Open access • Proceedings Article • DOI:10.1117/12.453534

Evaluating the optimal probability distribution for steganography under zero-error conditions — Source link

Gareth Brisbane, Reihaneh Safavi-Naini, Philip Ogunbona

Institutions: University of Wollongong, Motorola

Published on: 01 Jan 2003

Topics: Alice and Bob and Steganography

1-1-2002

Evaluating the optimal probability distribution for steganography under zero-error conditions

Gareth Brisbane
University of Wollongong

Reihaneh Safavi-Naini
University of Wollongong, rei@uow.edu.au

Philip Ogunbona
Motorola Australia Research Centre, philipo@uow.edu.au

Evaluating the optimal probability distribution for steganography under zero-error conditions

Abstract

Information hiding can be performed under the guise of a digital image. We consider the following scenario: Alice and Bob share an image and would like to use it as a cover image to communicate a message m. We are interested in answering two questions: What is the maximum amount of information that can be sent for a given level of degradation to an image? and How can this level of efficiency be achieved in practice? We require the recovered message to be the same as the embedded one. Our model begins with Alice compressing a message to obtain a binary sequence with uniform distribution. She then converts the binary sequence into a Q-ary sequence having a pre-defined distribution, and finally adding each symbol to a pixel. The distribution of the Q-ary sequence is chosen such that the amount of information is maximized for a given value of the signal to noise ratio. Bob recovers the sequence by subtracting the image data, and then converting the Q-ary string into the original binary string. We determine the optimal distribution analytically and provide a graphical representation of the variation of the amount of information with signal-to-noise ratio when Q varies.

Keywords

evaluating, steganography, distribution, conditions, probability, error, optimal, zero, under

Disciplines

Physical Sciences and Mathematics

Publication Details

Brisbane, G., Safavi-Naini, R. & Ogunbona, P. (2002). Evaluating the optimal probability distribution for steganography under zero-error conditions. Proceedings of SPIE - Mathematics of Data/Image Coding, Compression, and Encryption V, with Applications (pp. 145-155). The International Society for Optical Engineering.

Evaluating the optimal probability distribution for steganography under zero error conditions

Gareth Brisbane[a], Rei Safavi-Naini[a] and Philip Ogunbona[b]

[a]School of IT and CS, University of Wollongong, NSW, Australia
[b]Motorola Australia Research Centre, Sydney, NSW, Australia

ABSTRACT

Information hiding can be performed under the guise of a digital image. We consider the following scenario: Alice and Bob share an image and would like to use it as a cover image to communicate a message m. We are interested in answering two questions: *What is the maximum amount of information that can be sent for a given level of degradation to an image?* and *How can this level of efficiency be achieved in practice?* We require the recovered message to be the same as the embedded one.

Our model begins with Alice compressing a message to obtain a binary sequence with uniform distribution. She then converts the binary sequence into a Q-ary sequence having a pre-defined distribution, and finally adding each symbol to a pixel. The distribution of the Q-ary sequence is chosen such that the amount of information is maximized for a given value of the signal to noise ratio. Bob recovers the sequence by subtracting the image data, and then converting the Q-ary string into the original binary string.

We determine the optimal distribution analytically and provide a graphical representation of the variation of the amount of information with signal-to-noise ratio when Q varies.

Keywords: Steganography, image escrow, optimal, capacity

1. INTRODUCTION

The most common illustration for representing the various nuances of the steganographic problem was posed by Simmons.[1] He allegorized the problem as the desire for two prisoners, Alice and Bob, to communicate with each other. They had already anticipated their arrival in jail and so have already shared a short secret. Their messages are couriered by agents of the warden, Wendy. Wendy knows that they will try to co-ordinate their escape but wishes to catch them in the act. For this to happen, she allows communications in the hope that she will identify messages which contain convicting information. Alice and Bob, aware of this restriction, hide *stego-text* (hidden messages) within *cover-text* (an innocuous medium).

The subset of the problem which we are interested in is the *passive warden*. No alterations are made to the coversignal as an *active warden* could, nor send additional messages as might a *malicious warden*. Instead, the passive warden will only prohibit messages in the event that they do not have the appearance of a normal message.

The specific type of message we examine are digital images. Images have potential for information hiding due to the redundancy and irrelevancy of the image data where the latter is because of the limitations of Human Visual System (HVS). Compression algorithms exploit these properties to find a much shorter description of the data.

