ARMY RESEARCH LABORATORY

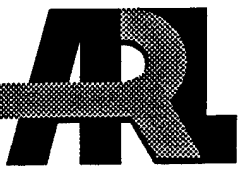


Image Steganography for Hidden Communication

by Lisa M. Marvel

ARL-TR-2200 April 2000

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5067

ARL-TR-2200 **April 2000**

Image Steganography for Hidden Communication

Lisa M. Marvel
Information Science and Technology Directorate, ARL

Abstract

Modern steganographic methods, which conceal the existence of communication, are needed to exploit contemporary modes of information exchange. Measures of performance for these methods are essential to compare specific algorithms and determine appropriate uses. This report develops a methodology for steganographic data hiding. The methodology encompasses derivation of a general theory of steganographic communication, including theoretical capacity bounds, and design of an actual data-hiding technique that used digital imagery as a cover. The technique promotes maximization of payload, allows error-free recovery of embedded date, and provides some resilience to removal while concealing the existence of the embedded information from the observer and the observer's resources (e.g., computer).

Table of Contents

List of Figures

List of Tables

INTENTIONALLY LEFT BLANK.

1. Introduction

As a society, humans have continually sought new and efficient ways to communicate. The earliest methods included cave drawings, smoke signals, and drums. Advancements of civilization introduced written language, telegraph, radio/television, and, most recently, electronic mail. In 1998, the U.S. Postal Service delivered 107 billion first-class items, a modest amount compared to the estimated **4 trillion** e-mail messages received by U.S. residents [1]. And this is only a fraction of what we can expect in the future. As more and more communication is conducted electronically, new needs, issues, and opportunities are born.

At times when we communicate, we prefer that only the intended recipient have the ability to decipher the contents of the communication. We want to keep the message *secret.* A common solution to this problem is the use of encryption to obscure the information content of the message. Today, our credit card numbers are encrypted within e-commerce orders sent over the Internet to prevent fraud. Military battle plans and specific target locations may be encrypted before transmission to preserve the element of surprise during wartime. The documentation of a company's new product design may be encrypted to curtail industrial espionage.

While encryption masks the meaning of a communication, instances exist where we would prefer that the entire communication process not be evident to any observer — that is, even the fact that communication is taking place is a secret. In this case, we want to keep the communication *hidden.* Steganographic techniques can be used to hide or *cover* the existence of communication with other data, intuitively referred to as *cover data.* Consider a sender who wants to convey information to a recipient but does not want anyone else to know that the two parties are communicating. The sender could use steganography to hide information within innocuous information — for example, a weather map that covers the existence of the communication. The weather map would then be made available on an open channel for anyone to access, but only the intended recipient would be aware of the hidden information and have the ability to extract it. Steganography is not meant as a replacement to cryptography, but rather an augmentation — information can be encrypted and then covertly communicated via steganographic means for added privacy. We can think of steganography as one more tool to convey information in a hidden manner.

Privacy is not the only motivation for steganography. By embedding one piece of data inside of another, the two become a single entity, thus eliminating the need to preserve a link between the two different pieces of data to prevent the chance of their separation. One application that exhibits the advantage of this facet of steganography is the embedding of patient information within the medical imagery. By doing so, a permanent association between these two information objects is created.

Steganography can also provide forward and backward compatibility. Consider a system, such as the FM radio receiver, which receives an analog communication signal. If improvements to the radio signal are made, say the addition of an in-band on-channel digital signal to provide better audio quality, new radio receivers would be built to take advantage of this enhancement [2]. However, radio stations could scarcely anticipate that everyone would

immediately purchase this new radio, so they must use steganography to embed the new digital information within the analog signal, leaving traditional receivers unaware that the hidden signal exists.

Additionally, information integrity can be provided using steganography to embed authentication information within the cover data. This is particularly advantageous in an age when the preservation and assurance of digital data is vital.

We can also use steganography to avoid communication restrictions such as those stipulated by firewalls. Many firewalls may not allow encrypted data, if detected, to move freely through the wall. However, the transmission of an innocuous cover that has hidden steganographic information may not experience such limitations.

A multitude of objects can serve as potential cover for steganographic communication. A simple written letter, a commentator's report, or a picture of a child may all seem commonplace. However, tiny pin pricks embedded within the letter, specific wording of the commentator's speech, or the color of the child's hat in the picture can serve to indicate a hidden message. In today's electronic world, the prevalence of multimedia introduces rich, new avenues for hiding communication using digital audio and imagery.

1.1 Problem Motivation

In this report, we address modern steganography. Although the topic of steganography has existed since ancient times, and many of the general assumptions still apply today, new techniques that exploit contemporary technology in addition to metrics to measure their performance should be developed.

Much of the recent work in steganography is in the area of invisible digital *watermarking*, motivated by the desire for copyright protection of multimedia on the Internet. The objective of digital watermarking is to embed a signature within a digital cover signal to signify origin or ownership. Once added, a watermark must be robust to removal attacks and reliably detected, even after typical transformations such as cropping, quantization, and scaling.

Thus far, less attention has been focused on another type of steganography, *data hiding*. The objective of data hiding is to imperceptibly embed a significant amount of data, much more than that of a signature or serial number, into the cover signal. It also differs from watermarking by pursuing the resistance to removal to a much lesser degree.

The potential applications for data hiding are numerous. Of course, the relay of hidden messages is an apparent usage, but today's technology stimulates even more subtle practices. In-band captioning, for example, can be used to embed textual or ancillary information within a cover. It can be employed to deposit creation and revision information within the cover data for the purpose of revision tracking, preventing the need to maintain two separate media. This type of consolidation could be used to join medical images with text, such as patient data, to promote patient safety and record consistency. Additionally, forward and backward compatibility information could be inserted within an audio or video signal to permit additional functionality such as multilingual playback while allowing legacy systems

to continue operation. Data hiding can be utilized as a technique for authentication and tamperproofing. For example, unauthorized alterations in the cover can be detected by hiding attribute information unique to the cover, such as the checksum of certain pixel values, within the cover itself. By computing the checksum at the receiver and comparing it to the extracted checksum, the receiver could determine whether or not the cover has been corrupted.

The objective of this report is to develop a methodology for steganographic data hiding. The research encompasses derivation of a general theory of steganographic communication, including theoretical capacity bounds, and design of an actual data-hiding technique that uses digital imagery as a cover. The technique promotes maximization of payload, allows error-free recovery of the embedded data, and provides some resistance to removal while concealing the existence of the embedded information for observers and their resources.

1.2 Overview of Research

Section 2 begins the presentation of our work by reviewing ancient steganography and surveying relevant literature in the areas. From this review, the need for a general methodology along with performance metrics by which to compare steganographic algorithms becomes evident. In section 3, we address these needs by developing a general steganographic communication theory based on information theory. We also derive theoretical bounds for the capacity of a particular class of steganographic systems, those which add the embedded information to the cover. We subsequently show how this capacity can be applied to image steganography, thus providing a much-needed performance metric. Section 4 introduces a complete system for image steganography, entitled Spread Spectrum Image Steganography, that adheres to our primary goals of high payload, invisibility, good signal recovery, and error resistance. One component of this system is a covert modulation technique that can function in a stand-alone manner. The steganographic capacity bounds developed in section 3 are used as a metric by which to gauge performance of this system. Finally, section 5 summarizes the work and presents additional ideas to stimulate future research.

2. Background

In this section, the reader is familiarized with the science of steganography. We approach this task by first furnishing examples from ancient history and then proceeding to current era. Recent research in modern steganography, with particular focus on data hiding, is then reviewed to provide a foundation for our work.

2.1 Historical Examples

Steganography is not a new science. Some of the first documented examples of steganography can be found in the *Histories* of Herodotus, where the father of history relates several stories from the times of ancient Greece [3]. One is that of Histiaeus, who wished to inform his allies when to revolt against the enemy. To do so, he shaved the head of a trusted servant and then tattooed a message on his scalp. After allowing time for the slave's hair to grow back, he was sent through enemy territory to the allies. To the observer, the slave appeared to be a harmless traveler passing thorough the area. However, upon arrival, the slave reported to the leader of the allies and indicated that his head should be shaved, thereby revealing the message.

In ancient times, one type of writing medium was a wooden tablet covered with wax. A person etched letters in the wax, and when he desired to remove the writing, the wax was melted to a smooth surface and the tablets reused. While exiled in Persia, Demeratus discovered that Greece was about to be invaded and wanted to convey a message of warning. However, the risk of exposure was great for Demeratus, so he concealed his message by writing directly on the wood and then covering it with wax. The seemingly blank tablets were then transported to Sparta where the message was literally uncovered and his allies forewarned.

A less elegant method of hidden communication was adopted by Harpagus, a Median noble. He disguised a messenger as a hunter and hid a message in the body of an unskinned hare. The hunter carried the hare as if it were recently caught. Anyone encountering the messenger/hunter, then would probably comment on his good fortune and be none-the-wiser. The message would then be delivered to the appropriate party without detection or interception.

Recent times have yielded more advanced techniques. The use of invisible inks is one such method, where messages are written using substances that subsequently disappear. The hidden message is revealed using heat or certain chemical reactions. Other methods may employ routine correspondence, such as the application of pin pricks in the vicinity of particular letters to spell out a secret message. Advances in photography produced microfilm that was used to transmit messages via carrier pigeon. Further developments in this area improved film and lenses that provided the ability to reduce the size of secret messages to a printed period. This technique, known as the microdot, was used by the Germans in World War II.

2.2 Modern Steganography

As more of today's communications occur electronically, there have been advancements utilizing digital multimedia signals as vehicles for steganographic communication. These signals, which are typically audio, video, or still imagery, are cover signals. Schemes where the original cover signal is needed to reveal the hidden information are *cover escrow* schemes. They can be useful for traitor-tracing [4]. In this scenario, copies of the cover signal are disseminated with the assignee's identification embedded within. If illegal copies of the signal are acquired, the source of the copy is established by subtracting the original cover data from the modified signal, exposing the offender's identity.

For many applications, it is impractical to require the possession of the unaltered cover signal for extraction of the hidden information. More pragmatic methods operate *blindly*. These blind schemes allow extraction of the embedded data from the modified cover signal without knowledge of the original cover information. Blind strategies are predominant among steganography of the present day.

A block diagram of a generic blind steganographic system (stegosystem), which uses an image as a cover, is depicted in Figure 1. A message is embedded in a digital image by the stegosystem encoder, which uses a key or password. The resulting image, or *stegoimage*, is transmitted over a channel to the receiver, where it is processed by the stegosystem decoder using the same key. During transmission, the stegoimage can be monitored by unintended viewers, who will notice only the transmittal of the innocuous image without discovering the existence of the hidden message.

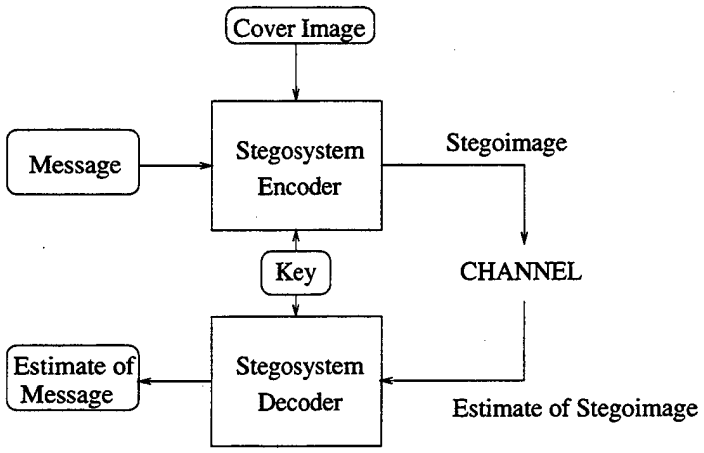


Figure 1. Overview of Blind Steganographic System.

Within the past few years, there has been a surge of research in the area of digital steganography, whose primary objectives are imperceptibility, removal resistance, and payload (the amount of data embedded within the cover). A majority of the work in the area has been in the development of invisible digital watermarking, whose thrust can be attributed to the desire for copyright protection of digital information on the Internet because the data can be reproduced perfectly. The objective of digital watermarking is to embed a signature or serial number within a digital cover to signify ownership. Once added, it must be resistant to

removal and reliably detected even after typical transformations. Another sector in the field of steganography is data hiding, where robustness to removal is secondary to the objective of maximizing the amount of data embedded within the cover. For both watermarking and data hiding, the embedded data must remain imperceptible to observers.

2.3 Information Theory Analogies

The act of digital steganography, also referred to as information hiding, can be characterized utilizing the theories of communication [5]. The parameters of hidden communication can be related to the characteristics of communication systems. For instance, the maximum amount of hidden data that can be hidden and successfully extracted is commensurate with the *capacity* of a communication channel. The imperceptibility, or undetectability, of hidden data is associated with a *signal-to-noise ratio* (SNR). In this context, the embedded message represents the information-bearing signal and the cover data is viewed as noise. Contrary to typical communication scenarios where a high SNR is desired, a low SNR is preferred for steganography and other systems that desire low probability of detection and low probability of interception. Low SNR corresponding to lower perceptibility that signifies greater concealment of the embedded signal. Additionally, the resistance to removal of the embedded information is analogous to the *jamming margin*. The measure of jamming resistance is used to describe a level of resistance to removal or destruction of the embedded signal, accidental or intentional.

It is not possible to simultaneously maximize removal resistance and payload capacity while adhering to the imperceptibility constraints (low SNR) imposed by steganography. Therefore, the acceptable balance of these items is dictated by the application. For example, data-hiding schemes forego removal resistance in favor of capacity and invisibility, whereas an invisible watermarking scheme, which does not require large capacity, would certainly advocate increased removal resistance. Finally, a steganographic scheme used as a method of covert communication would adopt the utmost undetectability while sacrificing resistance to removal and possibly capacity.

2.4 Survey of Relevant Literature

Digital steganography is currently an active research area, encompassing methods of copyright protection, image authentication, and hidden communications. Since our research pertains to data hiding, where emphasis is placed upon invisibility and the maximization of payload, we limit our discussion to steganographic methods with these commonalities.

2.4.1 Existing Methods

A simple method of data hiding involves the manipulation of the least significant bit (LSB) plane of the data. Various techniques, such as direct replacement of the cover LSBs with message bits or an arithmetic combination between the two, are used. Several examples

of LSB schemes can be found in Schyndel et al., Wolfgang et al., and Machado [6, 7, 8]. LSB manipulation software has been written for a variety of image formats and can be found in Mildbrandt [9]. These methods typically achieve both high payload and low perceptibility. However, because the fact that the data are hidden in the LSB may be known, LSB methods are vulnerable to extraction by unauthorized parties.

There are, of course, many approaches that function as cover escrow schemes, where it is necessary to possess the original cover signal in order to retrieve the hidden information. Examples of such schemes can be found in Cox et al., Podilchuk et al., and Swanson et al. [10, 11, 12]. Due to the fact that they place impractical escrow requirements on part of the recipient, we only consider the schemes of particular interest to the extent that they can be made to function blindly.

Several procedures for data hiding in multimedia can be found in Bender et al. [13]. One of these, entitled Patchwork, alters the statistics of the cover image. First, pairs of image regions are selected using a pseudorandom number generator. Once a pair is selected, the values of the pixel are altered so that the relationship between the regions reflect the hidden data. For instance, if all pixels of the first selected region are greater than those of the second, the hidden data bit is equal to 1. Although the modification is typically small and not perceptible, it is not restricted to the LSB. This scheme is somewhat robust to removal but has low payload.

Smith and Comiskey present several spread spectrum data-hiding methods [5]. These techniques utilize the binary message data, $b_i \in \{-1, 1\}$, to modulate a carrier signal, ϕ_i,

$$S(x, y) = \sum_i b_i \phi_i(x, y). \tag{1}$$

In the ideal case, the carried signal is a basis function that is orthogonal to the cover image N. In reality, the two may not be completely orthogonal,

$$\sum_{x,y} \phi_i(x, y) N(x, y) \approx 0. \tag{2}$$

The embedded signal, S, is then added to the cover image, N, to construct the stegoimage, D,

$$D(x, y) = S(x, y) + N(x, y). \tag{3}$$

The message is extracted via cross correlation between the stegoimage and the carrier regenerated by a local reference; hence, cover image escrow is unnecessary.

$$o_i = \sum_{x,y} D(x, y) \phi_i(x, y). \tag{4}$$

A thresholding operation is performed on o_i to determine the binary value of the embedded data bits. They are able to hide and recover 100 bits of information in a 320-pixel $\times$ 320-pixel grayscale image, an information rate (information bits/cover bits) of 0.0001.

A data-hiding scheme using the statistical properties of dithered imagery, images that represent each pixel value with a pattern of dots, is proposed by Tanaka, Nakamura, and

Matsui [14]. With this method, the dot patterns of the ordered dither pixels are controlled by the information bits to be concealed. This system accommodates 2 kilobytes of hidden information for a binary 256×256 image, yielding a payload of 1 information bit to 4 cover image bits (information rate of 0.25). An information rate of 0.1666 was obtained for trilevel images of the same size. The method has high payload but is restricted to dithered images and is not resistant to errors in the stegoimage.

Davern and Scott present an approach to image steganography utilizing fractal image compression operations [15]. The fractal image compression process splits the image into blocks, and those that are visually similar are identified. One bit of steganographic data is hidden by transforming one similar block into an approximation for another and then storing the approximation in the position of the original block in the stegoimage. The data are decoded using a visual key that specifies the position of the regions containing the hidden data. Unfortunately, the amount of data that can be hidden using this method is small (1 bit for each block) and easily corrupted. Additionally, the search for similar blocks in the encoder and the comparison process in the decoder are both computationally expensive operations.

Recent research by Swanson, Zhu, and Tewfik [16] utilizes an approach of perceptual masking to exploit characteristics of the human visual system (HVS) for data hiding. Perceptual masking refers to any situation where information in certain regions of an image is occluded by perceptually more prominent information in another part of the scene [17]. This masking is performed in either the spatial or frequency domain using techniques similar to Cox et al. [10] and Smith and Comisky [5] without cover image escrow. The payload potential of the system is not quantified.

A pertinent data-hiding technique using audio is presented in Neubaur, Herre, and Brandenburg [18]. This method provides a channel to convey hidden data within an uncompressed audio stream using direct-sequence spread spectrum binary phase shift keying modulation. An antipodal pseudorandom noise sequence (+1,–1 in this case) is generated and multiplied by the antipodal message bits to construct the embedded signal. The embedded signal is then up-converted to achieve a spectral shift such that its maximum contribution is fixed. Masking techniques similar to those proposed in Swanson, Zhu, and Tewfik [16] are then applied to derive weights to the embedded signal, which is then added to the audio samples. The decoder consists of a matched filter, synchronizer, and threshold decision unit. The matched filter is designed to match the pseudorandom noise sequence used for encoding. The output of the matched filter is thresholded to determine the value of the message bit. The average resulting payload for this system is 0.0007 information bits to audio bits, with an average message bit error rate of 0.05.

Westfeld and Wolf present a steganographic method for video conferencing [19]. They describe a system that embeds data into video that has undergone lossy compression via an H.261 video conferencing system [20]. The message bit is hidden by altering the phase of quantized DCT coefficients. Naturally, the payload of this system is dependent upon the characteristics of the video, and the quality of the video with the embedded data is difficult to estimate.

