b) *baboon*

Fig. 8. Empirical watermark detection error probability for the *lena* and *baboon* image using M-ary modulation. The curves show the detector improvement for large values of M. The detection error largely improves with increasing M, with exception for very small values of M.

References

[1] Frank Hartung and Martin Kutter. Multimedia watermarking techniques. *Proceedings IEEE:Special Issue on Identification and Protection of Multimedia Information*, July 1999.

[2] Carl W. Helstrom. *Probability and Stochastic Processes for Engineers.* Macmillan, 1991.

[3] Anil K. Jain. *Fundamentals of Digital Image Processing.* Prentice-Hall, 1989.

[4] Saleem A. Kassam. *Signal Detection in Non–Gaussian Noise.* Springer–Verlag, 1998. Chapter 2.

[5] Martin Kutter. Watermarking resisting to translation, rotation, and scaling. In Andrew G. Tescher, Bhasjaran Vasudev, V. Michel Bove, and Barbara Derryberry, editors, *Multimedia Systems and Applications*, volume 3528, pages 523–431, San Jose, CA, USA, November 1998. IS&T, The Society for Imaging Science and Technology and SPIE, The International Society for Optical Engineering, SPIE.

[6] Martin Kutter, F. Jordan, and Frank Bossen. Digital watermarking of color images using amplitude modulation. *Journal of Electronic Imaging*, 7(2):326–332, April 1998.

[7] J.R. Hernández F. Pérez-González and J.M. Rodríguez. The impact of channel coding on the performance of spatial watermarking for copyright protection. In *International Conference on Acoustic, Speech and Signal Processing (ICASP)*, volume 5, pages 2973–2976, May 1998.

[8] Raymond L. Pickholz, Donald L. Schilling, and Laurence B. Milstein. Theory of spread-spectrum communications–a tutorial. *IEEE Transactions on Communications*, 30(5):855–884, May 1982.

[9] Andrew J. Viterbi. *CDMA: Principles of Spread Spectrum Communication.* Addison-Wesley Publishing Company, 1995.

[10] Stephen G. Wilson. *Digital Modulation and Coding.* Prentice Hall, 1996.