An information hiding system consists of two algorithms: an *embedding algorithm* where stego-text is embedded in a cover-text, and an *extraction* algorithm where the stego-text is extracted from a cover-text. We describe an embedding process as a *zero-error* algorithm if all bits of the embedded data can be losslessly extracted from its cover-text, assuming that it has not otherwise been modified. Information hiding systems can be broadly divided into *image escrow systems* in which extracting the embedded message requires the knowledge of the original image, and *oblivious systems* which do not require the original.[2]

E-mail: {gareth, rei}@uow.edu.au, philip.ogunbona@motorola.com

Mathematics of Data/Image Coding, Compression, and Encryption V, with Applications,
Mark S. Schmalz, Editor, Proceedings of SPIE Vol. 4793 (2003) © 2003 SPIE · 0277-786X/03/$15.00

A steganographic technique is evaluated with respect to three criteria[2]: *embedding capacity*; *robustness*; and *imperceptibility*. The embedding capacity of an information hiding technique is defined as the maximum amount of stego-text that can be embedded in a given cover-text. The robustness of a technique refers to the ability of the technique in recovering the embedded stegotext after the cover-text is modified (either invisibly or visibly). Finally the *imperceptibility* of the technique is a measure of its effectiveness with respect to hiding the stegotext. There is a trade-off between satisfying these criteria. That is, increasing the embedding capacity will lower robustness and imperceptibility, and reducing it will tend to increase the two.

1.1. Other works on capacity

1.1.1. Channel capacity for oblivious techniques

Let N denote the signal power of an image X. Marvel and Boncelet's aim is to find the maximum quantity of information that can be embedded in an image, when the embedding is additive and in the pixel domain (pixel by pixel).[3] They assume an oblivious system, which does not require X for extraction of the stego-data. By modelling data hiding as sending a signal, S, through a noisy signal (N), they can use the channel capacity expression derived by Shannon,[4]

$$C = \frac{1}{2}\log_2\frac{S+N}{N}$$

They noted that because the image data is highly correlated, it cannot be represented as additive Gaussian noise. They proposed the use of *equivalent white Gaussian* noise to obtain an upper-bound for the channel capacity. To obtain the white noise equivalent to an arbitrary noise signal, the entropy of the given noise signal, in this case, N, must be determined. This is done by using CALIC, an image compression algorithm.[5] The results showed that the potential for oblivious information hiding was image dependent, and ranged from approximately 0.25bpp to 3.4bpp for a Signal to Noise Ratio (SNR) of -30dB.

1.1.2. Channel capacity for cover image escrow techniques

Barni et al.[6] consider the capacity of a class of data hiding techniques that operate in the frequency domain, such as those proposed by Cox et al.[7] and Barni et al.[8] Though not explicitly mentioned, both of these techniques are escrow image systems. In these systems the transform coefficients, for example the DCT and DFT in the two mentioned cases respectively, are modified. The modification of each coefficient is proportional to the size of the coefficients. The authors assimilate each coefficient to a channel through which a component of the watermark is transmitted. They argue that for such channels, the noise is neither additive nor Gaussian and so the frequently used expression for channel capacity, reproduced above, cannot be used. They also propose practical numerical methods for evaluating the capacity, indicating that approximately 0.0055bpp can be hidden, though no mention is made of the quality of Z, the output image.

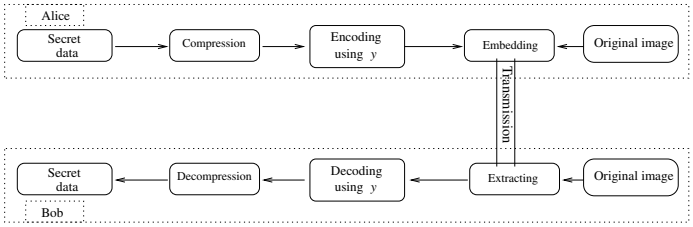
Ramkumar and Akansu[9] consider the capacity of the data hiding channel for both image escrow and oblivious systems and show that capacity can be substantially increased by decomposing the image using orthonormal transforms such as the DCT, Hartley, Hadamard and sub-band decomposition. They consider degradation mainly due to compression (followed by decompression) and show that the optimum choice of the transform depends on the required level of robustness.

2. THE INFORMATION HIDING MODEL

2.1. Definitions

We consider image escrow systems with zero error in recovery. Stego-data is embedded in the pixel domain, simply by adding it to the pixel values. We assume that the stego-data is a Q-ary string of symbols, where elements of the string are integers in the range $[-A, B]$. Elements of the stego-text data are added one-by-one to the pixels of the cover image. In extraction, the cover-text data is subtracted from the stego-text. We only consider grey scale images, although the results can be easily extended to images with more than one colour component.