A method of hiding speech in video is presented in Mukherjee, Chae, and Mitra [21]. This data-hiding technique embeds speech data, which has been compressed using vector quantization, into a digital video signal. Each video frame is transformed by an orthogonal wavelet transform [22]. The vector-quantized speech indices are then embedded into the wavelet coefficients. To make this scheme blind, the original video is altered by zeroing out the coefficients in one or more of the high-high bands and inserting the hidden data in this location of the transformed video fame. This alteration promotes extraction of the hidden data without knowledge of the original because the location of the hidden data is known. The video data is then inversely transformed before distribution. In the example for Mukherjee et al., they embed a quantized speech signal within a video stream and then attempt to corrupt the hidden data by compressing the stream using H.263 video coding [23]. The speech recovered after compression was deemed intelligible by the authors.

2.4.2 Existing Metrics

In data hiding, we have two primary objectives: the embedded data must be imperceptible to the observer, including the observer's resources such as computer analysis, and it should have the maximum payload possible.

It is difficult to quantify how imperceptible embedded data is. In the case of image steganography, the typical observer's detection resources include the HVS and, potentially, computer analysis. For most of the methods presented in the previous subsection using imagery, the imperceptibility of the embedded data is indicated by illustrating the original image and its counterpart with embedded data so that their visual differences, if any, can be determined. Additionally, the mean-squared-error (MSE) (5) or peak-signal-to-noise ratio (PSNR) (6) between the original and the stegoimage may be presented. The original image's pixels are represented as x_i and the stegoimage pixels as $\hat{x}_i$. The variable L reflects the peak signal level (L=255 for grayscale images).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2. \tag{5}$$

$$PSNR = 10\log_{10}\frac{L^2}{MSE}. \tag{6}$$

In audio steganography, the observer uses the human auditory system as well as computer analysis as detection devices. As a measure of comparison, the spectrograms of the original and modified signals are typically presented. A spectrogram is simply a plot of the frequency content of an audio signal as a function of time. In some cases, the quality of the audio with the embedded data was measured using a perceptual audio quality measure, noise-to-mask ratio, or the SNR,

$$SNR = \frac{\sigma^2_{signal}}{\sigma^2_{noise}}. \tag{7}$$

Here, σ^2 represents the sample variance of perspective signals. In (7), the embedded data is represented as the signal and the cover data as the noise. The SNR, which can also be

applied to images, provides a rough estimate of perceptibility because with all things being equal, the higher the SNR, the more perceptible the signal [5].

As far as a measurement of maximum payload, or capacity, currently published works use the capacity of an additive Gaussian channel [5, 24, 25, 21], shown in (8), derived by Shannon in [26]. Here, the power of the embedded signal is presented as S and the power of the cover data as N. The formulation assumes that the statistics of the cover data obey a Gaussian distribution, which is typically not the case. Many agree, the author included, that better models of channel noise will yield better capacity estimates [5].

$$C = \frac{1}{2}\log\left(1 + \frac{S}{N}\right). \tag{8}$$

2.5 Summary

In this section, we have introduced steganography, modern and historical, and provided an analogy with communication systems using information theory. The differing, and somewhat opposing, objectives of steganography have been presented. Of the two areas of steganography presented, watermarking strives to embed a small amount of data that is difficult to remove, while data hiding embeds a large amount of data that is less resistant to removal. Imperceptibility is a common goal in both.

Since our work is based in the area of data hiding, we have presented recent literature on existing data-hiding methods for imagery, audio, and video. Additionally, we have presented the metrics currently used to gauge the performance of these systems. The capacity of a data-hiding system has been traditionally compared to that of the additive Gaussian noise channel. However, the cover data are not well modeled by the Gaussian distribution because of the dependency among adjacent samples, and a more accurate metric must be derived.

In the text that follows, our research addresses capacity by developing a capacity metric that can be used as a performance criterion for the class of steganographic methods that add the embedded signal to the cover data. This measure, which can be used for all types of cover data including imagery and audio, is a more precise estimate of the maximum amount of payload that can be embedded and successfully extracted from the cover data.

Additionally, we develop a novel data-hiding technique that uses digital imagery as a cover signal, Spread Spectrum Image Steganography (SSIS). SSIS provides the ability to embed a significant amount of information within digital images while avoiding detection by an observer. The system is resistant to errors such as those caused by a noisy transmission channel or lossy image compression. The hidden data is recovered, with high probability, error free. Finally, the system operates blindly; the proposed recipient need only possess a key to reveal the hidden message; otherwise, the existence of the hidden information is undetectable by human or today's computer analysis.

3. Capacity of the Additive Steganographic Channel

We introduce the characteristics of the additive steganographic channel and derive capacity bounds by employing similarities to an arbitrary noise channel. These bounds are compared with the capacity of the additive white Gaussian noise (AWGN) channel. Although we concentrate on steganographic techniques that use digital imagery as a cover, the basic concepts presented here can also be extended to any in the class of steganographic systems in which the embedded signal is *added* to the cover data.

A direct link between communication theory and steganography can be established if we think of the data to be hidden as the information-bearing signal and the cover as the channel *through* which our information is conveyed. For image steganography, where the hidden data is embedded within an image, we can conceive the image as the *channel* by which our information is transmitted. Continuing this analogy, the capacity of the steganographic channel is defined as the maximum rate at which information can be reliably transmitted through this channel and is expressed as the maximum rate that can be successfully hidden and recovered [5]. Like many communication channels, the steganographic channel may cause the information-bearing signal to become distorted. Additionally, the ability to detect the message, in both a communication and steganographic manner, is related to the SNR (7), where the signal is the embedded signal and the noise is the cover. From a communication standpoint, to reliably decode the steganographic signal, the SNR should be high. This is at odds with the steganographic requirement that the SNR be low, indicating concealment. Therefore, a balance must be established where the embedded signal can be reliably decoded and yet remain invisible. Be that as it may, it is noteworthy to mention that the SNR is only a first-order approximation of concealment in terms of detection by the HVS, and more accurate perceptual models are under investigation.

3.1 The Channel Model

We introduce the channel model and give a brief overview of capacity, the maximum rate that information can be transmitted by a channel incurring an arbitrarily small number of errors using any encoding system [27]. Consider the channel model shown in Figure 2, where the input into the channel is represented by the variable X and the output is represented by variable Y. The capacity of this channel, C, is expressed as the mutual information, I, between X and Y maximized by all possible distributions of the input variable X (9).

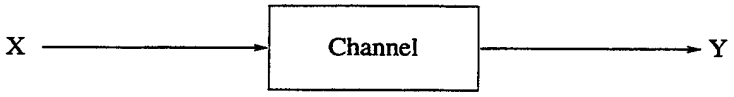


Figure 2. Communication Channel.

$$C = \max_{p(x)} I(X;Y). \tag{9}$$

Mutual information is expressed as the entropy, or information content, of X minus the information in X given by Y and is a symmetric entity as shown in (10).

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X). \tag{10}$$

$H(X)$ is the entropy of the random variable X expressed as

$$H(X) = -\sum_{p(x)} p(x) \log p(x), \tag{11}$$

and $H(X|Y)$ is the conditional entropy of X given Y,

$$H(X|Y) = -\sum_{k} P_{y=k} \sum_{j} P_{x=j|y=k} \log P_{x=j|y=k}. \tag{12}$$

For the noiseless channel, the channel output is equal to the channel input, $Y = X$; therefore, $H(X|Y) = 0$ in (10). So, the channel capacity is equal to the maximum information in X over all possible distributions of X, as in (13).

$$C_{noiseless} = \max_{p(x)} H(X). \tag{13}$$

Now let us look at the capacity of a channel that adds noise, represented as the variable Z, to the input signal such that $Y = X + Z$, shown in Figure 3. Using (10) and the assumption

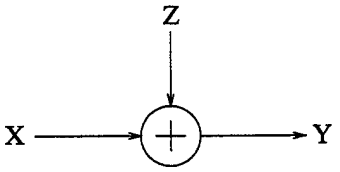


Figure 3. Additive Noise Channel.

that the input signal X and the noise Z are independent, the conditional entropy of Y given X is equal to the entropy of the noise Z. We use this result to determine the capacity of the additive noise channel.

$$C = \max_{p(x)} H(Y) - H(Z). \tag{14}$$

For the AWGN channel, X and Z are independent and Z has a Gaussian distribution with zero mean and variance N. The entropy of Z is then expressed as (15) [28].

$$H(Z) = \frac{1}{2} \log(2\pi e N). \tag{15}$$

To obtain the maximum mutual information over all possible distributions of the input, we assume that the channel input X also has a Gaussian distribution with some variance S.

Therefore, Y is the sum of the two Gaussian processes and has variance $S + N$. Then the AWGN channel capacity, C_g, conveying binary data is expressed as

$$\begin{aligned} C_g &= \frac{1}{2}\log_2[2\pi e(S+N)] - \frac{1}{2}\log_2(2\pi eN) \\ &= \frac{1}{2}\log_2\frac{S+N}{N}. \end{aligned} \tag{16}$$

It has been shown that the mutual information between the channel input and output for the additive channel is at a minimum when the channel noise is Gaussian with zero mean [26]. Consequently, the capacity of other additive non-Gaussian noise channels is lower-bounded by C_g (17). Equations 17 – 19 list the capacity of three such channels with various noise distributions [29].

$$C_g \leq C_{\text{Uniform}} \leq C_g + 0.2546, \tag{17}$$

$$C_g \leq C_{\text{Laplacian}} \leq C_g + 0.1044, \tag{18}$$

$$C_g \leq C_{\text{Trianular}} \leq C_g + 0.0333. \tag{19}$$

3.2 Steganographic Channel Capacity

Now let us consider steganographic systems in which the hidden information is embedded in some manner within a random noise signal, then *added* to the cover data on a sample-by-sample basis as in the methods described previously [5, 16, 30, 31, 32]. The preliminary concepts of Marvel, Boncelet, and Retter [32] were presented elsewhere [33, 34] and are fully described in section 4 of this report. Using the channel model presented in the previous subsection, the steganographic channel can be modeled as shown in Figure 4.

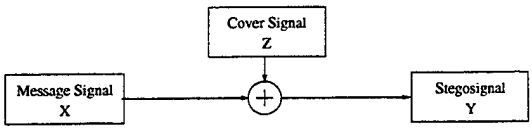


Figure 4. Additive Steganographic Channel.

To date, the image steganography capacity measure used for performance comparisons of these systems is that of a channel with AWGN, C_g (17) applied to steganography in Smith and Comisky [5]. However, a natural digital image, which consists of continuous tones, is not well modeled as a Gaussian noise process because of the underlying structure reflected in the high dependency among neighboring pixels. As with other non-Gaussian channels, (17–19), the capacity of the image steganographic channel is also lower bounded by the capacity of the Gaussian channel. Therefore, we show here and elsewhere [35, 36] that the channel capacity for an *arbitrary* noise channel [26] is better suited as the capacity of the image steganographic channel.

Shannon's theory states that the "randomness," or uncertainty, of a noise with an arbitrary distribution can be compared with that of white Gaussian noise using a measure of *entropy power*, also known as effective noise power. If an arbitrary noise Z has entropy $H(Z)$, the entropy power, average noise power of WGN having the same entropy, of Z is defined as

$$N_e(Z) = \frac{1}{2\pi e} e^{2H(Z)}. \tag{20}$$

Using (20) in conjunction with channel capacity for a channel with additive noise, the steganographic capacity is bounded by

$$C_g \leq C_{stego} \leq \frac{1}{2} log_2 \frac{S+N}{N_e}, \tag{21}$$

where N_e is the entropy power of the noise of the cover data. Because N_e is strictly less than N for all non-Gaussian channels, C_g (17) functions as a lower bound in (21). The upper bound is obtained by maximum mutual information between X and Y by assuming that Y has a Gaussian distribution with variance $S+N$ and channel noise Z is Gaussian with power N_e.

A tighter lower bound than C_g can be obtain using the entropy-power inequality [28],

$$C_g \leq \frac{1}{2} log_2 \frac{S+N_e}{N_e} \leq C_{stego} \leq \frac{1}{2} log_2 \frac{S+N}{N_e}. \tag{22}$$

In essence, the capacity of the arbitrary additive noise channel is lower bounded by the capacity of a WGN channel with the same entropy as the arbitrary noise and upper bounded by the channel with maximum entropy of the output (the output of the channel is Gaussian) under the same noise conditions.

Obviously, if Z is WGN, the entropy power (20) reduces to N and the bounds of (22) reduce to C_g as expected.

3.3 Capacity for Image Steganography

We use the bounds of C_{stego} (22) to determine the capacity bounds for the additive *image* steganographic channel. From the basic premise of the arbitrary noise channel, we can say that the cover image represents the channel noise, which perturbs the information-bearing signal much like AWGN with power N_e. The parameters of image steganography channel capacity are S, the power of the embedded signal consisting of random noise concealing the hidden information; N_e, the entropy power of the cover image; and N, the average power of the cover image. The entropy power of the cover is expressed as

$$N_e(image) = \frac{1}{2\pi e} e^{2H(image)}. \tag{23}$$

Because a natural image has continuous tones and is highly correlated in two dimensions, it can be considered a source with memory. The entropy of such a source is described as

$$H_n = -\frac{1}{n}\int \cdots \int p(x_1, \cdots, x_n) \log p(x_1, \cdots, x_n) dx_1, \cdots, dx_n, \tag{24}$$

with

$$H = \lim_{n \to \infty} H_n. \tag{25}$$

To compute this entropy, the probability distribution of the image must be defined. Many image models exist [37, 38, 39, 40], although there is not one that is consistently accepted. Consequently, without a good model of the image, computation of the probability distribution and thus the true entropy (25) is not possible. However, we can use the well-known result from channel coding that the average codeword length per source symbol is greater than or equal to the entropy of that source (26) [41], to estimate the upper bound on the true entropy.

$$\bar{l}_{image} \geq H(image). \tag{26}$$

Consequently, the average bitrate in bits per pixel (bpp) produced by the state-of-the-art lossless image compression algorithm CALIC [42] is an upper bound on the cover image entropy. The CALIC algorithm provides an average lossless bitrate of 2.99 bpp for the 18 8-bit (grayscale) test images selected by the International Standards Organization (ISO). Using the CALIC bitrate as an estimate of the image entropy, $H(image)$, to obtain a value for the entropy power in (23), we can calculate both upper and lower capacity bounds for the image steganographic channel.

As a generic result, the average power among several test images and the average CALIC bitrate were used to calculate the steganographic capacity bounds for a range of SNR values shown in Figure 5. The capacity for the Gaussian channel for the also appears in this graph. The average CALIC bitrate is 4.9588, and the average image power is 2284.7. The solid line in the figure represents the upper bound on the steganographic capacity with the dashed line indicating the lower bound from (22). The dotted line portrays the Gaussian capacity bound (17). Notice that as the SNR increases, so does the capacity. Conversely, as the SNR decreases, the lower capacity bounds go to 0. Furthermore, note the disparity between the Gaussian bound and the upper and lower image steganographic capacity bounds. The true steganographic capacity bound lies somewhere between the upper and lower bound and more accurately reflects the maximum amount of information that can be embedded within the cover.

The bounds were also calculated for two continuous-tone grayscale images that are used for demonstration later in this report. Results are plotted against the steganographic SNR and are shown in Figures 6 and 7. Figure 6 shows the upper and lower bounds for steganographic channel capacity for a subsampled version of the popular Lena image. The CALIC bitrate of 4.6321 bpp was used as an estimate of the cover image entropy. Figure 7 displays theoretical capacity curves and performance of our steganographic technique, SSIS, for the Eiger image with a CALIC bitrate of 5.2366 bpp.

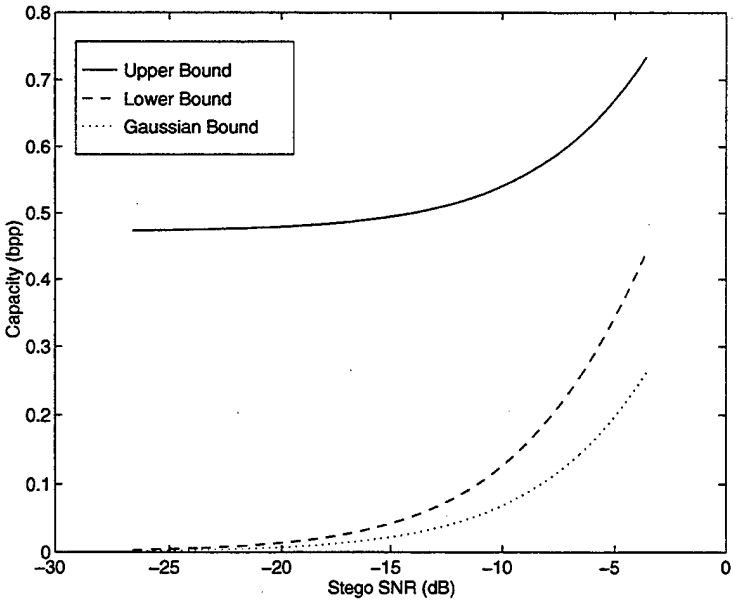


Figure 5. Generic Steganographic Channel Capacity.

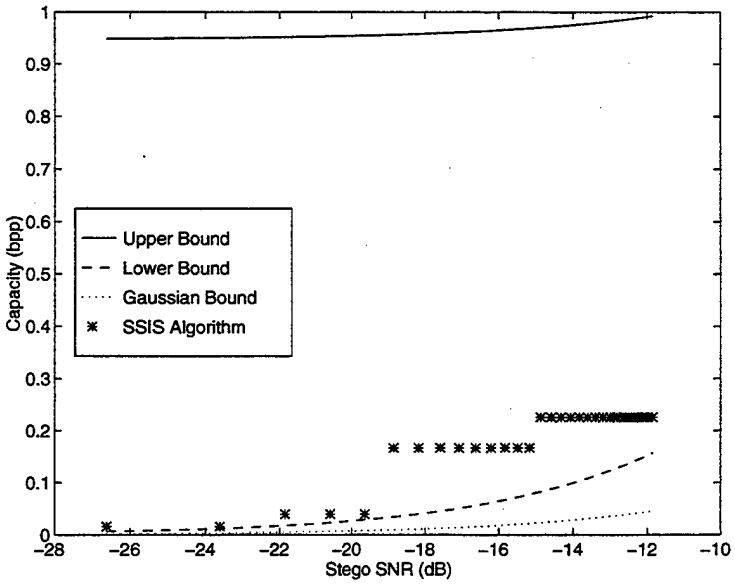


Figure 6. Steganography Capacity for the Lena Image.

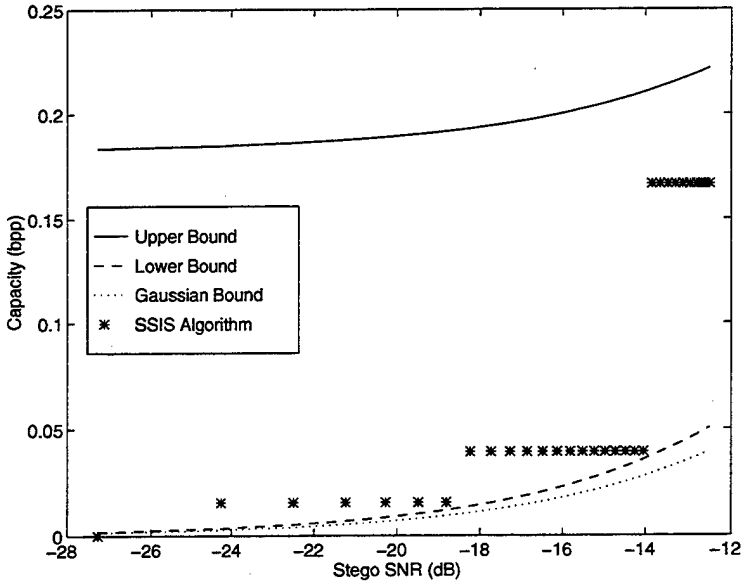


Figure 7. Steganography Capacity for the Eiger Image.