Figure 1. Hidden communication through an image



In the case of an 8 bit scale, the addition operation is clipped to the range $[0, 255]$ and so if the sum is above 255, because of clipping the information will be lost. Similarly if the result is less than zero it will be rounded to zero and again the embedded data will be lost.

The main question is "How much information can be hidden within an image, assuming that the warden is passive, and given a maximum level of degradation to the image?".

Let an image X consist of pixels, $x_1, x_2, ..., x_{U \times V}$, where U is the number of rows and V is the number of columns. The value of $x_i \in [0, R], \forall i \in [1, U \times V]$. We assume that we are dealing with only grey scale images. If colour images are used, then if they are stored independently, the capacity for hiding information is increased by the number of extra colour planes*.

Initially, the *secret data* which Alice wishes to transmit is a binary sequence. It is then compressed to produce a *message*, which is a binary sequence with $p(0) = p(1) = 1/2$, a uniform distribution. It is a reasonable assumption that such a sequence will be the output of any optimal compression algorithm, so we assume the message to be embedded will be in this form. Our calculations only assume this distribution of the message to eliminate any bias in the secret data. To use the embedding technique described above, the binary sequence must be converted into a Q-ary sequence with probability distribution y that will be chosen to maximize the amount of embedded information.

The full diagram of the process is shown in figure 1.

2.1.1. Alice's components

The *coder* takes the binary uniformly distributed sequence and generates a Q-ary sequence with probability distribution y. It is assumed that the symbols are in the range $[-A, B]$, where $Q = B + A + 1$.

This coder can be implemented using an *entropy decoder* algorithm, such as *arithmetic coding* or *Huffman coding*. A binary entropy coder takes a sequence over an alphabet, Λ, together with a probability distribution, y, and produces a binary output that is uniformly distributed. The decoder performs the inverse: it uses the same model as the encoder to convert a binary, uniformly distributed input to the original sequence. With this description, it is clear that the coder of the information hiding system described above can be constructed by an entropy decoder, for example an *arithmetic decoder* whose parameters (size of the alphabet and associated probability distribution) are determined by the probability distribution required by the embedder.

Although it is theoretically possible to use other entropy coders with optimal performance, in practice the arithmetic coder provides the best performance.[10] Thus, the arithmetic coding algorithm is used for our experiments. The goal of the encoding section is therefore identifying the model that gives the best performance form a data hiding perspective. This can be equated with designing a source that matches a transmission channel, in this case, the image X.

The *embedder* adds the encoded message sequence to the pixel values of X. This is done by adding each symbol, m_i, derived from the encoding of the message sequence, to a pixel, x_i, in the image. That is, $z_i = x_i + m_i$. We assume that clipping takes place, i.e. $z_i = 0$ when $z_i < 0$ and $z_i = R$ when $z_i > R$. This results in the loss of embedded information in the clipped values.

*This is not entirely true, as compression of colour channels makes use of this correlation.

Embedding may be in all pixels, or a portion of the pixels subject to a specific criterion, known by both Alice and Bob. Embedding in all pixels is likely to introduce error in message recovery because of the clipping effect described above. This error can be avoided if a subset of the pixels were used, that is, all pixels with values in the range $[A, R-B]$. Thus, even if the most outlying numeric symbols, $-A$ and B, were to be embedded, z_i would still remain in the range $[0, R]$, ensuring that the embedding algorithm is error-free.

Modifying pixel values causes damage to the original image. A metric for measuring the damage is the Peak Signal to Noise Ratio(PSNR):

$$\text{PSNR} = 20\log_{10}\frac{255}{\sqrt{\text{MSE}}} \tag{1}$$

where the MSE is the Mean Squared Error. The PSNR is not always a good measurement[11] but it is the most commonly used one and will be used in the rest of this paper for measuring imperceptibility of embedding.

2.1.2. Bob's components

The *extractor* takes Z and X, and by subtracting pixel values, recovers the embedded message, using the formula $\hat{m}_i = z_i - x_i, \forall i \in [1, U \times V]$, where $\hat{m}_i$ is either m_i or an erroneous value of m_i due to clipping.

The *decoder* is the inverse of the encoder, converting the message symbols back to the form of the message sequence. If the entropy coder is the arithmetic coding algorithm, then decoding will only succeed if no errors are present in $\hat{m}$.

2.1.3. Error correction

As noted previously, a requirement is that the message can be recovered error-free. Error correction works by adding redundancy to data. We noted above that error free embedding is possible if a subset of points is used for embedding. An alternative method is to use an error correcting code to correct the errors due to clipping. That is, use an error correcting code in Alice's component, just before the embedding. However this will alter the probability distribution, y. We note that error correcting codes cannot be used before the encoder (entropy decoder) as the corresponding decoder will be after the entropy encoder. This means that the input to the entropy encoder will have errors and because of the sensitivity of such encoders to error, the whole output will be in error. In the rest of this paper we assume error freeness is obtained by restricting the embedding to a subset of symbols.

3. EMBEDDING CAPACITY

3.1. Definition of the problem

With the above model, we can state the problem of finding the embedding capacity as an optimization problem. That is, for a fixed level of distortion, measured by the PSNR, we want to find the range $[-A, B]$ and probability distribution y of the symbols which maximizes the amount of embedded information.

Let y_i denote the probability of symbol i where $i \in [-A, B]$. Finding the embedding capacity is equivalent to finding $y_i, i \in [-A, B]$ that maximizes $H(y) = -\sum_{i=-A}^{B} y_i \log y_i, \forall y_i$ for a given value of $M(y)$. That is

$$\texttt{Maximize}_y H(y)$$

subject to

$$M(y) = \sum_{i=-A}^{B} i^2 y_i = c \tag{2}$$

$$P(y) = \sum_{i=-A}^{B} y_i = 1 \tag{3}$$

$$0 \leq y_i \leq 1 \tag{4}$$

$H(y)$ represents the entropy of the distribution, $M(y)$ represents the MSE, and $P(y)$ is the term which equates the sum of the probability distribution.

3.2. Deriving the optimal distribution

Using LaGrange multipliers, the following simultaneous equations are derived:

$$\frac{\partial H}{\partial y_i} = \frac{\partial M}{\partial y_i}\lambda_M + \frac{\partial P}{\partial y_i}\lambda_P \quad (5)$$

$$\frac{\partial H}{\partial y_{-i}} = \frac{\partial M}{\partial y_{-i}}\lambda_M + \frac{\partial P}{\partial y_{-i}}\lambda_P \quad (6)$$

$$\frac{\partial H}{\partial y_0} = \frac{\partial M}{\partial y_0}\lambda_M + \frac{\partial P}{\partial y_0}\lambda_P \quad (7)$$

From the maximization definition, the following identities are also known:

$$\sum_{i=-A}^{B} i^2 y_i - c = 0 \quad (8)$$

$$\sum_{i=-A}^{B} y_i - 1 = 0 \quad (9)$$

From the partial derivative equations 5, 6, and 7, where $y_i > 0$

$$-1 - \log y_i = i^2\lambda_M + \lambda_P \quad (10)$$

$$-1 - \log y_{-i} = (-i)^2\lambda_M + \lambda_P \quad (11)$$

$$-1 - \log y_0 = \lambda_P \quad (12)$$

LEMMA 3.1. $y_i = y_{-i}, \forall i \in Z, i \in [-A, B]$

From equation 11,

$$\begin{aligned} -1 - \log y_{-i} &= i^2\lambda_M + \lambda_P \\ &= -1 - \log y_i \\ y_{-i} &= y_i \end{aligned} \quad (13)$$

where $y_i, y_{-i} > 0$. However, $0 \leq y_i \leq 1$. Thus, for the case where $y_i = 0$, suppose $y_{-i} = d$. If $d > 0$, then $y_i = y_{-i}$ by equation 13 above. Therefore, $d = 0$ because $y_{-i} \in [0, 1]$. Therefore, $y_i = y_{-i}, \forall i \in Z, i \in [-A, B]$.

LEMMA 3.2. $A = B$

From lemma 3.1, $y_i = y_{-i}$. Suppose $A < B$. Then $y_j = 0, j \in [-A-1, -B]$. From lemma 3.1, $y_{-j} = y_j = 0$. Therefore, all symbols in the range $[A+1, B]$ have no probability of occurring and can be excluded from the probability distribution y. Therefore, $B = A$. Because A and B are arbitrary, the same logic can be used to show that if $B < A$, then $A = B$. Therefore in all cases, $A = B$. For simplification, we now let $q = A = B$, so $Q = 2q + 1$.

LEMMA 3.3. $\overline{y} = 0$

From lemma 3.1, $y_i = y_{-i}, \forall i \in Z$. Thus,

$$\overline{y} = \sum_{i=-q}^{q} i y_i \quad (14)$$

$$= \sum_{i=-q}^{-1} i y_i + \sum_{i=0}^{0} i y_i + \sum_{i=1}^{q} i y_i \quad (15)$$

$$= \sum_{i=1}^{q} -i y_i + 0 + \sum_{i=1}^{q} i y_i \quad (16)$$

$$= 0 \quad (17)$$

Lemma 3.3 implies that in the optimal case, the average intensities of the original and embedded images is equal: $\overline{Z} = \overline{X}$.