Again, the solid line represents the upper bound and the dashed line indicates the lower bound on the steganographic capacity. The dotted line portrays the Gaussian capacity bound. The asterisks in the graphs depict the performance of the image steganography method, SSIS, we have presented elsewhere [32]. Note that the performance of the SSIS system falls within the theoretic upper and lower capacity bounds. Since the maximum rate must be less than or equal to the capacity, the SSIS performance establishes new lower capacity bound for that particular steganographic channel (image). The SSIS technique, which incorporated spread spectrum, channel estimation, and error-control coding, is fully described in section 4.

As previously mentioned, SNR can be used as a first-order approximation of the degree of concealment. Although the permissible power of the information signal, S, may vary with the local characteristics of the cover image, an overall steganographic SNR of less than –10 to –15 dB was found to be acceptable for these particular images. Therefore, capacity values for less than –10 to –15 dB SNR are most relevant.

3.4 Summary

In this section, we have reviewed the capacity bounds for noiseless, AWGN, and arbitrarily noisy channels and have used these bounds to model and derive the bounds for steganographic channel capacity. These new bounds have been approximated by estimating the entropy of the cover using the bitrate of the current state-of-the-art lossless image compression algorithm, and a generic graph has been constructed to show the relation of the bounds. Finally, the bounds have been calculated for two images, then compared with the

capacity of the AWGN channel and the actual performance of an image steganography system. It has been shown that the AWGN channel capacity, which has been the performance metric used for steganography, greatly understates the steganographic channel capacity and that the new bound more accurately reflects the capacity for steganographic systems.

4. Spread Spectrum Image Steganography (SSIS)

Our method of SSIS is a steganographic communication method that uses digital imagery as a cover signal. It is not to be considered a watermarking method, but rather a data-hiding method that provides the ability to embed a significant amount of information within digital images while avoiding detection by an observer. Hence, more emphasis is placed on the maximization of payload and invisibility with less focus on resistance to removal. Furthermore, SSIS is a blind scheme where the original image is not needed to extract the hidden information. The recipient need only possess a key to reveal the hidden message; otherwise, even the existence of the hidden information is undetectable by current means.

Techniques of error-control coding, channel estimation, and spread spectrum communication are combined within the SSIS system. The fundamental concept is the embedding of the hidden information within samples of a noise-like waveform that is then added to a digital cover image. This waveform is typical of the noise inherent in the image acquisition process and, if kept at low levels, is not perceptible to the human eye or by computer analysis (without access to the original image). To successfully decode the message, channel estimation techniques and error-control coding are employed. In SSIS, channel estimation consists of image restoration techniques (since the channel in this case is an image) that approximate the original cover image from the stegoimage, thus allowing the receiver to function blindly. This approximation is then used to acquire an estimate of the embedded signal that has been added to the cover. Finally, because the added noise is of low power and the restoration and signal detection processes are not perfect, the estimation of the embedded signal may have errors that will result in a message bit error rate (BER) that is rather high. To correct these errors, an error-control code (ECC) is applied to the message signal before embedding.

The major processes of the stegosystem encoder are portrayed in Figure 8. Within the system, the message is optionally encrypted with key 1 and then encoded via a low-rate ECC, producing the encoded message, m. The sender enters key 2 into a pseudorandom noise generator, producing n, a real-valued sequence whose samples have a Gaussian distribution. The modulation scheme combines the message with the noise sequence, thereby composing the embedded signal, s, that is then input to an interleaver, which uses key 3. This resulting signal is then added to the cover image, f, to produce the stegoimage, h, which is appropriately quantized and clipped to preserve the typical dynamic range of the cover image (0–255 for grayscale images). The stegoimage is then transmitted in some manner to the recipient.

The stegoimage is passed through the transmission channel and received by the recipient, who, maintaining the same keys as the sender, uses the stegosystem decoder, shown in Figure 9, to extract the hidden information. The decoder first tries to estimate the channel (cover image) using image restoration techniques to produce an estimate of the original cover image, $\hat{f}$, from the received stegoimage, $\hat{h}$. The difference between $\hat{h}$ and $\hat{f}$ is fed into a deinterleaver to construct an estimate of the embedded signal, $\hat{s}$. With key 2, the noise-like sequence, n, is regenerated, and the embedded signal is then demodulated, thereby constructing an estimate of the encoded message, $\hat{m}$. The ECC decodes the encoded message which is decrypted (if encrypted) using key 1, and revealed to the recipient.

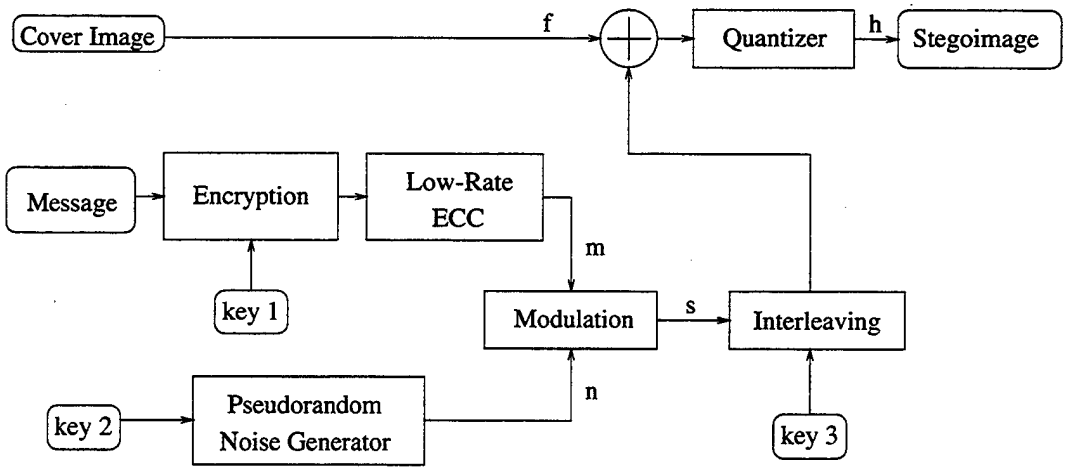


Figure 8. SSIS Encoder.

The interleaver in this scheme, which reorders the embedded signal before it is added to the cover image, serves a dual function. The first is to distribute a group or *burst* of errors uniformly among many codewords, thus allowing errors to occur almost independently within a codeword and increasing the probability that the number of errors occurring in any one codeword will not exceed the error-correcting capability of the code [43]. Secondly, since the interleaver requires a key to stipulate the interleaving algorithm, this key can serve as another level of ambiguity in order to establish the proper order of the embedded signal before decoding.

SSIS uses noise inherent to digital imagery to hide information within the image. Wideband thermal noise is common in imagery of natural scenes captured by photoelectronic systems, such as CCD arrays, and can be modeled as AWGN [44]. In SSIS, a variation of spread spectrum techniques is used to embed the message within AWGN, and, because the message may be encoded using a low-rate error-control code, the encoding has a similar spreading effect as few message bits are spread among the many encoder output symbols. This additional noise that conceals the hidden message is a natural phenomenon of digital imagery and, therefore, if kept at typical levels, may not be noticed by the casual observer or detected by computer analysis.

The major components of the encoder and decoder are described in the following subsections. Section 4.1 delineates the two modulation techniques used within SSIS. Section 4.2 provides insight into the channel estimation stage. Section 4.3 details many types of error-control coding used within our steganography system. Section 4.4 displays system performance when operating in the noiseless transmission channel, additive noise transmission channel, and noise caused by lossy image compression. Finally, the system performance is compared to the capacity bounds developed in the previous section.

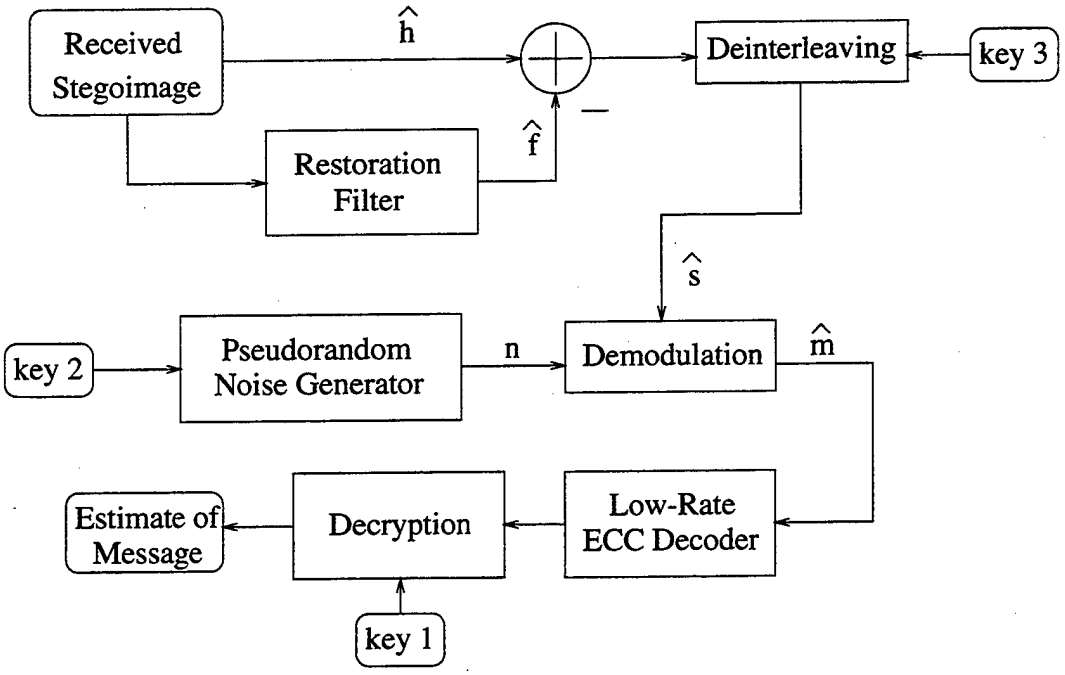


Figure 9. SSIS Decoder.

4.1 Modulation

It is the central philosophy of SSIS that the embedded signal added to the cover image must have the same characteristics as noise inherent to the image — namely, low-power AWGN — so the hidden information will remain invisible. Consequently, for cover images that consist of natural scenes, the embedded signal is constructed by modulating the message data with a white Gaussian noise-like waveform in such a way that the modulated signal maintains the Gaussian distribution.

The concept of a stored reference spread spectrum communication system [45] is used to enable independent generation of identical pseudorandom waveforms at both the transmitter and receiver. Therefore, both the sender and receiver must possess the same key [46] and identical waveform generators.

For our system, we use discrete real-valued samples of the noise-like waveform. The waveform is generated by a pseudorandom number generator and has a Gaussian distribution. The number generator meets the requirements of a random sequence as specified in Knuth [47].

Two types of modulation are used within SSIS: a simple sign (antipodal) modulation technique and a piecewise linear technique for improved detection performance. In this particular subsection, we refer to *detection* in the communication sense, correctly interpreting the transmitted signal, and not in the steganographic sense, as a measure of detecting the existence of the embedded signal by an observer.

4.1.1 Sign-Detector System

Let us begin by describing the sign modulation technique, the initial modulation method used within our system that is similar to the one used in Hartung and Girod [30]. Assume that the message signal, m, is a bilevel signal consisting of $\{-1, +1\}$ and the noise sequence, n, is a sequence of real numbers that has a Gaussian distribution with zero mean and sample variance, σ^2. The two signals are modulated by simple multiplication, as in (27),

$$s(m, n) = m * n. \tag{27}$$

The embedded signal, s, is a sequence of real numbers possessing a Gaussian distribution with zero mean and sample variance, σ^2.

The sign of both the embedded signal sample and the original noise sequence sample determines the value of the message bit. Since the noise sequence is symmetric about zero, a change in sign preserves the Gaussian distribution of the signal. The demodulation process is elementary. The sequence n is replicated at the receiver, and the sign of this sequence is compared to the sign of the received embedded sequence, $\hat{s}$, to recover an estimated value of the message sequence, $\hat{m}$, as shown in (28).

$$sign\left(\frac{\hat{s}}{n}\right) = \hat{m}. \tag{28}$$

Even though this modulation method meets the necessary requirements of producing a Gaussian sequence regardless of the distribution of the message sequence, a major deficiency lies within the detection of this signal in the presence of noise for the transmission process. Because the embedded signal must follow a Gaussian distribution, many of the sample values occur in the vicinity of zero, with fewer samples at the tails. Moreover, only the variation of the sign of samples indicates the value of the encoded message bits. Although the distance, D (29), between the values of the embedded signal for both values of m is large for extreme values of the Gaussian waveform, it is much more often small, in accordance with the Gaussian distribution.

$$D = |s(n, m = -1) - s(n, m = +1)|. \tag{29}$$

In most instances, when the embedded signal is exposed to external noise from the effects of the cover image or the actual transmission channel, correct detection of the encoded message sequence is unlikely. As with many communication signals that may be exposed to noise, we desire to have the points within our signal constellation as far apart as possible, reflecting a large minimum distance.

Consider a simple bipolar modulation scheme, shown in Figure 10. The modulation points a_i and b_i are separated by d_{min}. This minimum distance is defined as the smallest Euclidean distance between any pair of distinct points in the signal constellation, as in

$$d_{min} = \min_{all\ i} |a_i - b_i|. \tag{30}$$

Communication constellations are typically compared by their d_{min}. If the modulated signal incurs noise or distortion, the larger the minimum distance, the more distortion the modulated signal can incur and still be demodulated correctly with high probability. However, if the distortion is greater than the threshold value $\lfloor d_{min}/2 \rfloor$, then a demodulation error will occur.

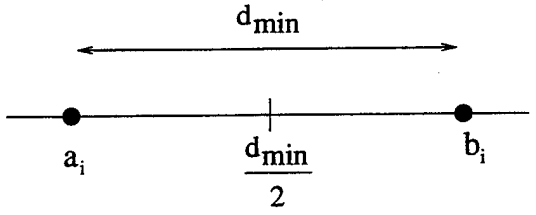


Figure 10. Bipolar Modulation.

For steganography, the same concepts apply. With steganography, the maximum allowable power of the embedded signal is dictated by the need for concealment. The minimum distance for the steganographic system is

$$d_{min} = \min_{all\ i} |s(n_i, m_i = -1) - s(n_i, m_i = +1)| . \tag{31}$$

Therefore, we develop a modulation technique in which the minimum distance is maximized to promote more reliable detection and thus yield fewer errors in our estimate of $\hat{s}$.

4.1.2 Piecewise Linear Modulation Scheme

Improved detection performance in the presence of noise motivated our search for a more effective modulation scheme that would provide a large minimum distance. Under the constraint that the embedded signal maintains a Gaussian distribution, the new modulation technique should modulate keyed pseudorandom values with the bilevel message bits and produce a sequence of real numbers that follow a Gaussian distribution and yield a large minimum distance.

The search for a function to transform one Gaussian random sequence into another, under the additional constraint that the distance between the two is maximized, is a daunting task. At the onset, one would have to consider the class of bijections (one-to-one and onto) of the real line. Consequently, we simplified the problem by exploiting the relationship between the Gaussian and uniform distributions.

Beginning with transformations in the uniform domain [0,1] mapping $\mathcal{U}[0,1]$ to $\mathcal{U}[0,1]$ and knowing that uniform variables may be transformed to any continuous distribution, including the Gaussian by inverting the cumulative distribution function (cdf) [48], we search for a transformation to maximize d_{min} of (31). Once a transformation is obtained, mapping one uniform to another in a one-to-one and onto fashion, the uniform value and its transform could then be converted to Gaussian values and used as modulation points.

Let us look at the transformation that produces the modulation points for the sign modulation method that is presented in section 4.1.1. Assume that $U = \{u_0, u_1, \ldots, u_n\}$ are

$\mathcal{U}(0,1)$ variates. Now assume that the transform f of (32), depicted graphically in Figure 11, produces $f(U) = \{f(u_0), f(u_1), \ldots, f(u_n)\}$, which is also uniformly distributed on (0,1).

$$f(u) = 1 - u \qquad 0 \leq u \leq 1. \tag{32}$$

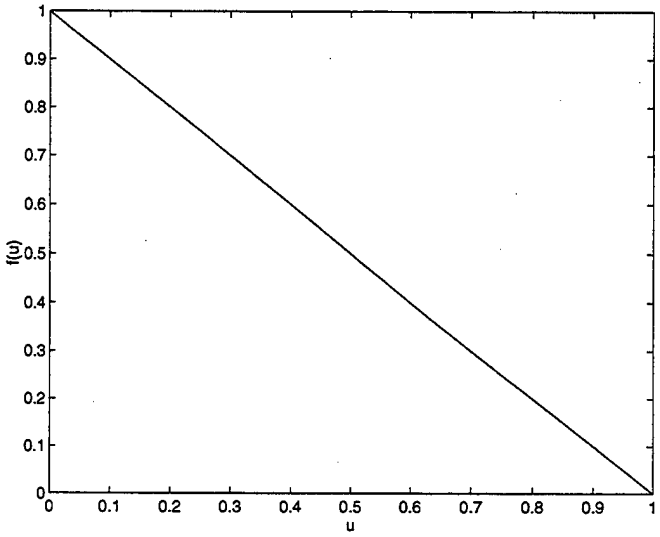


Figure 11. Transformation for Sign Detector Modulation.

The embedded signal is constructed as

$$s(u,m) = \begin{cases} \Phi^{-1}(u) & m = -1 \\ \Phi^{-1}(f(u)) & m = +1, \end{cases} \tag{33}$$

where Φ^{-1} denotes the inverse cdf for a Gaussian random variable.

From Figure 11, observe that at the point $u = \frac{1}{2}$, the transformation $f(u) = u$, and by (29) and (33), the distance $D = 0$ and the message m cannot be recovered.

Now consider a transformation, uniform to uniform, that has a single discontinuity at $u = \frac{1}{2}$. Given a pseudorandom sequence $U = \{u_0, u_1, \ldots, u_n\}$ from $U(0,1)$, generate a sequence $g(U) = \{g(u_0), g(u_1), \ldots, g(u_n)\}$ under the transformation $g(\cdot)$, of (34), shown graphically in Figure 12.

$$g(u) = \begin{cases} u + \frac{1}{2} & 0 \leq u < \frac{1}{2} \\ u - \frac{1}{2} & \frac{1}{2} < u \leq 1 \\ 0 & otherwise. \end{cases} \tag{34}$$

To encode, each element of the embedded signal sequence, s, is formed by selecting from U or $g(U)$, arbitrated by the elements of the modulating sequence, m, and transformed to a Gaussian random value, as in (35).

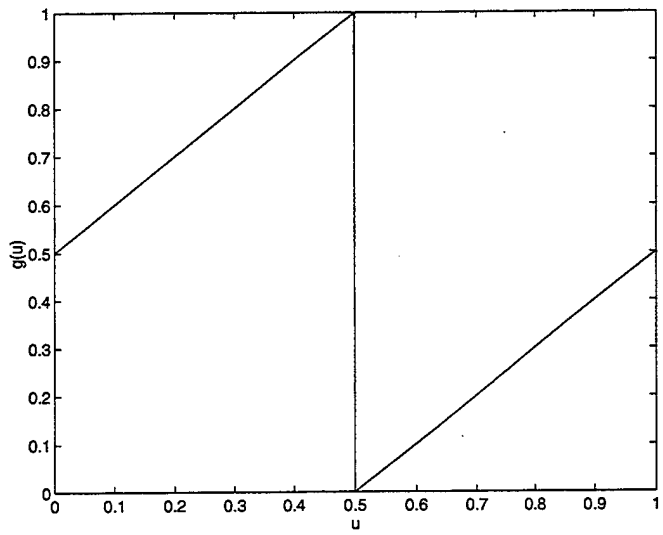


Figure 12. Transformation for Piecewise Linear Modulation.

$$s(u,m) = \begin{cases} \Phi^{-1}(u) & m = -1 \\ \Phi^{-1}(g(u)) & m = +1. \end{cases} \tag{35}$$

Demodulation of the embedded sequence is accomplished by regenerating both $\Phi^{-1}(U)$ and $\Phi^{-1}(g(U))$ at the receiver and calculating a threshold as the midpoint between the modulated values. The estimated embedded signal is then compared to this threshold to determine the value of the encoded message.