LEMMA 3.4. $y_i = y_{-i} = y_0 2^{-i^2 \lambda_M}, \forall i \in Z$

From equation 12,

$$y_0 = 2^{-(\lambda_P + 1)} \tag{18}$$

where $y_0 > 0$. From equation 10,

$$y_i = 2^{-(i^2 \lambda_M + \lambda_P + 1)} \tag{19}$$

$$y_i = y_0 2^{-i^2 \lambda_M} \tag{20}$$

Thus, from equation 20 and lemma 3.1, in the optimal case, $y_i = y_{-i} = y_0 2^{-i^2 \lambda_M}, \forall i \in Z$.

THEOREM 3.5. *The solution to* $\sum_{i=1}^{q} (2^{-\lambda_M i^2}(c - i^2)) + \frac{c}{2} = 0$ *provides the optimal solution.*

From lemma 3.4, a relationship is defined between the value of y_0 and the remainder of the distribution. Now, the value of λ_M can be calculated. Equations 8 and 9 can now be rewritten as

$$\begin{aligned} \sum_i i^2 y_0 2^{-i^2 \lambda_M} - c &= 0 \\ \sum_i y_0 2^{-i^2 \lambda_M} - 1 &= 0 \end{aligned}$$

Therefore,

$$\begin{aligned} y_0 &= \frac{c}{\sum_{i=-q}^{q} i^2 2^{-i^2 \lambda_M}} \\ y_0 &= \frac{1}{\sum_{i=-q}^{q} 2^{-i^2 \lambda_M}} \\ \sum_{i=-q}^{q} i^2 x^{i^2} &= c \sum_{i=-q}^{q} x^{i^2} \end{aligned} \tag{21}$$

Where $x = 2^{-\lambda_M}, x > 0$. Rewriting equations 21 and 20,

$$y_0 = \frac{1}{\sum_{i=-q}^{q} x^{i^2}} \tag{22}$$

$$y_i = y_0 x^{i^2} \tag{23}$$

$$\begin{aligned} \sum_{i=-q}^{q} x^{i^2} (i^2 - c) &= 0 \\ 2 \sum_{i=1}^{q} (x^{i^2} (i^2 - c)) - c &= 0 \\ \sum_{i=1}^{q} (x^{i^2} (c - i^2)) + \frac{c}{2} &= 0 \end{aligned} \tag{24}$$

Because $\sum$ is involved, the equation cannot be inverted. However, when Q is known, q can be determined and the polynomial can be evaluated numerically, yielding a value for x. As y_i is dependent on x, y can be derived as the optimal solution for the general case.

3.3. Reducing the search region

LEMMA 3.6. *$c \in [0, q^2)$ for a solution to be found.*

From equation 24, the largest co-efficient of the polynomial is q^2. When $c \geq q^2$, $c - i^2 > 0, \forall i \in [1, q]$ and $\frac{c}{2} > 0$, rendering the equation insoluble for $x \geq 0$. Likewise, when $c < 0$, $c - i^2 < 0$ and $\frac{c}{2} < 0$, which bounds $c \in [0, q^2)$.

LEMMA 3.7. $c = 0 \Rightarrow H(y) = 0$

Suppose $c = 0$. For equation 8 to hold true, $y_i = 0, y_0 = 1, i \neq 0, i \in Z$. The entropy of this distribution is $H(y) = -1 \times \log 1 = 0$. Thus $c = 0 \Rightarrow H(y) = 0$.

LEMMA 3.8. *The distribution with the maximum entropy can be found if and only if $x = 1$.*

The maximum entropy for a distribution with Q symbols occurs when the distribution is uniform.[4] That is, all symbols have equal probability: $y_i = y_0 = \frac{1}{Q}$.

From equation 20, $2^{-i^2\lambda_M} = 1$, or $y_i = y_0 = 0$. The second option is not possible because equation 9 would not hold. Therefore, a uniform distribution implies $x = 1$.

Assuming $x = 1$,

$$\begin{aligned} 2^{-i^2\lambda_M} &= 1 \\ -i^2\lambda_M &= 0 \\ \lambda_M &= 0 \end{aligned} \tag{25}$$

i cannot be zero in all cases unless the trivial solution, $Q = 1$, is being considered. From equations 25, 10 and 12,

$$\begin{aligned} i^2\lambda_M + \lambda_P &= \lambda_P \\ -1 - \log y_i &= -1 - \log y_0 \\ y_i &= y_0 \end{aligned}$$

This is the uniform distribution, which has the maximum entropy. Thus, $x = 1$ implies and is the only case where the maximum entropy is reached.

LEMMA 3.9. *The distribution with the maximum entropy can be found if and only if $c = \frac{q(q+1)}{3}$.*

From lemma 3.8, the maximum entropy can be found if and only if $x = 1$. Therefore, equation 24 becomes:

$$\begin{aligned} \sum_{i=1}^{q}(c - i^2) + \frac{c}{2} &= 0 \\ q \times c + \frac{c}{2} &= \frac{q(q+1)(2q+1)}{6} \\ \frac{3c}{2q+1} &= q(q+1)(2q+1) \\ c &= \frac{q(q+1)}{3} \end{aligned}$$

Thus, as $x = 1 \Leftrightarrow c = \frac{q(q+1)}{3}$, from lemma 3.8, the lemma is proved. We label the point $c = \frac{q(q+1)}{3}$ as the *turning point* for a given Q.

LEMMA 3.10. *The curve representing the family of optimal distributions is continuous and can be completely described in the range $x \in (0, 1]$.*

To derive the optimal distribution, a positive real root must exist from equation 24, which is a polynomial of order x^{q^2}. Lemma 3.9 shows that $c = \frac{q(q+1)}{3}$ gives the maximum entropy for a given q. Also, if $c = 0$,

then $x = 0$ is a solution to equation 24, although the optimal distribution cannot be derived from this because $x = 2^{-i^2\lambda_M}$. However, the equation still produces a real root. Thus, for a given value of q, the root is defined for $x = 0$ and $x = 1$.

The Zero Exclusion Condition,[12] is helpful in solving this problem. The parameterized polynomial

$$p(x,q) = \sum_{i=0}^{n} a_i(q)x^i, q \in Q \subset R^l \tag{26}$$

where $a_i(q)$ continuously depends on $q, i = 0, 1, ..., n$. If $deg(p(x,q))$ is invariant, every root $x_i(q)$ of $p(x,q)$ continuously depends on the parameter q. Thus, the pointwise continuous property of $a_i(q)$ implies that $x_i(q)$ is also pointwise continuous. In our case, $a_i(q) = (c - i)$, where $i = k^2$, for some $k \in N$ and $a_i(q) = 0$, otherwise. Thus, as c is continuous in the range $c \in (-\infty, \infty)$, so also all of the roots $x_i(q)$ are continuously joined. Therefore, as $x = 0$ and $x = 1$ are roots for the values of $c = 0$ and $\frac{q(q+1)}{3}$ respectively, then by the sandwich theorem, zeroes exist for all points in the range $[0, 1]$. Because $x = 0$ is a special case, it is excluded from the general statement to avoid difficulty.

Thus, the family of optimal distributions is continuous for $c \in (0, \frac{q(q+1)}{3}]$ and covers every possible bit rate through lemmas 3.7 and 3.9.†

3.4. Clipping

Until this point, the effects of clipping have not been incorporated into the model. Because information needs to be communicated reliably, some pixels are deemed to be unsuitable for embedding information. Any pixels which might be clipped cannot guarantee reliable communication, so some symbols are excluded from the embedding and extraction processes. This provides independence between the decision to use a given pixel and the next symbol to be embedded.

Assuming a uniform distribution for the intensities of pixels within I, the proportion of pixels, K, that will be used for embedding is given by

$$\begin{aligned} K &= \frac{R - (Q-1)}{R} \\ &= 1 - \frac{Q-1}{R} \end{aligned}$$

where R represents the number of discrete values. In general, R is 256. Now, for the $1 - K$ proportion of pixels that are not used, MSE$= 0$. Thus, in order to achieve the intended MSE for the image, the MSE is scaled for the modified pixels: $\hat{c} = \frac{c}{K}$. This provides us with a greater ability to hide information within those pixels but the bit rate must also be scaled, where $\hat{H}(y) = K \times H(y)$. As clipping is avoided, the embedding algorithm is error-free.

3.5. Results

Figure 2 shows the amount of data that can be hidden with respect to the MSE and Q, where embedding is error-free. The performance after the turning point is represented by a dotted line. It is clear that increasing the size of the distribution allows for better performance as the MSE increases. When the MSE is small, shorter alphabets perform better due a larger value of K.

The surface of the optimal distribution is illustrated in figure 3. As the MSE increases, the distribution converges to the uniform distribution from a distribution with one symbol. As illustrated in figure 2, when $c = 0$, no noise is permitted, leaving the only option as $y_0 = 1, y_i = 0, i \neq 0$. For a fixed size alphabet, the maximum entropy is achieved when all symbols are equally likely.[4] The maximum in figure 2 therefore occurs predictably when the distribution is uniform, as shown in figure 3.