We next demonstrate that our transformation $g(\cdot)$ is an optimal solution in the search for the transformation that maximizes the minimum distance between the two possible modulated values, as in (36),

$$g* = \arg\max_{g} \min_{0<u<1} \left| \Phi^{-1}(u) - \Phi^{-1}(g(u)) \right|. \tag{36}$$

To confirm this, we first establish that the transform in (34) does in fact produce a random sequence that corresponds to a uniform distribution. This fact is important to assure that when the inverse Gaussian cdf, $\Phi^{-1}(\cdot)$, is applied, the inverse transformed value will follow a Gaussian distribution. Then we show that the transform in (34), is an optimal solution for maximizing the minimum distance in the uniform space. We then proceed to quantify the value of minimum distance between $\Phi^{-1}(U)$ and $\Phi^{-1}(g(U))$. Finally, we present a numerically demonstration that suggests this minimum distance is in fact an optimal solution to our problem (36).

4.1.2.1 Distribution of Transformed Variable. We have introduced the stochastic constraint that the transform $g(\cdot)$ (34) preserve the uniform probability distribution of the input. This preservation is necessary because the output of the transform is then converted

to a Gaussian variable using the inverse cdf. If the transformation does not produce a sequence that is uniformly distributed, the output of the inverse cdf will not have a Gaussian distribution.

We substantiate that $g(U)$ is distributed uniformly on the unit interval by establishing that the moment-generating function for the uniform distribution function is equal to the moment-generating function for $g(\cdot)$.

Definition: If $f(x)$ is the probability distribution function of the random variable X, the moment-generating function [49] of $g(X)$ is given by:

$$M_{g(x)}(t) = \int_{-\infty}^{\infty} e^{tg(x)} f(x) dx. \tag{37}$$

Now let $u \sim U(0,1)$ and let

$$g(u) = \begin{cases} u + \frac{1}{2} & 0 \leq u < \frac{1}{2} \\ u - \frac{1}{2} & \frac{1}{2} < u \leq 1 \\ 0 & otherwise. \end{cases} \tag{38}$$

Then the moment-generating function for $g(U)$ is

$$\begin{aligned} M_{g(u)}(t) &= \int_{-\infty}^{\infty} e^{tg(u)} f(u) du, \\ &= \int_0^{\frac{1}{2}} e^{t(u+\frac{1}{2})} \cdot 1 \cdot du + \int_{\frac{1}{2}}^{1} e^{t(u-\frac{1}{2})} \cdot 1 \cdot du, \end{aligned} \tag{39}$$

where

$$\int_0^{\frac{1}{2}} e^{t(u+\frac{1}{2})} du = e^{\frac{t}{2}} \int_0^{\frac{1}{2}} e^{tu} du = e^{\frac{t}{2}} (\frac{1}{t} e^{tu}) \Big|_0^{\frac{1}{2}} = \frac{e^{\frac{t}{2}}}{t} (e^{\frac{t}{2}} - 1), \tag{40}$$

and

$$\int_{\frac{1}{2}}^{1} e^{t(u-\frac{1}{2})} du = e^{-\frac{t}{2}} \int_{\frac{1}{2}}^{1} e^{tu} du = e^{-\frac{t}{2}} (\frac{1}{t} e^{tu}) \Big|_{\frac{1}{2}}^{1} = \frac{e^{-\frac{t}{2}}}{t} (e^{t} - e^{\frac{t}{2}}), \tag{41}$$

which results in

$$\begin{aligned} M_{g(u)}(t) &= \frac{1}{t}(e^t - e^{\frac{t}{2}} + e^{\frac{t}{2}} - 1) \\ &= \frac{1}{t}(e^t - 1). \end{aligned} \tag{42}$$

Consider the Uniqueness Theorem [50]. Let X and Y be two random variables with moment-generating functions $M_X(t)$ and $M_Y(t)$, respectively. If $M_X(t) = M_Y(t)$ for all values of t, then X and Y have the same probability distribution.

The moment-generating function for a uniform probability distribution on the interval (0,1) [51] is given as

$$M_{U(0,1)}(t) = \frac{1}{t}(e^t - 1). \tag{43}$$

Note that the moment-generating function of (42) is the same moment-generating function as that for a uniform distribution $U(0,1)$ (43). By the Uniqueness Theorem, we have shown that the transformation $g(U)$ produces a uniform random variable.

4.1.2.2 Optimality - Uniform Distribution. **Theorem:** The function $g(\cdot)$ is an optimal solution maximizing the minimum distance between U and $g(U)$.

Let $\mathcal{F}$, with $f \in \mathcal{F}$, denote the class of bijections (one-to-one and onto) of the unit interval $[0,1]$. For $x = \frac{1}{2}$, the distance is expressed by the inequality

$$|x - f(x)| = \left|\frac{1}{2} - f(\frac{1}{2})\right| \leq \frac{1}{2} \quad \forall\, f \in \mathcal{F}; \tag{44}$$

therefore,

$$\max_{f \in \mathcal{F}} \min_{x \in [0,1]} |x - f(x)| \leq \frac{1}{2}. \tag{45}$$

Now consider our transformation of (34), since distance is

$$|x - g(x)| = \frac{1}{2} \quad \forall\, x \ \in [0,1]. \tag{46}$$

Then

$$\min_{x \in [0,1]} |x - g(x)| = \frac{1}{2}, \tag{47}$$

but $g \in \mathcal{F}$; therefore, $g(\cdot)$ is an optimal solution for (45).

By the aforementioned, we have shown that the piecewise linear function $g(\cdot)$ is an optimal transformation from one uniform random variable to another, which maximizes the minimum distance between the two.

4.1.2.3 Minimum Euclidean Distance. Our initial goal was to generate two normally distributed random variables for each message bit. Because the variables would represent the two possible values of the message bit, the variables must have maximum Euclidean distance. Ordinarily, to find the maximum or minimum of a continuous distance function, we would solve for the points where the derivative is equal to 0. However, the inverse Gaussian cdf, Φ^{-1}, does not exist in closed form, so we are denied the used of a direct method to determine the minimum value of the distance.

We begin by expressing the distance between the possible modulation values of our piecewise linear transformation as

$$D = \left|\Phi^{-1}(u) - \Phi^{-1}(g(u))\right|. \tag{48}$$

Now consider the distance for the range of $0 \leq u < \frac{1}{2}$

$$D_{0 \leq u < \frac{1}{2}} = \Phi^{-1}(u) - \Phi^{-1}(u + \frac{1}{2}) \qquad 0 \leq u < \frac{1}{2}, \tag{49}$$

Definition: The relationship between the normal value, x_u, and the uth percentile can be expressed as

$$\Phi(x_u) = \int_{-\infty}^{x_u} \phi(t)dt = u; \qquad \Phi^{-1}(u) = x_u \qquad 0 \leq u \leq 1. \tag{50}$$

with $\phi(\cdot)$ representing the Gaussian density function.

Let

$$u \in (0, \frac{1}{4}); \qquad g(u) \in (\frac{1}{2}, \frac{3}{4}). \tag{51}$$

$$\Phi^{-1}(u) = x_u; \qquad \Phi^{-1}(u + \frac{1}{2}) = x_{(u+\frac{1}{2})}. \tag{52}$$

$$\Phi^{-1}(u) - \Phi^{-1}(g(u)) \;=\; x_u - x_{(u+\frac{1}{2})}. \tag{53}$$

Now consider for some $\epsilon > 0$ and arbitrarily small,

$$u + \epsilon \in (0, \frac{1}{4}); \quad g(u+\epsilon) = u + \frac{1}{2} + \epsilon \in (\frac{1}{2}, \frac{3}{4}) \tag{54}$$

$\forall\ \epsilon > 0$ satisfying (54),

$$\exists\ \ \delta = \delta(u, \epsilon) \quad \int_a^b \phi(t)dt = \epsilon, \qquad \int_c^d \phi(t)dt = \epsilon, \tag{55}$$

where

$$a = x_u \qquad b = x_{(u+\epsilon)} = x_u + \delta_1(\epsilon), \tag{56}$$

$$c = x_{u+\frac{1}{2}} \quad d = x_{(u+\frac{1}{2}+\epsilon)} = x_{(u+\frac{1}{2})} + \delta_2(\epsilon). \tag{57}$$

But for $u \in (0, \frac{1}{4})$, $\phi(x)$ for $x_u < x < x_u + \delta_1(\epsilon)$ is everywhere less than $\phi(x)$ for $x_{(u+\frac{1}{2})} < x < x_{(u+\frac{1}{2})} + \delta_2(\epsilon)$.

Using the Mean-Value Theorem [52], Let f be continuous on the closed interval $[a, b]$. Then there is some number x' such that $a \leq x' \leq b$ and

$$\int_a^b f(x)dx = f(x')(b-a). \tag{58}$$

Hence, since the integrals of (55) are equal to ϵ, then using the Mean-Value Theorem: $\phi(x')(b-a) = \phi(x'')(d-c)$ with $\phi(x') < \phi(x'')$; therefore, $b - a > d - c$. In other words,

$$x_u - x_{(u+\frac{1}{2})} < x_{(u+\epsilon)} - x_{(u+\frac{1}{2}+\epsilon)}. \tag{59}$$

But $u \in (0, \frac{1}{4})$ is arbitrary and $\epsilon > 0$ is arbitrarily small; therefore, by (59), $\Phi^{-1}(u) - \Phi^{-1}(g(u))$ is strictly monotonically increasing on the interval $(0, \frac{1}{4})$.

Now consider the value $u = \frac{1}{4}$. At $u = \frac{1}{4}$, $g(u) = \frac{3}{4}$ and $\Phi^{-1}(\frac{1}{4}) = x_{\frac{1}{4}}$, depicted in Figure 13, but recall, by symmetry, that $x_\alpha \equiv -x_{1-\alpha}$ for $0 \leq \alpha \leq 1$, so $\Phi^{-1}(\frac{3}{4}) = -x_{\frac{1}{4}}$. Therefore,

$$\Phi^{-1}(\frac{1}{4}) - \Phi^{-1}(\frac{3}{4}) = x_{\frac{1}{4}} - (-x_{\frac{1}{4}}) = 2x_{\frac{1}{4}}. \tag{60}$$

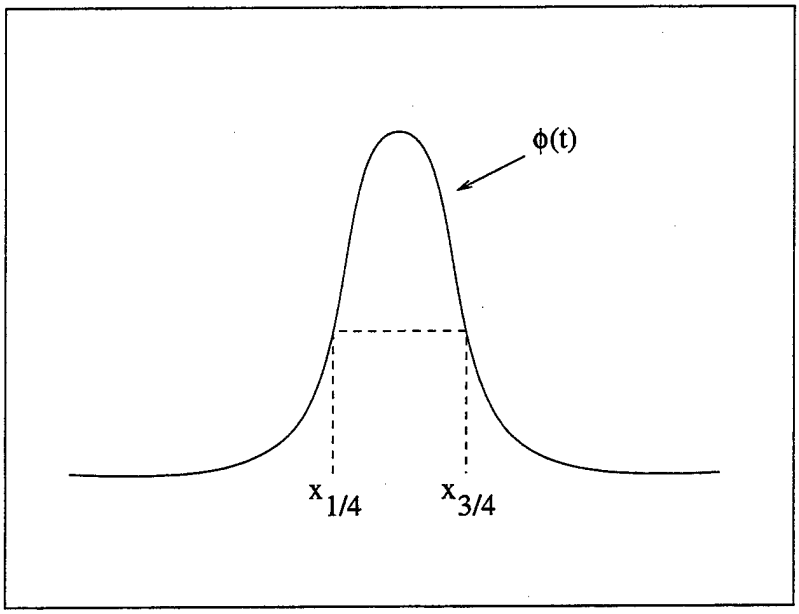


Figure 13. Distance Between Modulation Points, $u = \frac{1}{4}$.

We would like to express this distance as a function of the standard normal variate, $x \sim N(0, \sigma^2)$, appearing in (60). To do so, we make use of the familiar transformation

$$\frac{x-\mu}{\sigma} = z, \tag{61}$$

where $z \sim N(0,1)$. By virtue of symmetry, a completely analogous argument holds for the remaining subintervals $(\frac{1}{4}, \frac{1}{2})$, $(\frac{1}{2}, \frac{3}{4})$, and $(\frac{3}{4}, 1)$. A plot of the signed distance, (29), is shown in Figure 14, with the noise power $\sigma^2 = 1$. As can be seen, the minimums occur at $u = \frac{1}{4}$ and $u = \frac{3}{4}$, with $d_{min} = 2x_{\frac{1}{4}} = 2\sigma z_{\frac{1}{4}} = 1.35$ $(\sigma = 1)$, consistent with (60).

To compare the distance measure for this modulation with the sign modulation of section 4.1.1, a graph of the Euclidean distance for the sign modulation method is shown in Figure 15. The minimum distance for the sign modulator occurs at $u = \frac{1}{2}$ and is equal to 0, which is significantly less than the minimum distance for our piecewise linear modulation.

At the decoder, the modulated values can be correctly detected if the distortion does not exceed the minimum distance of $\lfloor \frac{2\sigma z_{0.25}}{2} \rfloor$. Furthermore, as intuition would reason, this minimum distance is proportional to the power of the noise signal, σ^2, and can be adjusted to achieve desired performance.

4.1.2.4 Results for Gaussian Distribution. We now discuss the specific properties that our target function must possess. As an alternative approach to confirming that our function preserve the distribution of the input sequence, a direct method using properties of calculus and probability is used.

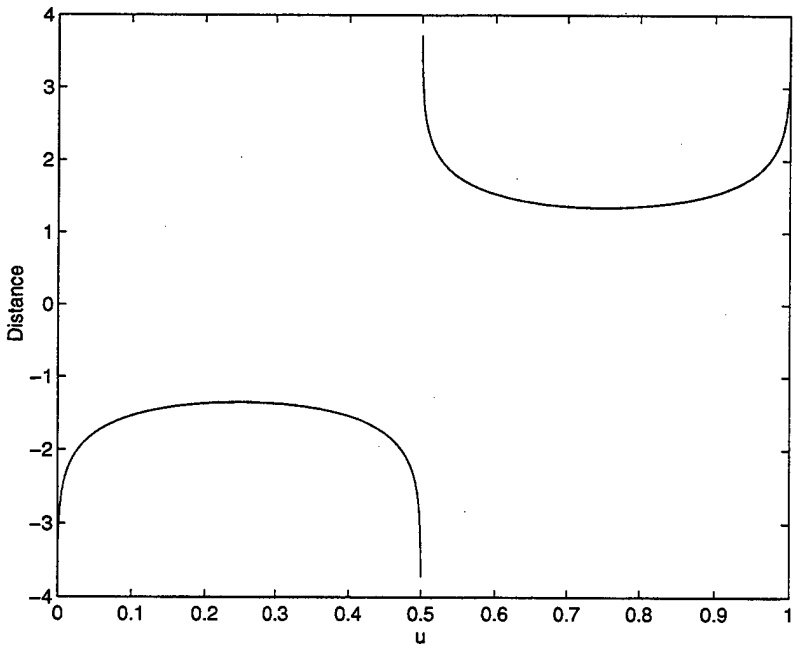


Figure 14. Piecewise Linear Modulation, Euclidean Distance.

To begin, let X be a continuous random variable with density $f_X(x)$ and let $Y = g(X)$ be an invertible transformation of X. To determine the density of the transformed variable, $f_Y(y)$, notice that

$$P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)). \tag{62}$$

Then rewriting this relationship, we have

$$\int_{-\infty}^{y} f_Y(t)dt = \int_{-\infty}^{g^{-1}(y)} f_X(t)dt. \tag{63}$$

After differentiating both sides of (63) with respect to y, we obtain

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \frac{d\, g^{-1}(y)}{dy}, \tag{64}$$

the expression for the probability distribution of the transformed variable.

Suppose we require the $f_Y(y) = 1,\ 0 \leq y \leq 1$, (i.e., $g(X) \sim U(0,1)$). This means that $f_X(g^{-1}(y)) = 1$ for $0 \leq y \leq 1$. But from (64), we have that $\frac{d\, g^{-1}(y)}{dy}$ must be equal to one (actually, $\left|\frac{d\, g^{-1}(y)}{dy}\right| = 1$, to allow for both monotonic increase or decreasing $g(X)$ [53]). This means that for any invertible transformation that preserves the probability distribution of the input, the transformation must have a slope equal to ± 1. This gives us a better understanding of how the graph of the transformation should look in order to preserve the probability properties.

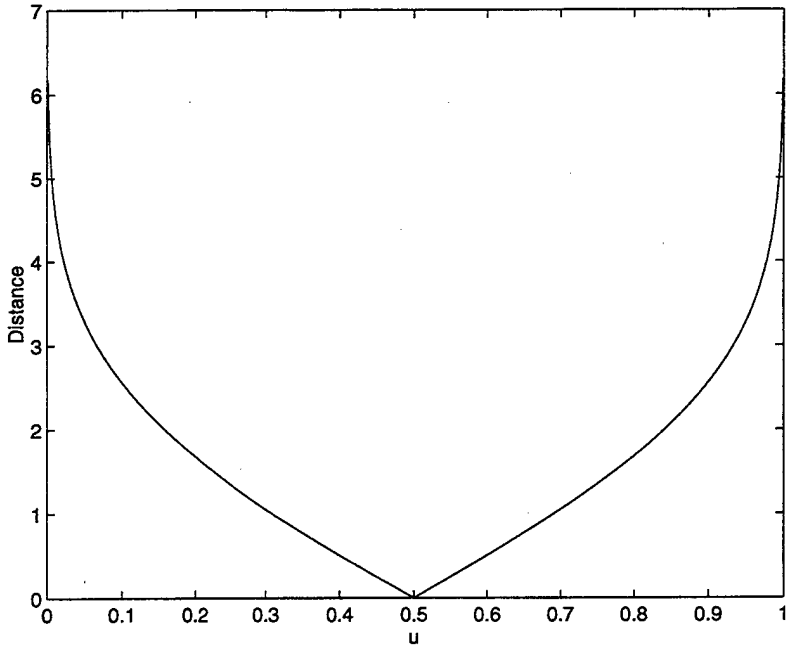


Figure 15. Sign Modulation, Euclidean Distance.

For our case, X is uniform on the interval $[0,1]$ and $Y = g(X)$, where $g(X)$ is defined by (34). Now consider the value of $\frac{d\,g^{-1}(y)}{dy}$ for our transformation $g(\cdot)$. The inverse $g^{-1}(y)$ has a discontinuity at $y = \frac{1}{2}$, so the derivative $\frac{d\,g^{-1}(y)}{dy}$ does not exist at that point; however, $\frac{d\,g^{-1}(y)}{dy} = 1$ for $0 < y < 1$ and $y \neq \frac{1}{2}$. Therefore, $f_Y(y) = f_X(g^{-1}(y)) = 1$ for $0 < y < 1$ with $y \neq \frac{1}{2}$. In other words, the random variable Y is distributed uniformly on the unit interval $[0,1]$ except for a set of measure zero (at $y = \frac{1}{2}$). Moreover, any transformation of the uniform random variable X, where $Y = g(X)$, must be of the form $Y = X + c$, where c is a constant, for the stochastic constraint to be satisfied. Since our transform, $g(\cdot)$, is of this form, we have shown again that it will preserve the distribution of the input.