† Although it appears true, it has not been proven that the optimal distribution is monotonically increasing in capacity within this range.

Figure 2. The relationship between the number of symbols and performance

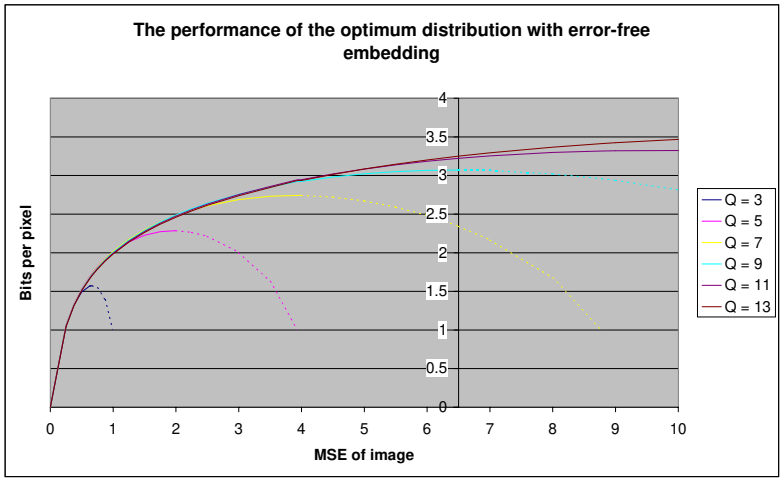


Figure 3. The relationship between the number of symbols and performance in the range identified by lemma 3.10

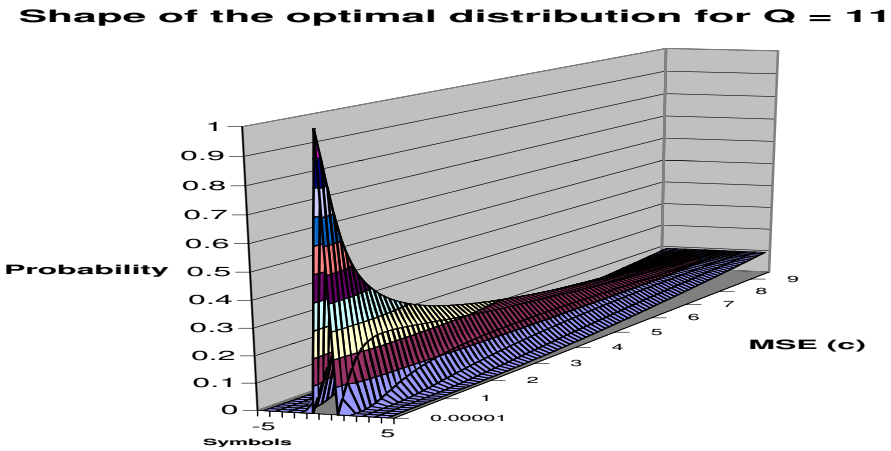


Table 1. Embedding 50 random files in images using the optimal distribution for 40dB

Image	Av. PSNR (dB)	Av. bits/pixel/colour	Boundary prop. (%)
Peppers	39.99	3.252	4.496
Baboon	39.79	3.402	0.115
Airplane	39.79	3.405	0.076
Cameraman	39.79	3.399	0.034
Man	40.26	3.059	10.229
Boat	39.79	3.405	0.018
Average	39.90	3.320	2.495

3.6. Conclusion

The optimal distribution for a PSNR of 40dB can now be determined, using $c = 6.5025$. Across all values of Q and assuming a uniform distribution for the image, the optimum result is when $Q = 13$, where $H(y) = 3.25$bpp.

4. VERIFICATION

4.1. Experiments

We implemented this model in order to verify the limits above. As each trial is purely deterministic, given an input and cover image, we therefore vary both of these in order to give reliability to our claims. The cover images used were "Peppers", "Baboon", and "Airplane" each with 256 by 256 pixels and 3 colours, as well as the grey-scale images of "cameraman", "man", and "boat" which ranged up to 1,024 by 1,024 pixels. These were chosen because they are standard test images in the data hiding field, as well as their differing levels of contrast.

To provide variation in the input, random files of length 100,000 bytes were generated and then compressed (to eliminate any possibility of the random generator being unreliable) before being used as the input data. Our objective was to verify that compressed data is able to be embedded within an image at 3.25bpp (per colour) as derived above. Our initial assumption was only that images have a uniform distribution (and in fact this only is only needed to model the boundary points correctly). Therefore, even though this is quite unlikely to be true in practice, it is expected that it will be insufficient to cause any practical problems.