Thus far, we have shown that our candidate function $g(U)$ adheres to the stochastic constraint and is optimal in maximizing the minimum distance in the uniform domain. Furthermore, we have quantified the minimum distance in the Gaussian domain, which is a function of the noise power σ^2. At this point in our investigation, we have some familiarity with the way in which the optimal transform must behave. In fact, we know the optimal transform must have a slope equal to ± 1 in order to preserve the stochastic constraint. This restricts our class of functions $\mathcal{F}$ to functions of the form $Y = g(X) = X + c$, where c is a constant.

We have considered the sign modulation transformation and showed that the minimum distance is equal to 0. We have allowed a single discontinuity at $u = \frac{1}{2}$ for our candidate function, resulting in a minimum distance of $2\sigma z_{\frac{1}{4}}$. What if we allow the location of this single discontinuity to vary along $[0,1]$. Is it possible to achieve a greater minimum distance? To

address this question, we have conducted a numerical investigation in which the minimum of the distance (48), was evaluated over a class of transformations $g(\cdot)$ from (65) with Δ varying between 0 and 1.

$$g(u) = \begin{cases} u + (1 - \Delta) & u \leq \Delta \\ u - \Delta & \Delta > u \\ 0 & otherwise. \end{cases} \tag{65}$$

Figure 16 illustrates the results of this numerical investigation. From this graph, we see that the minimum distance reaches an apex when the discontinuity location is $\frac{1}{2}$ (as is the case with our transform $g(U)$), resulting in a minimum distance equal to 1.35 for $\sigma = 1$.

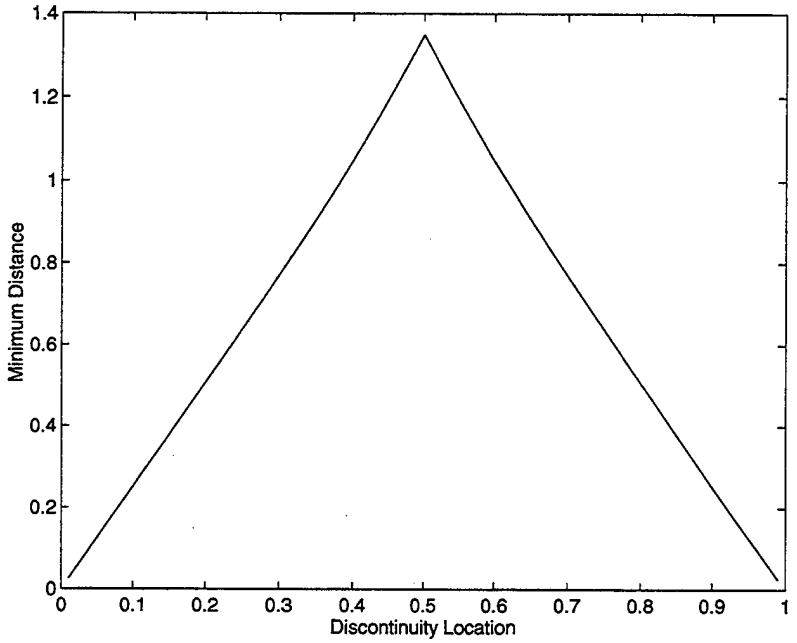


Figure 16. Minimum Distance vs. Discontinuity Location.

Our numerical results strongly suggest that our mapping is an optimal choice from all possible mappings in the class $\mathcal{F}$ with a single discontinuity that provides the largest minimum distance.

4.1.3 Summary

Under the fundamental premise of SSIS, the embedded information must have the same probability distribution as noise that appears naturally in the cover. In this subsection, we have discussed two techniques to combine the information-bearing signal with the characteristic pseudorandom noise. The initial sign modulation technique has been presented and its detection performance discussed. We have shown that for improved recovery of the embedded signal, it is necessary to maximize the minimum distance between the modulated values, and a second modulation technique has been developed with this goal in mind. This technique has been illustrated and its performance quantified. In addition, we have presented

the conjecture that the piecewise linear transformation of this method is an optimal solution to maximizing the minimum distance between the modulation points.

For the SSIS system, the message signal is modulated using the piecewise linear technique to construct the embedded signal, s, which is then added to the digital cover image. The image is subsequently quantized and clipped to become the stegoimage. The selection of the power of the embedded signal, σ^2, is based on the balance between human perception and that needed for reliable decoding.

Although an intruder may be aware of the general strategy of the system, he/she would have difficulty establishing whether the noise in the stegoimage is attributed to the image itself, the transmission channel, or an embedded signal. Without the necessary keys, the modulated signal is statistically indistinguishable from white Gaussian noise and is as secure, in a cryptographic sense, as the pseudorandom number generator used within the system.

4.2 Channel Estimation

For the SSIS system, the hidden data is embedded within WGN that is subsequently added to the image pixels. If the recipient were to possess a copy of the original cover image, these pixels could be subtracted from those of the stegoimage to result in a near perfect reproduction of our embedded signal, neglecting quantization error. This situation, where the cover image characteristics are known, is analogous to transmission through an essentially noiseless channel. However, for our scenario we require a blind system where the recipient does not need to possess the original cover in order to recover the hidden information. Consequently, at the receiver, the embedded signal must be extracted from the stegoimage without any a priori channel information.

4.2.1 Embedded Signal Recovery

Because a receiver does not have the original image, it must approximate the image from the stegoimage; in other words, the receiver must estimate the channel, where the channel in question is a natural grayscale cover image to which WGN has been added. Given that this type of image is rich in low-frequency content [44], we can use image restoration techniques to construct an approximation of the original image. This approximation can then be subtracted from the stegoimage, leaving an estimate of our embedded signal. In this way, we eliminate the need for the recipient to possess a copy of the cover image and yield a reasonable estimate of the embedded signal for decoding purposes.

The restored image can be obtained with a variety of image processing filters such as mean or median filters [44], wavelet shrinkage techniques [54], or adaptive filters [55]. The embedded signal estimate, derived from the stegoimage using image restoration, and an identical copy of the pseudorandom wideband waveform used at the encoder allow demodulation of the hidden data.

4.2.2 Filter Selection

It is plausible to assume that the best-performing filter in this context would be the one that provides the lowest overall MSE between the filtered image and the original cover image, thus providing a restored image that was much like the original cover image in a mean-squared sense. However, through experimentation, we have found that that the MSE was not the appropriate fidelity criterion for our system given our ultimate objective to obtain the most accurate estimate of the embedded signal.

The performance of several restoration techniques was evaluated within the SSIS system. To cite an example, Table 1 exhibits the MSE and resulting embedded signal BER (the BER before the error-control decoder) for a sampling of the filters tested using the Lena image as a cover with an embedded signal power (variance) of 20. From our empirical data, a sample of which is shown in this table, the filter that typically produces a restored image with the lowest MSE is the adaptive Wiener (AW) filter, implemented using Lee's algorithm [56]. However, this filter does not provide the lowest embedded signal BER. The alpha-trimmed mean (ATM) filter presented in Bednar and Watt [57] provides the lowest embedded signal BER of the tested filters.

Table 1. SSIS Image Restoration Filter Performance.

Image Restoration	MSE	Embedded Signal BER
Mean Filter	43.72	0.191
ATM Filter	36.43	0.189
Median Filter	23.50	0.208
AW Filter	10.62	0.254

In the following subsections, we discuss two of the restoration filters from Table 1: the AW filter, which provides the lowest MSE, and the ATM filter, which produces the lowest embedded signal BER. Additionally, we compare the two filters.

4.2.2.1 AW Filter. Initially, the AW filter was used by SSIS to reduce the amount of low-level additive random noise in the stegoimage. Wiener filtering preserves the signal while eliminating noise in the degraded image. Due to linear independence between the cover image and the embedded signal, the optimal linear minimum MSE estimate of the original image is obtained by filtering with an AW filter [55]. The frequency response of the filter is dependent upon the power spectra of the original image and noise as shown in (66), where P_f is the power spectrum of the original image f and P_s is the power spectra of the embedded signal, s. The filter reflects a bivariate spectrum because the image data is evaluated in two dimensions.

$$H(\omega_1, \omega_2) = \frac{P_f(\omega_1, \omega_2)}{P_f(\omega_1, \omega_2) + P_s(\omega_1, \omega_2)}. \tag{66}$$

The power spectrum of the AWGN, which is constant and independent of ω_1 and ω_2, is known at the receiver from the regenerated sequence. Although the embedded signal characteristics do not change within the stegoimage, the image characteristics do change from one region to another. For instance, consider an image with smooth background areas and a detailed foreground; the power spectrum will differ significantly in these areas. To compensate for the changing image characteristics, the AW filter is a space-variant filter whose filter coefficients change as a function of the local image statistics. Adaption to the local image characteristics can be performed on a pixel-by-pixel or block-by-block basis.

The power spectrum of the original image is not known at the receiver and, therefore, must be estimated from the received stegoimage, $\hat{h}$. If we assume that the original image signal, $f(n_1, n_2)$, of a small local region of the image is stationary, it can be reasonably modeled as (67), where m_f and σ_f are the local mean and standard deviation of the original image, and w is a zero mean, white noise process with unit variance [58][59].

$$f(n_1, n_2) = m_f + \sigma_f w(n_1, n_2). \tag{67}$$

To estimate these parameters from the stegoimage, consider that when the mean of the embedded signal is zero, as is the case with the AWGN embedded signal, m_f is identical to the mean of the local region of the stegoimage, m_h. Additionally, because s is additive, σ_h^2 can be defined as (68) and an estimate of $\hat{\sigma}_f^2$ can be obtained by (69), where $\sigma_{\hat{h}}^2$ is the variance of the local region of the received stegoimage. Within this local region, the transfer function of the space-variant Wiener filter is given by (70), and the restored image, $\hat{f}$, is obtained as in (71).

$$\sigma_h^2 = \sigma_f^2 + \sigma_s^2. \tag{68}$$

$$\hat{\sigma}_f^2(n_1, n_2) = \begin{cases} \sigma_{\hat{h}}^2(n_1, n_2) - \sigma_s^2, & if\ \sigma_{\hat{h}}^2(n_1, n_2) > \sigma_s^2 \\ 0, & otherwise. \end{cases} \tag{69}$$

$$H(\omega_1, \omega_2) = \frac{P_f(\omega_1, \omega_2)}{P_f(\omega_1, \omega_2) + P_s(\omega_1, \omega_2)} = \frac{\hat{\sigma}_f^2}{\hat{\sigma}_f^2 + \sigma_s^2}. \tag{70}$$

$$\hat{f}(n_1, n_2) = m_{\hat{h}} + (\hat{h}(n_1, n_2) - m_{\hat{h}}) * \frac{\hat{\sigma}_f^2}{\hat{\sigma}_f^2 + \sigma_s^2} \delta(n_1, n_2). \tag{71}$$

The resultant restored image is scaled according to the relation between $\hat{\sigma}_f^2$, which is estimated from the local region statistics of the stegoimage, and the predetermined σ_s^2. If σ_s^2 is much greater than the contrast of the degraded image, the contrast is assumed to be primarily due to s and is significantly attenuated. Conversely, when the estimated $\hat{\sigma}_f^2$ is greater than σ_s^2, the local contrast is attributed to the original image and little processing is done [55].

4.2.2.2 ATM Filter. The ATM filters form a group of filters whose properties vary between the mean and the median by way of a tuning parameter, l [57]. The particular implementation of the ATM filter used here is an order statistics filter of length N operating on the sequence $\{x_j : j = k - M, \cdots, k, \cdots, k + M\}$, where k is the center sample and N is odd. The output of the filter, y_k, is given by

$$y_k = \frac{1}{N - 2l} \sum_{i=1+l}^{N-l} x_{(i)}^k, \tag{72}$$

where $x_{(i)}^k$ is formed from the elements of x_j arranged in increasing order,

$$x_{(1)}^k \leq x_{(2)}^k \leq \cdots \leq x_{(N)}^k. \tag{73}$$

The filter parameters were selected as those that commonly provided the lowest BER for a collection of images with varying characteristics. The SSIS ATM filter implementation used a length of $N = 9$ indicating a 3×3 pixel window with the tuning parameter $l = 1$. With the tuning parameter l set to one, the minimum and maximum values of the input is eliminated, as in

$$y_k = \frac{1}{N - 2} \sum_{i=2}^{N-1} x_{(i)}^k. \tag{74}$$

In essence, this filter estimates the center pixel by "trimming" the minimum and maximum values within the window, thereby disregarding outliers and subsequently taking the mean of the remaining pixels to smooth the data.

4.2.2.3 Filter Comparison. In this subsection, we demonstrate the effectiveness of these restoration filters within our steganography system. Figure 17 shows a test image that consists of a rectified sine wave covering the left half of the image and a smooth area on the right half. The pixel values for a single row of this image are illustrated in Figure 18. The image is indicative of the edges and flat regions encountered in natural imagery.

We embedded information in the test image using the SSIS system and a embedded signal (stegosignal) power of 20 to produce the stegoimage shown in Figure 19. We then used two different decoders, differing only in image restoration filter, to decode the hidden information. Figure 20 shows the error maps for both filters; a white pixel indicates that an embedded signal decoding error occurred at that location, and a black pixel reflects no error. As can be seen in this figure, many errors are incurred using the decoder with the AW restoration filter. In fact, the embedded signal BER is 0.2494 even though the MSE between the stegoimage filtered by the AW filter and the original cover is equal to 10.15. For the ATM restoration filter, the MSE is 55.62; however, the embedded signal BER is only 0.0699, indicated by the small number of white pixels.

From this experiment, we demonstrate that although the errors between the original image and the Wiener filtered image were small, they were very frequent in number, and

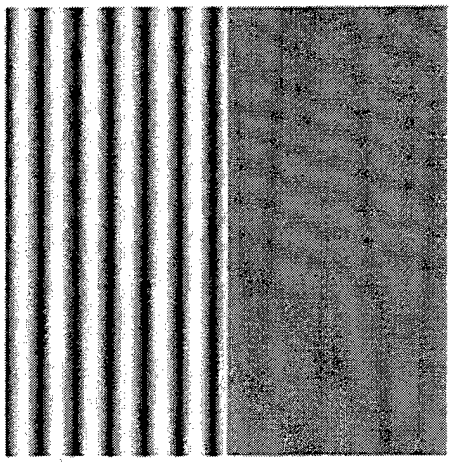


Figure 17. Test Image.

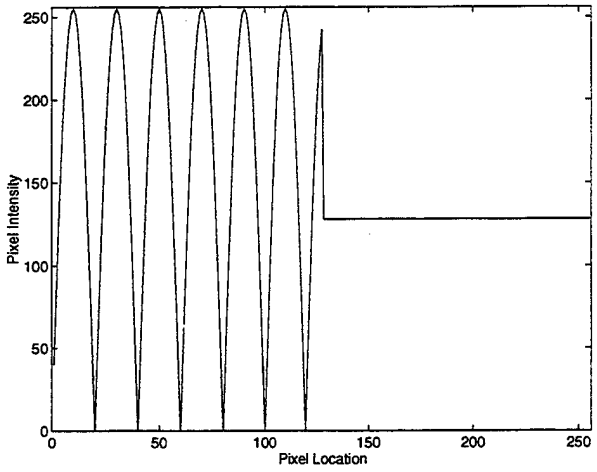


Figure 18. Row of Pixels From Test Image.

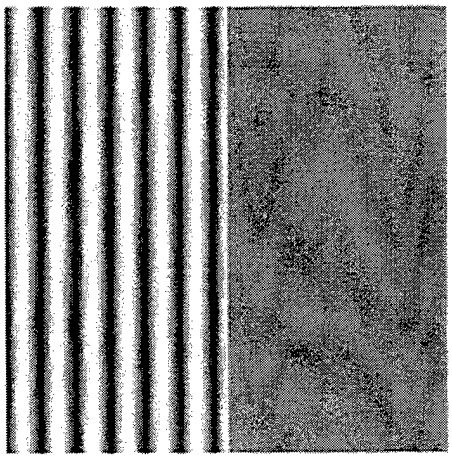
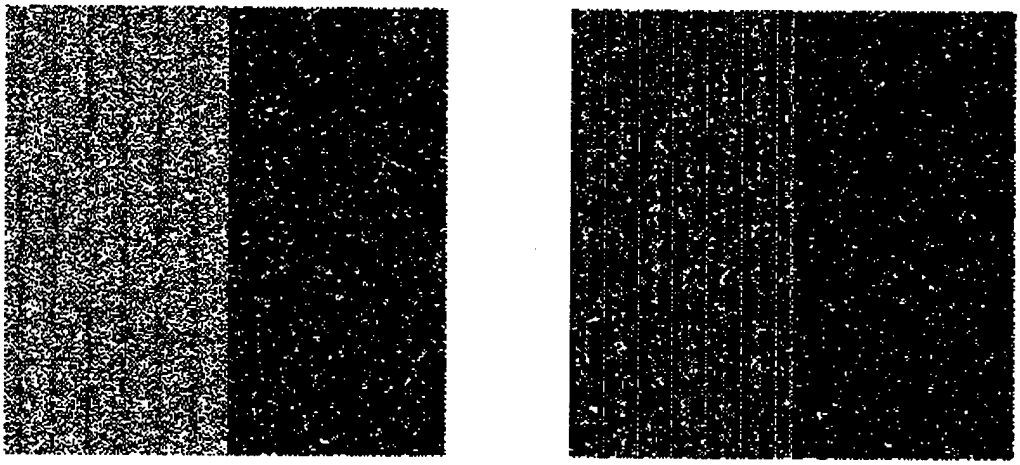


Figure 19. Stegoimage - Test Image.



AW Filter ATM Filter

Figure 20. Embedded Signal Error Map.

the errors encountered using the ATM filter, although much larger in magnitude, were less numerous. In the decoding process, once the error threshold has been traversed, a decoding error occurs and the magnitude of the error is of no consequence. The only relevant factor in the embedded signal BER is that an error has been made; thus, the MSE is not indicative of the filter performance within this system. Let us take a closer look at this phenomenon by comparing the embedded signal estimation for both decoders with the original embedded signal and the decoding threshold. Figure 21 shows actual values for a partial row from our test image.

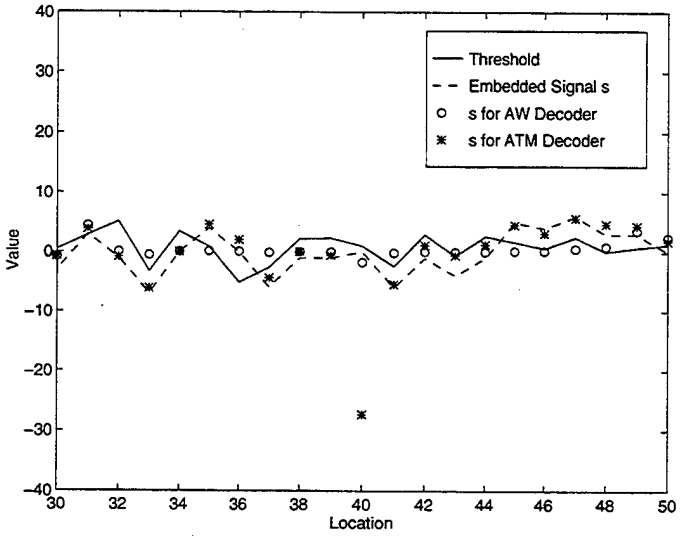


Figure 21. Decoded Values.

The solid line represents the threshold that determines the decoder mapping (whether the hidden data will be decoded as a -1 or a $+1$). The dotted line shows the values of the actual embedded signal, while the "o" represents the embedded signal estimate from the AW

decoder and the "*" represents the embedded signal estimate produced by the ATM decoder. From this plot, we can see that the ATM decoder typically remains on the same side of the threshold, above or below, as the original embedded signal, thus indicating correct decoding. However, the AW decoder values may, at times, lie on the *wrong* side of the threshold, as is the case with locations 45–47, resulting in a decoding error. Also note that at location 40, the ATM decoder has made a very poor estimate of the embedded signal, increasing the MSE without causing a decoding error. Conversely, at location 33, the AW filter value, though close in distance to the embedded signal, lies on the other side of the threshold and would result in a decoding error. From this comparison and the results of experiments using other typical images, we can conclude that the ATM filter provides better overall detection of the embedded signal by more accurately estimating our channel.

4.2.3 Summary

In this subsection, we have discussed the importance of channel estimation to our blind steganographic scheme. We have presented the performance of various restoration techniques, and two specific methods have been compared in detail. Additionally, we have evaluated two fidelity criteria: the MSE and the embedded signal BER. Because our ultimate goal was to minimize the embedded signal BER and not the MSE, we have concluded that the ATM restoration filter offered the best performance in the SSIS system. This filter disregards outliers, smoothes the effects of noise, and more accurately estimates the channel.

The estimate of the embedded signal is compared with an identical copy of the pseudorandom sequence used at the encoder. Even though the channel estimation may yield good performance, the recovery of such a low power signal, necessary to provide the degree of invisibility essential for a steganographic system, may not provide error-free decoding of the embedded signal. Therefore, to compensate for the performance of the embedded signal estimation, we have incorporated the use of error-control coding as discussed in the next subsection.

4.3 Error-Control Coding

For our steganographic system to function as desired, it must be able to decode the hidden message with a small probability of error. However, because the channel estimation does not perfectly replicate the channel, the estimate of the embedded signal is likely to contain errors. This error will cause the estimate of the embedded signal produced by the demodulator to have a substantial number of bit errors, indicated by a high embedded signal BER. For most cover image/stegopower combinations, the BER may be high, at times greater than 25% (see Table 1). To compensate, error-control coding is integrated into the SSIS system.

Any code that is capable of correcting the signal estimation BER can be used within SSIS. This BER is influenced by the stegosignal power, cover image characteristics, and attributes of the transmission channel. In this subsection, we consider operations only for a noiseless transmission channel; the noisy transmission channel is addressed in section 4.4.

The entire decoding process can be simulated at the encoder, thus allowing selection of the proper code by the SSIS system. This assures that the hidden message will be recovered, with high probability, free of errors for the noiseless transmission channel. When the transmission channel is expected to be noisy, an appropriate code may be selected to correct for the additional channel errors. Similarly, error correction may be able to compensate for the errors generated by low levels of noise resulting from lossy image compression of the stegoimage. For our purposes, noise caused by lossy image compression is considered part of the transmission channel.

Of course, the drawback of using codes is that, because they add redundancy by adding extra bits, the amount of payload (hidden data) that can be embedded within an image is reduced. For instance, if the BER is high, a low-rate code must be used to correct the errors in the embedded signal, thereby severely limiting the payload size. However, without them we cannot get closer to capacity. Since the objective of steganography for data hiding is to embed a large amount of data, choosing the highest rate code with the required correction capability is of significant consequence.

It should be noted that interleaving of the hidden data is vital to many error-control operations with short block lengths. By reordering the data, the interleaver disperses a long error burst uniformly over many codewords, causing the errors within each codeword to occur almost independently and giving the code a better opportunity to correct the errors in all blocks. The interleaver does not alter the payload amount, but its benefit to the coding function merits mention.

In this subsection, we discuss the types of codes used within the SSIS system and their specific performance. Additionally, it is shown how *side information* provided by the channel (stegoimage) can be used to assist the decoder in error correction.

4.3.1 Low Rate Error-Control Codes

To correct a large number of errors in the embedded signal, a low-rate code must be used. For some cover images, the SSIS embedded signal BER was exceptionally high, greater than 30%. Therefore, we needed to correct a large number of bit errors. We used a selection of the many codes presented in [60] correct this large number of errors. These codes were derived by expanding traditional Reed-Solomon (RS) codes with a symbol alphabet equal to $Q = 2^j$ to j binary symbols. The selected codes have many low-weight parity checks and are decoded using an iterative decoder developed by Retter [61]. The iterative decoder is based on an idea from Bossert and Hergert [62] using low-weight parity checks. These codes and decoder are capable of correcting many more bit errors than the traditional RS decoders.

For instance, using the (255,4) RS code, whose minimum symbol distance is equal to the BCH bound of $(N + 1 - K) = 252 = d_{min}$, the conventional RS decoder can correct up to $\lfloor \frac{d_{min}-1}{2} \rfloor$ symbol errors. The symbols alphabet contains 256 symbols. This code can be expanded to a (2040,32) binary code because each symbol can be expanded to an 8-bit binary number and decoded with a conventional RS decoder, but any pattern of errors that affected more than 125 symbols would be uncorrectable. The average binary minimum

distance of the (2040,32) codes in Retter [60] is 863.4, suggesting that about 431 binary errors are correctable with the appropriate binary decoder. However, the iterative decoders of Retter [61] can correct virtually all error patterns with weights less than 700.

Table 2 lists these powerful codes, BER capability, the possible SSIS payload in bpp, and the information rate (information bits/image bits). We select the code from this table that is capable of correcting the specific embedded signal BER to encode and decode the message data. However, because (as previously mentioned) these codes are of low rate, the SSIS payload was significantly restricted.

Table 2. Binary Expansion of RS Codes.

Original RS Code	Binary Code	BER Correcting Capability	Payload (bpp)	Information Rate (bits)
(31,8)	(155,40)	0.12	0.2581	0.0323
(63,6)	(378,36)	0.21	0.0952	0.0119
(127,5)	(889,35)	0.27	0.0393	0.0049
(255,4)	(2040,32)	0.34	0.0156	0.0019

4.3.2 Maximum Likelihood Decoding

Since one of our primary goals is to increase the amount of payload, we looked for higher rate codes that were capable of correcting many errors. We explored the use of the binary expansion of the (31,5) RS code decoded via a hard-decision maximum-likelihood (ML) decoder. The binary expansion of this code is a (155,25) binary code that has 2^{25} (approximately 32 million) codewords. The ML decoder compares each block to be decoded with each of the 32 million possible codewords and selects one as the decoded value based on a chosen cost function such as the minimum Hamming distance (the number of places in which two codewords differ). Due to the relatively small number of codewords in this code, it may be feasible to do such comparisons. The code can correct 43 out of 155 bits (0.277 BER) with good performance (successful decoding 98% of the time) and has a payload of 0.1612 bpp and an information rate of 0.0201. This code may replace the (378,36) and the (889,35) codes (in Table 2) in the SSIS system to obtain higher payload for the range of error-correcting capability, provided that slower decoding (due to the 32 million comparisons for each codeword) is permissible.

4.3.3 Soft-Decision Decoding

Again, in an attempt to increase the payload and get closer to the channel capacity, we tried convolutional codes. The output of a convolutional code depends not only on the corresponding input but also on m previous inputs [43]. To decode the convolutional codes, the Viterbi algorithm was used. This algorithm uses a trellis representation of the encoder

state machine mapped along time axis. A metric, such as the Hamming distance, was used to select the most probable path, indicating the sequence of encoded bits, through a trellis.

The rate 1/6 convolutional code, with constraint length 15, developed by the Jet Propulsion Laboratory for the Mars Pathfinder Telemetry Link [63], was implemented using the Viterbi algorithm [64]. Note that there are several definitions of constraint length [65]; in this case, the constraint length is the number of information bits upon which each output bit depends. The payload of this code is 0.1666 bpp with an information rate of 0.0208 (48 image bits to 1 information bit). This is a slight increase in rate compared to the ML code discussed previously, but decoding is much faster.

To demonstrate decoder performance, a stegoimage constructed from the Eiger cover image with a stegosignal power of 80, shown in Figure 22, is used. The demodulated embedded signal extracted from the stegoimage has a BER of 0.23. Using the rate 1/6 convolutional code along with the Viterbi decoder, the BER is improved from 0.23 to 0.0117. Although this performance is favorable, it is possible for the Viterbi algorithm to perform even better with soft inputs. In general, soft-decision decoders outperform hard-decision decoders [43].

Original Eiger Image

Eiger80 Stegoimage

Figure 22. Eiger Image, Original and Stegoimage.

Forney showed that the Viterbi algorithm was a maximum-likelihood decoding algorithm for convolutional codes [66], meaning that the decoder output selected is always the code sequence that gives the largest value of the log-likelihood function. The Viterbi algorithm uses the log of the likelihood ratio, shown in (75), as a metric to select the most probable encoded bit sequence. In (75), x represents the observation and H_1 and H_0 the hypotheses [67].

$$\log L(x) = \log \left(\frac{P(x|H_1)}{P(x|H_0)} \right). \tag{75}$$

Each hypothesis represents the event that the source-generated one of two possible values. The values occur at the source with probability P_0 and P_1, respectively. If the likelihood

ratio, L, is greater than the corresponding ratio of probability of occurrence for the generated values,

$$L = \frac{P(x|H_1)}{P(x|H_0)} > \frac{P_0}{P_1}, \tag{76}$$

then we can decide that hypothesis H_1 is true; otherwise, we decide that hypothesis H_0 is true. When the source generated values are equally likely, $P_0 = P_1$, H_1 is selected if L is greater than 1. Because the logarithm is a monotonic function, taking the log of L results in an equivalent test. So, if the log-likelihood ratio for the given observation x is positive, we should decide in favor of H_1; if negative, we decide in favor of H_0; and if equal to 0, H_1 and H_0 are equally likely.

Again, let us briefly look at the bipolar modulation system whose source generates two possible values, ± 1, operating in AWGN with unit variance. The relationship between the conditional probability density functions for the value of the observation x given both hypotheses is exhibited in the graph of Figure 23. The log of the likelihood ratio as a function of x is shown in Figure 24. Notice that at $x = 0$, the hypotheses are equally likely and the corresponding log-likelihood ratio is 0. As x moves toward one of the generated values of -1 or $+1$, the log-likelihood ratio decreases, or increases, respectively.

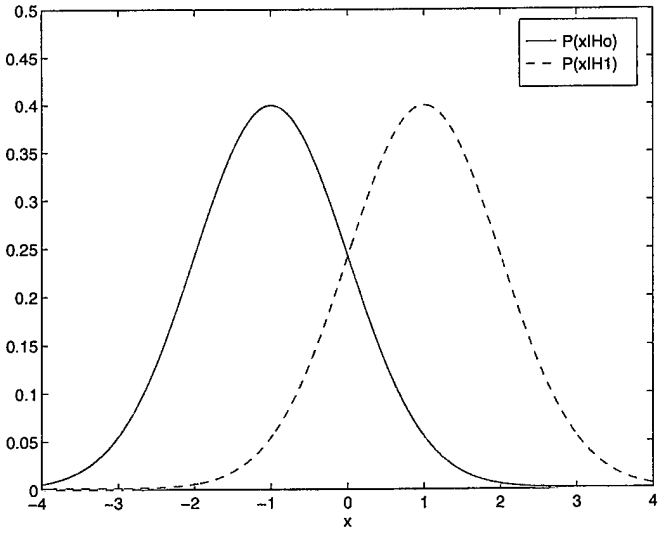


Figure 23. Bipolar Modulation in AWGN.

To ascertain the performance of soft-decision decoding for the SSIS system, it is necessary to calculate the likelihood ratio of the observed embedded signal to be used as input to our soft-decision Viterbi decoder. This task is not easily accomplished given that the SSIS generated values at the source, themselves, are samples of a WGN process. Even the relationship between the possible modulated values is not constant (although the minimum distance has been determined in section 4.1.2). This being the case, we cannot look at the conditional distributions as we have for the bipolar modulation system. However, the generated source can be considered independently.

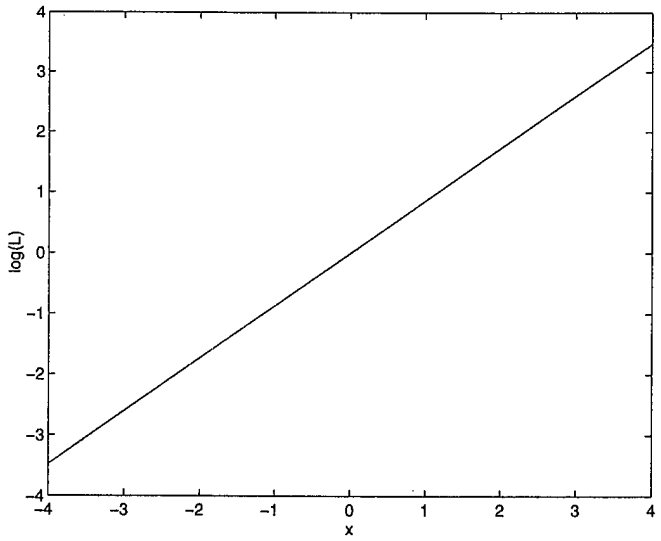


Figure 24. Bipolar Modulation in AWGN, Log-Likelihood Ratio.

Remember that the SSIS system transmits the embedded signal, s, through a channel, the cover image. At the encoder, the embedded signal takes on one of two possible values, dependent upon the hidden message bit. If the hidden message bit is a 0, s_0 is sent through the image and hypothesis H_0 is true; otherwise, s_1 is sent and hypothesis H_1 is true. At the decoder, an estimate of the embedded signal is recovered from the image. During the recovery process, the embedded signal has incurred distortion and is observed as $\hat{s}$. Although $s \sim N(0, \sigma^2)$ and varies with time, at a single instant in time (or space, as is the case with an image), we can observe the deviation of $\hat{s}$ from s and the encoder by simulating the decoder. By normalizing this observation, the conditional probabilities can be calculated for each hypothesis and used to determine the likelihood ratio. Furthermore, since the distortion in $\hat{s}$ is caused by the cover image functioning as a channel, conditional probabilities should vary with each image.

Let us restate this through the following equations which represent the relevant SSIS encoder and decoder functions. The embedded signal, which has been modulated with the message bit, is added to the cover image, f, at the encoder to construct the stegoimage, h:

$$h = f + s. \tag{77}$$

At the decoder, h is received, assuming a noiseless channel, and filtered with the channel estimation filter, denoted as $\mathcal{F}$, in an effort to construct an estimate of the original cover image

$$\hat{f} = \mathcal{F}(h) = \mathcal{F}(f + s). \tag{78}$$

The embedded signal is then recovered as

$$\begin{aligned} \hat{s} &= h - \mathcal{F}(h) \\ &= f + s - \mathcal{F}(f+s) \\ &= s + (f - \mathcal{F}(f+s)) \\ &= s + \eta, \end{aligned} \tag{79}$$

where

$$\eta = f - \mathcal{F}(f+s) = f - \hat{f} \tag{80}$$

is the distortion incident upon the embedded signal. Because $\hat{f}$ is constructed by filtering the stegoimage and is an approximation of the original cover image, f and $\hat{f}$ are highly correlated. It is known that the difference between two highly correlated sources (the stegoimage and restored stegoimage, in this case) can be represented by the double exponential, or Laplacian distribution, shown in (81) [44].

$$f_x(x) = \frac{1}{2\lambda} \exp^{-|\frac{x-\alpha}{\lambda}|}, \tag{81}$$

where α represents the mean of the distribution and the variance, σ^2, is represented by

$$\sigma^2 = 2\lambda^2. \tag{82}$$

Consequently, we attempt to model the distortion in $\hat{s}$ using this Laplacian distribution to generate the conditional probabilities needed to calculate the likelihood ratio. To consider the hypothesis that -1 was embedded, H_0, we use the distribution of the deviation from the various values of s given that s_0 was embedded, $P(\hat{s} - s_0|H_0)$. The variance of the conditional distribution is used to obtain the Laplacian parameter, λ, using the relationship in (82), and the mean of our model, α, is set to 0 because the expected value of $\hat{s} - s_0$ is 0.

Unfortunately, we do not have $P(\hat{s} - s_0|H_0)$ at the decoder, and thus we must estimate the value of λ from the variance of the difference of the stegoimage and its filtered version, $(f - \hat{f})$, which we will call $\hat{\sigma}^2_{(f-\hat{f})}$.

To illustrate the accuracy of our Laplacian model, we computed the actual distribution of $P(\hat{s} - s_0|H_0)$ for the Eiger80 image. The sample variance of the empirical distribution was then used to compute the Laplacian parameter, λ, for the model. Lastly, λ was estimated as it would be at the decoder using the stegoimage. Figure 25 illustrates the association between the actual distribution, the Laplacian model with the true lambda value calculated from the distribution, and the Laplacian model with estimated λ that would be used at the decoder to evaluate the likelihood ratio. Figure 26 shows the error between the distribution and the model and the error between the distribution and the model using the estimated Laplacian parameter.

As the embedded signal power σ_s^2 is reduced, the error between the actual distribution and the estimated model is diminished. For instance, for the Eiger80 stegoimage, $\sigma^2_{(f-\hat{f})}$, the actual distribution variance is approximately 350 and $\hat{\sigma}^2_{(h-\hat{f})}$, the variance estimated from

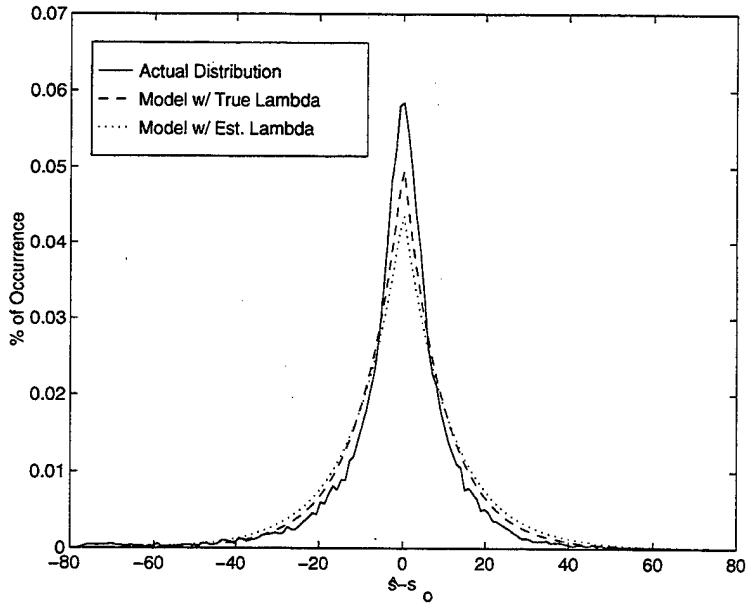


Figure 25. Comparison of Distribution to Laplacian Model.