4.2. Results

If an image has a uniform distribution then the percentage of intensities in this section will be $\frac{Q-1}{256} = 4.69\%$. When it does not have this exact proportion, the results will differ by a small amount. When the PSNR is 40dB, the MSE is 6.5025. After accounting for the expected effects of the boundary section, this is scaled to 6.8223 as shown in subsection 3.4. If no pixels have intensities in this boundary section, then 100% of pixels will be used for data hiding, giving a PSNR of 39.79dB. So in the worst case scenario (in terms of damage) with this distribution, will embed 3.41bpp. There is no practical means of identifying the other extreme, as if the image data was totally white (or black), then no data would be able to be embedded under this model.

The results for the three images are shown in table 1. Although the target PSNR of 40dB is achieved in only one image, the results of the other images are still very close to the expected capacity (3.25bpp) and damage (40dB). Also, with most of the images having a negligible proportion in the boundary section, the PSNR and capacity are approximately equal to the estimates given above.

5. CONCLUSION

We have devised a model that allows the calculation of the optimal distribution in the spatial domain without noise, for a given Mean Squared Error. Thus, the best performance from any steganographic method will be 3.250 bpp per colour, for a PSNR of 40dB. Local variations in the cover image may provide more than this,

but this is the limit for the general case. This performance will diminish if any robustness is required, or if less information can be shared.

There are still some open questions in this area: "Is it possible to use some form of error correction to recover from clipping errors introduced in transmission"; and "How can information be transmitted robustly?". For example, suppose that a symbol did not represent a single value but instead a range, then some sort of robustness might be granted. For example, when embedding three symbols, "-1" could refer to the range $[-4, -2]$, "0" to $[-1, 1]$, and "1" to $[2, 4]$. However, given that tampering usually involves pixels changing by more than a couple of points of intensity, a superior approach would have to be some form of error correction in the symbol domain. As noted in the introduction, some research already examines these questions of robust steganography[9 2 .6]

Though we have confined ourselves to the spatial domain, the use of the frequency domain will not provide any further increase in capacity as the evaluation of the PSNR must be carried out in the spatial domain. The introduction of any message can always be evaluated in the spatial domain, so any technique that uses the frequency domain is still constrained to these limits. We have succeeded in evaluating the maximum reliable rate of transmission of information through the medium of an image from a theoretical perspective, measured the level of steganographic security this provides, as well as illustrating this performance in practice.

ACKNOWLEDGMENTS

This research work is partly funded by the Motorola Australian Research Centre, through the Australian Government SPIRT grant scheme.

REFERENCES

1. G. J. Simmons, "The prisoners' problem and the subliminal channel," in *Advances in Cryptology, Proceedings of CRYPTO '83*, pp. 51–67, Plenum Press, 1984.
2. J. R. Smith and B. O. Comiskey, "Modulation and information hiding in images," in *Workshop on Information Hiding*, **1174**, (Isaac Newton Institute, University of Cambridge, UK), May 1996.
3. L. M. Marvel and J. Charles G. Boncelet, "Capacity of the additive steganographic channel," 1999.
4. C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal* , pp. 379–423, 1948.
5. L. Marvel, C. Boncelet, and J. Retter, "Spread spectrum image steganography," in *IEEE Transactions on Image Processing*, pp. 1075–1083, 1999.
6. M. Barni, F. Bartolini, A. D. Rosa, and A. Piva, "Capacity of the watermark-channel: How many bits can be hidden within a digital image," in *Security and watermarking of multimedia contents, Society of Photo-optical Instrumentation Engineers (SPIE)* **3657**, pp. 437–448, (San Jose, California), January 1999.
7. I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "A secure, robust watermark for multimedia," in *Information hiding: first international workshop, Cambridge, U.K., May 30–June 1, 1996: proceedings, Lecture Notes in Computer Science* **1174**, pp. 185–206, Springer-Verlag, 1996.
8. M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "Copyright protection of digital images by embedded unperceivable marks," in *Image and Vision Computing*, pp. 897–906, 1998.
9. M. Ramkumar and A. Akansu, "Theoretical capacity measures for data hiding in compressed images," in *Multimedia systems and applications*, Nov. 1998.
10. T. C. Bell, J. G. Cleary, and I. H. Witten, *Text Compression*, Prentice Hall, 1990.
11. B. Girod, "What's wrong with mean-squared error?," in *Digital Images and Human Vision*, M.I.T. press, 1993.
12. V. L. Kharitonov, "Asymptotic stability of an equilibrium position of a family of systems of linear differential equations," *Differential'nye Uraveniya* , pp. 1483–1485, 1978.