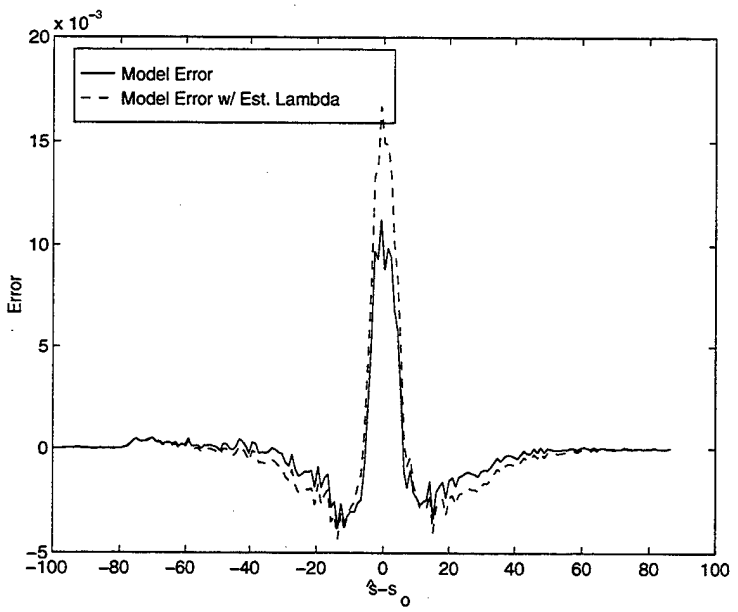


Figure 26. Laplacian Model Error.

the difference of stegoimage and its filtered version, is 406. When the stegoimage power is decreased to 40, the $\hat{\sigma}^2_{(h-\hat{f})}$ is decreased to 369. Hence, as the stegosignal power decreases, there is less distortion in our model.

By using this Laplacian model to calculate the log-likelihood ratio as the soft input for the soft-decision Viterbi decoder for the Eiger80 stegoimage, the BER is reduced from 0.0117, for the hard-decision Viterbi decoder, to 0.0057, which is an improvement of 51%.

4.3.4 Incorporation of Side Information

While seeking to further enhance the coding performance to increase payload, we examined the error map, which specifies the decoding error locations, at the decoder and detected a correlation between decoding errors and the edges within the image. It was then determined that the most unreliable portion of the data occurred around the edges, where the channel estimation performed poorly. For instance, Figure 27 shows the error map for the Eiger80 stegoimage (white pixels representing decoding errors and black pixels indicating correct decoding). Notice many white pixels occur on the outline of the mountain Eiger, and if the error map is examined closely, one can even discern the window and roof line of the buildings in the foreground. Using this relationship, we can modify the soft-decision decoder to use edge information that has been extracted from the received stegoimage as channel side information to better correct bit errors.

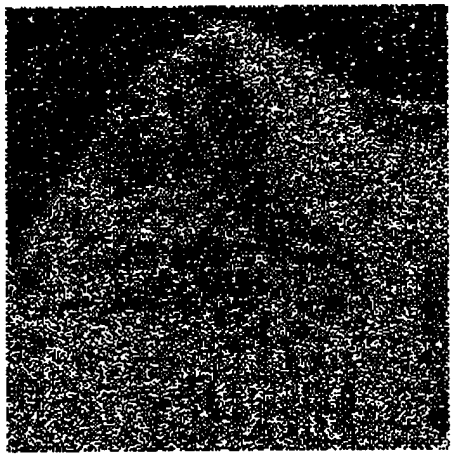


Figure 27. Error Map for Eiger Stegoimage.

We apply this side information to both the hard- and soft-decoder input in a variety of ways. First we used an extracted edge map to signify *erasures* in the hard-decision data. Then in place of edge detection, we used the results of filtering the stegoimage with a variance filter (a filter that calculates the local variance or energy in a group of pixels)[68]. The variance filter indicates edge regions with high variance, while low values of variance are indicative of smooth (nonedge) regions. We used this variance information to *weight* the log-likelihood ratio. Finally, we used a priori information to create a probability of error from the hard-decision data for a given variance as soft-decision data.

4.3.4.1 Erasures. We began first by using the edge map to indicate erasures in our hard-decision data, as shown in the discrete channel depicted in Figure 28.

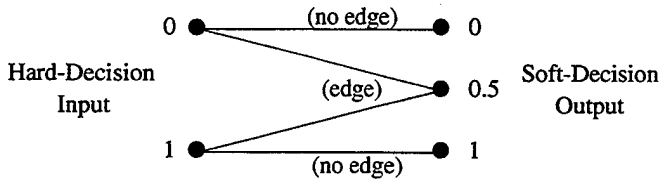


Figure 28. Erasure Channel.

If an edge occurred at a particular location, the corresponding embedded signal bit is considered suspect and denoted as an erasure to a soft-decision decoder. Assuming that our bilevel message bits take on values of $\{0, +1\}$ and by setting the received value of an erasure to 0.5, we indicate the event in which the bit is equal to 1 and the event in which the bit is 0 are equally likely. This knowledge is used by the decoder to select the most probable codeword given the nonerasure inputs for that particular block. These three possible values, 0, 0.5, and 1, are then converted to the corresponding log-likelihood ratios (75) before input to the decoder using the following equation [69]:

$$log\ L = \log\left(\frac{1}{p(x)} - 1\right). \tag{83}$$

The edge map is obtained by using edge-detection techniques such as thresholding the output of the stegoimage that has been filtered by Sobel operators for edge detection [68] or a local variance filter,

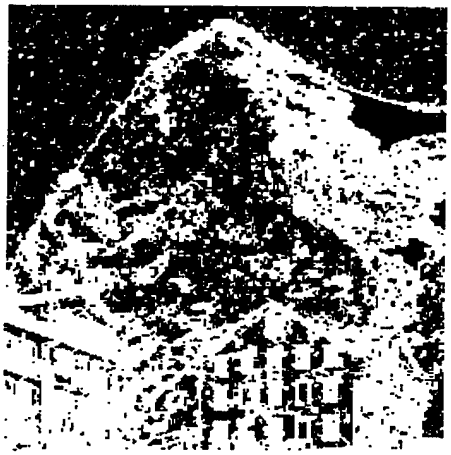
$$\sigma^2(x,y) = \frac{1}{(2X+1)(2Y+1)} \sum_{m=-X}^{X} \sum_{n=-Y}^{Y} [f(x+m, y+n) - \mu(x,y)]^2, \tag{84}$$

where $\mu(x,y)$ is equal to the local mean of the $(2X+1) \times (2Y+1)$ pixel window. For our system, a local neighborhood of size $(2X+1) = 3$ and $(2Y+1) = 3$ was used. If the variance in a particular pixel's local neighborhood is high, indicating an edge, there is a low probability that the demodulated embedded signal value in that location is dependable. On the other hand, if the local variance is low, reflecting a smooth region, the probability that the demodulated signal is correct is considerably higher. The objective of thresholding the output of the edge detection filter is to manufacture an approximation of the binary error map that was shown in Figure 27.

Figure 29 shows the output of a local variance filter for the Eiger80 stegoimage along with the output of the variance filter that has been thresholded at three levels. The variance image has been scaled and equalized to properly depict the relationship of the local variance values. A hard threshold was applied to the variance image at the 50th, 75th, and 85th percentile to obtain the other binary images in the figure, which can be used as *erasure maps* to denote the locations of erasures.

We can define the best-choice threshold as the threshold level at which the erasure map, the thresholded output of the edge detection filter, most resembles the error map. We search

Variance



Threshold = 50%

Threshold = 75%

Threshold = 85%

Figure 29. Thresholded Local Variance.

for the best threshold value that gives a high probability of a *hit*, true detection of an error (85), and a high probability of a true negative, true detection of no error (86), while resulting in a low probability of a *miss*, false detection of no error (87), and a false alarm or false positive, false detection of an error (88).

$$P(hit) = P(edge|error) = \frac{P(edge \bigcap error)}{P(error)}. \quad (85)$$

$$P(true\ negative) = P(!edge|!error) = \frac{P(!edge \bigcap !error)}{P(!error)}. \quad (86)$$

$$P(miss) = P(!edge|error) = \frac{P(!edge \bigcap error)}{P(error)}. \quad (87)$$

$$P(false\ alarm) = P(edge|!error) = \frac{P(edge \bigcap !error)}{P(!error)}. \quad (88)$$

The best-choice threshold cannot be computed without knowledge of the actual error map, which is unknown to the decoder. However, it is possible to obtain this threshold value from the information available at the encoder by using the error map to search for the best-choice threshold using the aforementioned equations. Once found, the threshold can then be embedded within the stegoimage along with the hidden message and used to reconstruct the approximate error map from the edge data obtained at the decoder. The overhead costs incurred by including the threshold within the stegoimage would be minimal, although extra steps should be taken to assure that the threshold is accurately recovered.

To keep our system consistently blind, an erasure map was constructed at the decoder by arbitrarily selecting a threshold for the variance image. In the case of the Eiger80 image, the use of side information at the decoder improved the BER from 0.0117 for hard-decision decoding to 0.007 for hard-decision decoding with erasures. However, we found that inconsistent decoding performance resulted from arbitrary threshold selection.

4.3.4.2 Variance Weights. To eliminate the thresholding issue encountered when using the edge map as erasure input, we decided to use the entire range of local variance measurements of the stegoimage to weigh the confidence of the log-likelihood ratio. Again, if the local variance of a pixel is high, the demodulated embedded signal value of that location is considered less reliable than that of a value whose location has a lower local variance.

The soft-decision decoder input is formed by first filtering the stegoimage with a local variance filter to produce σ_x^2; a single index is used here for brevity. Each local variance value is then scaled to a percentage of the entire variance range for that image. These weights are used to weight the output of the log-likelihood estimator, ll. The input for the soft-decision decoder is then

$$y(x) = \begin{cases} ll - |\frac{\sigma_x^2 - \sigma_{min}^2}{\sigma_{max}^2 - \sigma_{min}^2}| * ll & ll > 0 \\ ll + |\frac{\sigma_x^2 - \sigma_{min}^2}{\sigma_{max}^2 - \sigma_{min}^2}| * ll & ll > 0. \end{cases} \quad (89)$$

For instance, if the local variance measurement at a specific location is 100% of the entire variance range, the log-likelihood ratio would be set to 0, reflecting the uncertainty of the demodulated embedded signal value. Conversely, if the local variance is close to 0, the weighted log-likelihood ratio is approximately equal to the unweighted log-likelihood ratio.

After conducting simulations for several images and stegosignal power combinations, it was confirmed that the decoder performance utilizing side information via local variance weights with soft-decision decoding is more consistent than that of blind thresholding of edges for erasure inputs. When this side information technique was used for the decoding of the Eiger80 stegoimage, the soft-decision BER was reduced from 0.0057 to 0, total error-free recovery of the hidden data. This method operates completely blind, producing reliable results without a priori knowledge such as the threshold value needed for erasures.

4.3.4.3 Probability of Error Given a Variance. Our best soft-decision performance obtained thus far was accomplished by using side information from the local variance that has been weighted by its true probability of error. This true probability of error is calculated for each 1% of the variance range by calculating the ratio of the number of errors that occurred in that range to the total number of locations in that variance range. For instance, if $P(error|variance = v)$ is very high, it is likely that an error will occur in the location where the local variance is equal to v.

This $P(error|variance)$ relationship is used to influence the hard-decision data as soft input for the Viterbi decoder. Of course, this technique is not practical for implementation because, again, the error map is not available at the decoder, but by calculating the $P(error|variance)$ distribution for several images with similar characteristics (i.e., smooth sky, detailed foreground, etc.), a generalized table can be constructed. The table would then become an inherent object of the SSIS system and be used to assist the decoder in making the most use of side information extracted from the stegoimage. Using this procedure for the Eiger80 image, the decoder BER was decreased from 0.0117 for hard-decision decoding and from 0.007 for hard-decision decoding with erasures to 0.

4.3.4.4 Comparison of Side Information Sources. To better compare these side information techniques, the Eiger cover image was used to generate stegoimages with a wide range of stegosignal powers, varying from 5 to 150 in steps of 5. For each of these 30 stegoimages, we decoded using each of the techniques described previously. A graph illustrating the relationship of the decoder output BER to the steganographic SNR (Stego-SNR) is shown in Figure 30. Low SNR reflects low stegosignal power. As an aside, note that the Stego-SNR for the Eiger80 image is –15.226.

From this graph, we can see that as the Stego-SNR is increased, the BER decreases for all decoder implementations. This result is intuitive, in that as the power of our signal becomes stronger, fewer bit errors should occur. Next we notice that when the Stego-SNR is very low, the BER value for all decoders approaches 0.5, indicating that the decoders are making decoding errors approximately one half of the time and are therefore providing no information about the embedded signal. As the stegosignal power is increased, the varying performance

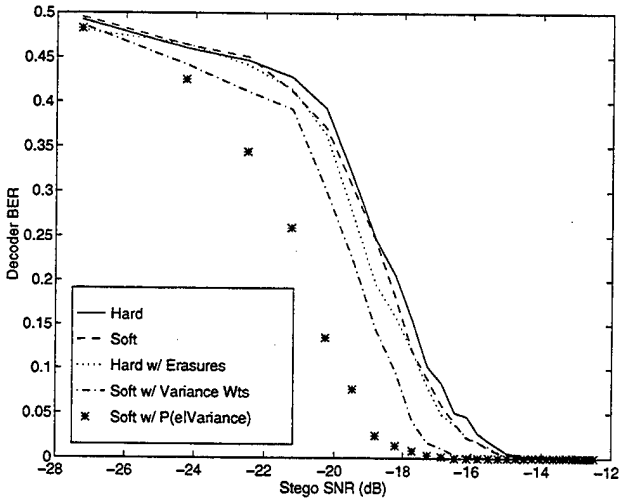


Figure 30. Comparison of Side Information Techniques.

of the different decoders is evident. The solid line represents the performance of the hard-decision Viterbi decoder without side information. The dashed line shows performance of the soft-decision decoding via the log-likelihood ratio and reflects an average increase in coding gain (over the hard-decision decoder) of 0.3 dB in Stego-SNR. The dotted line indicates the improvement obtained by using erasure side information to the Viterbi decoder with an average coding gain of 0.5 dB Stego-SNR. The dashed-dotted line illustrates the improvement of using the variance to weight the log-likelihood ratio, providing a coding gain of 1.3 dB Stego-SNR over hard-decision decoder. Finally, the best performance using generic tables of $P(error|variance)$ for soft-decision decoding is represented by the asterisks. An average coding gain of 2.5 dB in Stego-SNR is achieved using the $P(error|variance)$ as input to the soft-decision decoder.

4.3.5 Turbo Codes

We also investigated the performance of turbo codes, a recent development in error control [70, 71], in our system. Turbo codes, also known as parallel concatenated systematic convolutional codes that use two binary convolutional encoders and an interleaver, have been shown [70] to operate very close to Shannon's limit with reasonable decoding complexity.

Figure 31 shows a simple diagram of the turbo code encoder using rate 1/3 convolutional encoders. The fundamental idea is to encode a message bit using the first encoder to generate 3 bits, consisting of the message bit and 2 parity bits. The message bits are interleaved and used as input to the second encoder, producing an additional sets of 2 parity bits. Each code is systematic (the information bits are part of both encoded sequences); we need only to send one copy of the information. The received message and corresponding parity checks are then decoded via an iterative turbo decoder, as shown in Figure 32. The decoder has

a soft-input/soft-output decoder for each of the encoders along with a deinterleaver. These decoders take turns operating on the received data, forming and exchanging estimates of the message bit [72].

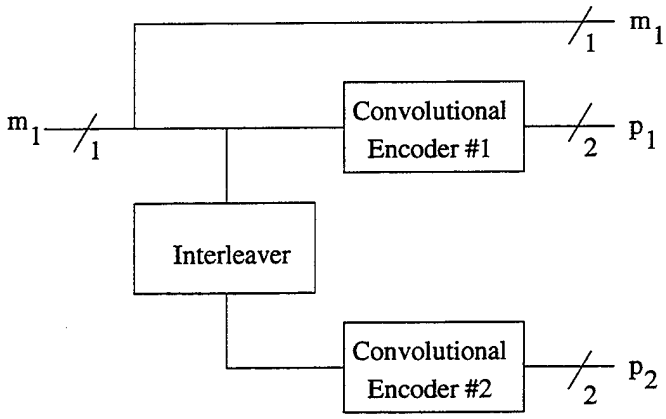


Figure 31. Turbo Coding Encoder.

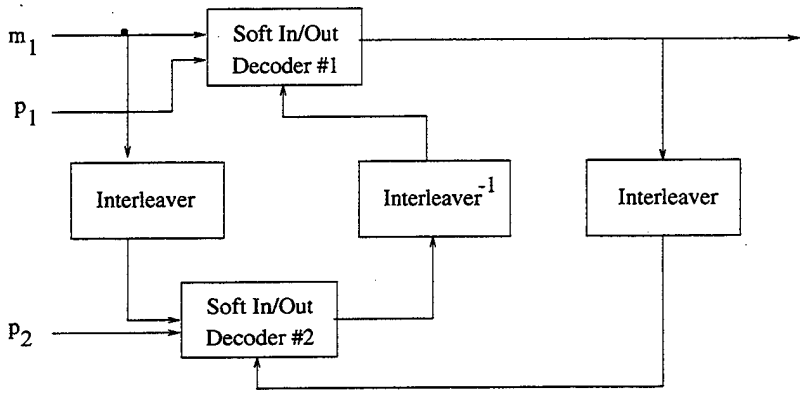


Figure 32. Turbo Coding Decoder.

We used a decoder developed by Bahl, Cocke, Jelinek, and Raviv (BCJR) [73] whose performance is only slightly better than the Viterbi decoder but is optimized to correct for bit errors rather than frame errors. Additionally, this decoder produces soft output, which is necessary for the turbo code implementations.

For our purposes, two rate 1/3 convolutional codes were selected along with a block interleaver to construct the turbo encoder, which has a composite rate of 1/5. The BCJR soft-input/soft-output decoders are used to implement an iterative turbo decoder. Table 3 shows the iterative correcting performance of the turbo decoder for the Eiger stegoimage along with the embedded signal BER for four values of stegosignal power. The payload for a 256×256 image using this code is 1638 bytes (an information rate of 0.025), which is greater than the 1364-byte payload obtained using the rate 1/6 convolutional code with an information rate of 0.0208.

For the Eiger80 image, the turbo decoder produced a message BER of 0.0105 while the rate 1/6 convolutional code using the hard-decision Viterbi decoder produced a BER of

Table 3. Iterative Turbo Decoder Performance.

Stegosignal Power	Stego SNR	Embedded Signal BER	Message BER for Number of Iterations			
			1	2	3	4
80	−15.2260	0.2345	0.0989	0.0343	0.0157	0.0105
90	−14.7145	0.2292	0.0704	0.0069	0.0028	0.0030
100	−14.2569	0.2229	0.0592	0.0047	0.0025	0.0025
110	−13.8430	0.2155	0.0389	0.0013	0.0000	0.0000

0.0117, both without the application of side information. Although the decoded message BER is comparable for these two coders, a higher payload is achieved with the turbo code whose information rate is 0.025 (40 image bits to 1 information bit) versus the information rate of 0.0208 (48 image bits to 1 information bit) for the hard-decision rate 1/6 convolutional code.

Naturally, it may be possible to improve these results by using different convolutional codes selected from a search of good convolutional codes for turbo coding [74] and implementing a more sophisticated interleaver. However, the performance of this encoder/decoder exhibits the basic attributes of turbo coding within SSIS.

4.3.6 Summary

In this subsection, we have described the role of error control within SSIS and discussed the use of several codes and decoders that have been incorporated into the system. The correlation between edges in the image and decoding errors along with its use as channel side information, which we have used to improve the decoder BER, has been cited. Several methods of incorporating side information has been specified and compared. In addition, the error correction performance of a simple turbo code has been illustrated. Of the methods presented here, the rate 1/6 convolutional code using the soft-decision Viterbi algorithm weighted by the variance of the stegoimage satisfies the requirement for a target decoded message BER of 0 and operates in a completely blind fashion.

For most cover images with an acceptable (in a steganographic sense) stegosignal power, the hidden message could not be recovered error free without the use of error correction. In fact, the ability to correct the recovered embedded signal permits the stegosignal to have the required low power.

4.4 System Performance

The purpose of this subsection is to demonstrate the performance of the SSIS algorithm and discuss issues that affect performance over noiseless and noisy transmission channels. In addition, we compare the SSIS payload to the theoretical capacity developed in section 3.

4.4.1 General Performance

Six images with a variety of image characteristics are used to demonstrate the performance of SSIS. The original digital 256×256 images, each containing 64 kilobytes, appear in Figure 33. The Barbara and Lena images are commonly used in the image processing community. These are subsampled from the original 512×512 images. The Lena image has some high-frequency regions around the feathers in her hat, along with some smooth regions on her face and shoulder. The Barbara image exhibits several high-frequency areas reflected in the regions with diagonal lines. The Castle, Eiger, and Ulm images* represent typical scenic images with a horizon, smooth sky, and detailed foreground. The Castle image is a picture of the Neuschwanstein castle in Bavaria; the Eiger image (used in the previous subsections) is the north face of the Eiger, taken from Kleine Scheidegg, Switzerland; and the Ulm image is a picturesque view of Ulm taken from the Danube in Germany. Our last test image is the Tank image, which contains a LAV-25 light armored vehicle followed by an M-551A1 Sheridan armored reconnaissance vehicle in Saudi Arabia. The image was taken in January 1991 by SGT Nathan Webster and obtained by the author from the U.S. Army web site. It has a significant amount of smooth area with hills of sand in the foreground and detailed items in both the mid- and background.

Thirty stegoimages were generated using the SSIS encoder for each of our six test images, with stegopower varying from 5 to 150, by steps of 5. The encoder used the piecewise linear modulation technique described in section 4.1.2 and the alpha-trimmed mean filter of section 4.2.2.2 for channel estimation. Figure 34 shows a graph displaying the relationship between the Stego-SNR and the embedded signal BER for each image. All graphs illustrate the general trend of decreasing BER for increasing Stego-SNR. Note that because the images have different characteristics, the specific values of embedded signal BER and corresponding Stego-SNR differ for each image.

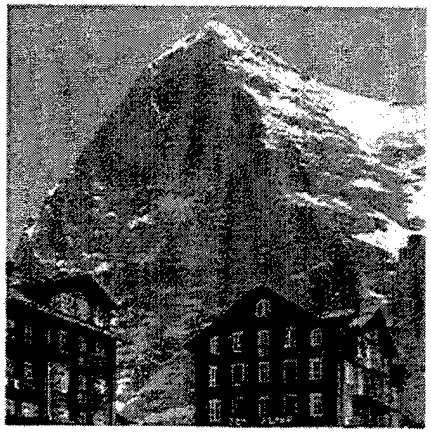
From this data, we can get a sense of the steganographic capabilities of each image. Basically, if the embedded signal BER is relatively low (steganographically speaking, less than 0.22 or so) and the corresponding Stego-SNR is low (less than, say, 20 dB), the image is a good candidate for steganography.

To present the effectiveness of the incorporation of side information in the error-control decoder, we have constructed plots for each test image displaying the decoder BER over the range of applicable Stego-SNR. These plots are shown in Figure 35 for the rate 1/6 code discussed in section 4.3.3. The code enables a payload of 1364 bytes for images of this size. As we mention previously, any error-control code that is capable of correcting the embedded signal BER can be used — if higher rate codes can be used, the payload amount will increase accordingly.

*The Castle, Eiger, and Ulm images are provided courtesy of Dr. Charles Retter, U.S. Army Research Laboratory.

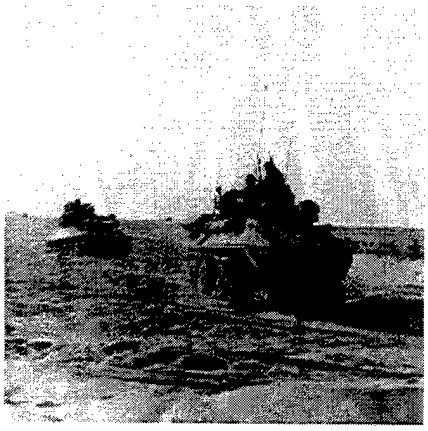
Barbara

Castle



Eiger

Lena



Tank

Ulm

Figure 33. Original Test Images.

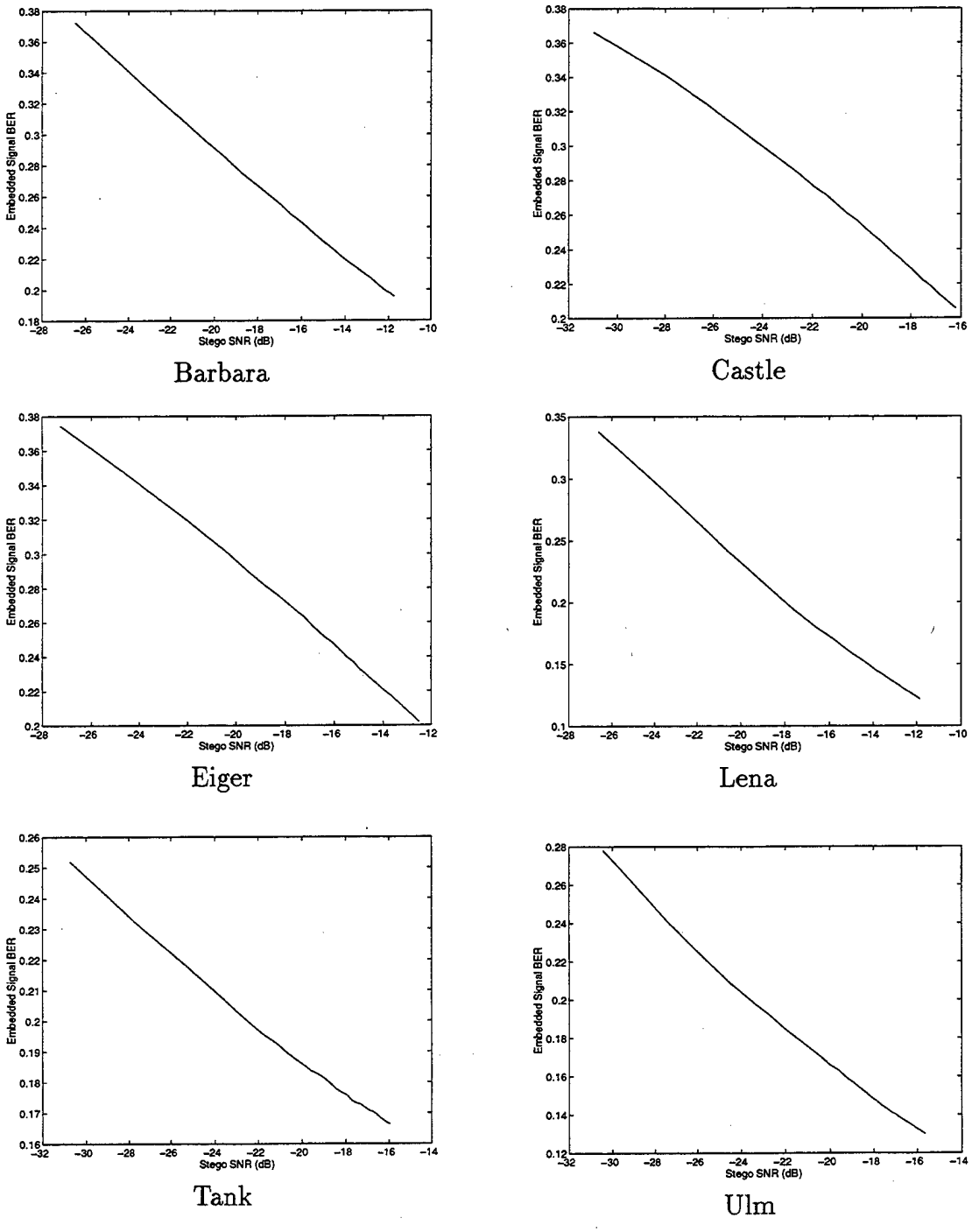
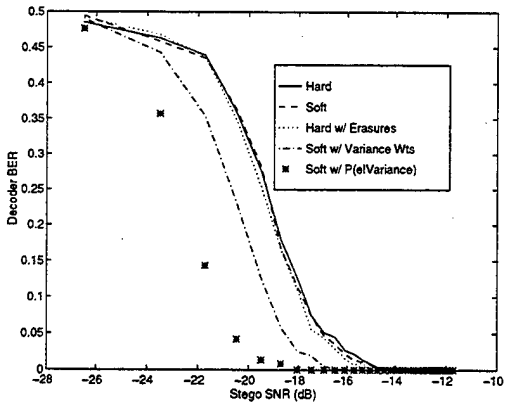
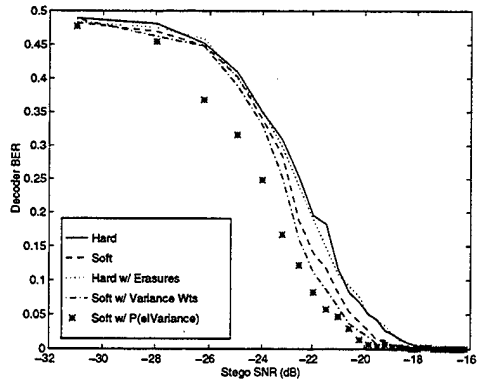


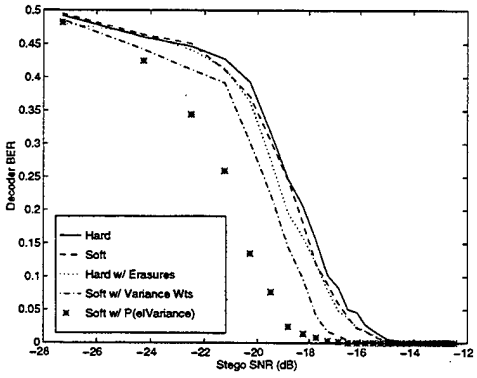
Figure 34. Embedded Signal BER.



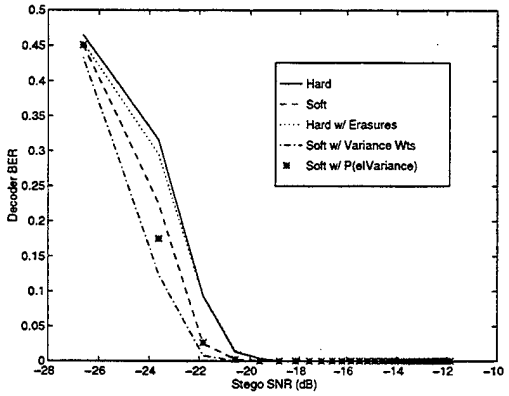
Barbara



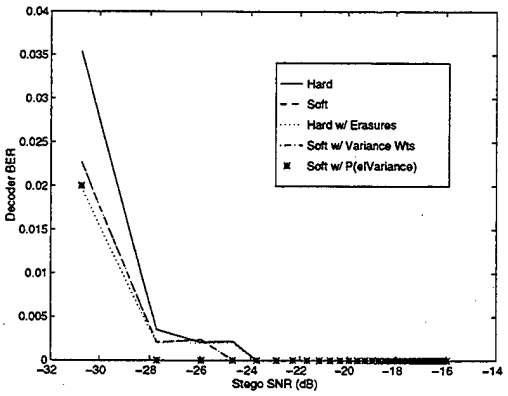
Castle



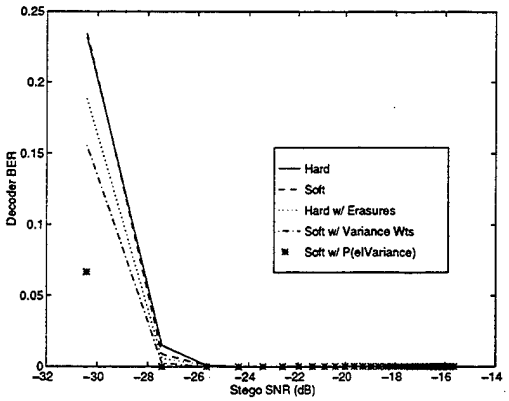
Eiger



Lena



Tank



Ulm

Figure 35. Decoder Performance.

4.4.2 Performance for a Fixed Payload

To compare the quality of the stegoimages and view the degradation caused by embedding a hidden message, it is necessary to look at images with a specific payload size. Fixing the payload is equivalent to specifying an embedded signal BER because the payload is determined by the ECC, which in turn has been selected to correct a specific level of BER.

Here, we compare stegoimages that have an embedded signal BER between the values of 0.21 and 0.22. For each test image, the stegosignal power has been adjusted to achieve an embedded signal BER in this range. The stegosignal power required to accomplish this goal varied with each image. For the Tank and Ulm images, a stegosignal power of 20 was needed to achieve embedded signal BER values of 0.2141 and 0.2079 with Stego-SNR values of –24.72 and –24.39, respectively. To obtain this level of embedded signal BER from the Barbara image, a stegopower value of 90, resulting in a Stego-SNR of –13.94, is required. Similar stegopower and Stego-SNR values are obtained for the Eiger and Castle images. Figure 36 displays the stegoimages along with the various stegopowers that are necessary to achieve an embedded signal BER between 0.22 and 0.21.

Since the BER is equivalent, all of these stegoimages require the same level of error correction and, therefore, carry an identical amount of payload. However, as one can see from the stegoimages, the image quality of the stegoimage compared to the original image varies. Since the stegosignal consists of AWGN, some of the test images look noisy depending on the image characteristics. For instance, the stegosignal is readily apparent in the sky of the Castle image because of its high power value, 120. However, even with a stegosignal power of 90, the presence of the stegosignal in the Barbara image is not obvious because of the amount of variation within the image. Yet the stegosignal, whose power is 110, is vaguely noticeable in the Eiger image. In the Lena and Tank image, the signal power is low and the stegoimages look fairly close to the original. Arguably, the best stegoimage is the Ulm stegoimage. There are just enough smooth areas in this image for the channel estimation to perform well, but also enough variation within that smooth area to help disguise the presence of the stegosignal. The MSE of the original image to the stegoimage is approximately equal to the noted stegosignal power.

4.4.3 Clipping

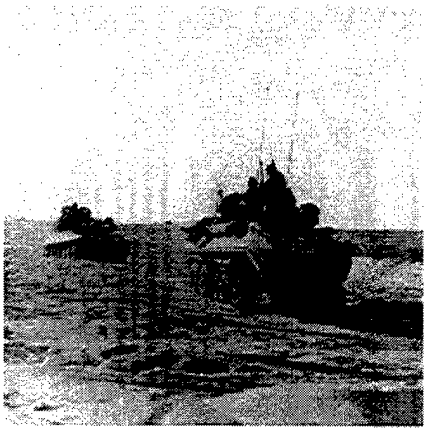
It is possible that many more decoding errors may occur for some cover images that have a large number of pixels lying at extreme colormap values, near 0 and 255 for grayscale, than for those which have a more centralized histogram. These errors occur when the addition of the stegosignal to the pixel value causes the stegoimage pixels to saturate the colormap. These values must then be truncated, or clipped, so that the stegoimage pixels are within the acceptable dynamic range of the colormap. This clipping may cause some or all of the embedded signal information to be lost. A grayscale stegoimage containing pixels with negative values or values greater than 255 would surely alert the suspicious observer, and thus defeat our goal of operating in a stealthy manner. Of course, the stegosignal power is also a factor to the degree at which clipping affects the embedded signal BER. When the

Barbara, Stegopower = 90

Castle, Stegopower = 120

Eiger, Stegopower = 110

Lena, Stegopower = 30



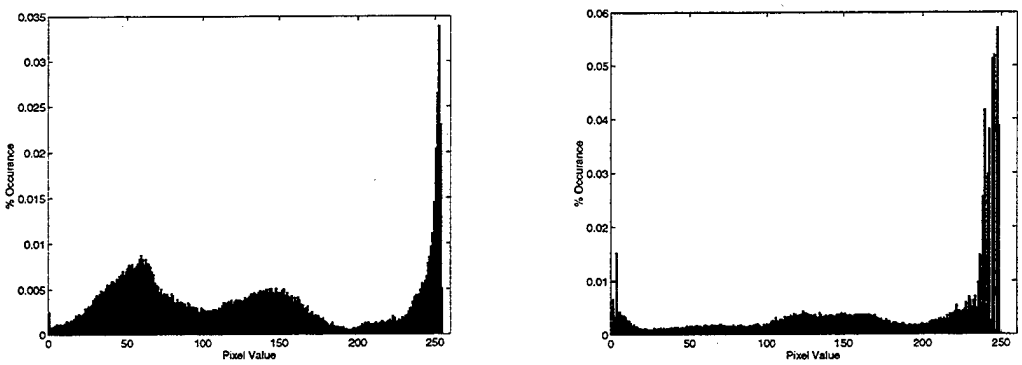
Tank, Stegopower = 20

Ulm, Stegopower = 20

Figure 36. Stegoimages with Embedded Signal BER $\approx$ 0.22.

stegosignal power is not great enough to cause saturation or the histogram is centralized, clipping may not be an issue.

Take, for instance, the Castle and Tank images, whose histograms are shown in Figure 37. Both images have similar histograms and comparably low values for a majority of the pixel range, but each has a substantial peak near colormap saturation, 255.



Castle Tank

Figure 37. Original Image Histograms.

The stegosignal power needed to attain an embedded signal BER of approximately 0.22, thereby achieving a constant payload, is 120 for the Castle image and 20 for the Tank image. The large stegosignal power for the Castle stegoimage causes 4,468 pixels, approximately 7.1%, to be saturated. The relatively low stegosignal power of 20 for the Tank stegoimage causes less than 1.9% of the pixel values to be saturated. Consequently, clipping is a more significant factor in the Castle stegoimage than the Tank stegoimage.

4.4.4 Variations in Transmission Channel

Throughout this report, results have been presented for the the noiseless transmission channel. Subsequently, the only noise incident upon the embedded signal is that caused by the cover image. This noiseless channel operation is reasonable when images are transmitted in a lossless fashion over today's reliable wired networks. Let us now consider the case where the embedded signal may be corrupted by other noise sources such as that of a wireless channel or the noise caused by lossy image compression.

4.4.4.1 Additive Channel Noise. Many SSIS stegoimages can be made resistant to low levels of additive noise by selecting the proper stegopower and coding. To demonstrate this resistance, we generated an Ulm stegoimage using a stegosignal power of 80 and added various levels of white Gaussian noise. This added noise is independent of the embedded signal and simulates a noisy transmission channel. As a baseline for performance comparison, the noiseless channel transmission of this stegoimage yields an embedded signal BER of 0.1515.

The Ulm stegoimage along with stegoimages that have been corrupted with AWGN with powers of 10, 20, and 30 are shown in Figure 38. All stegoimages have been decoded with the standard SSIS decoder using hard-decision decoding without side information. When adding transmission channel noise with a power of 10, the embedded signal BER is increased from 0.1515 to 0.1840. This increase is still within the acceptable BER range for the 1/6 rate convolutional code, which would correct this BER to 0. The channel noise power introduced in the next image has an increased power of 20, causing the embedded signal BER to increase to 0.2094 — still within the BER range for this code. However, when the additive noise power is increased to 30, the resulting embedded signal BER becomes 0.2312 and the decoded message BER increases from 0 to 0.003542. Consequently, a more powerful code, soft-decision decoding, a decoder incorporating side information, or an increase in stegosignal power must be used to achieve error-free message recovery at this level of transmission channel noise.

Ulm80 Stegoimage

Ulm80, AWGN Power = 10

Ulm80, AWGN Power = 20

Ulm80, AWGN Power = 30

Figure 38. Stegoimage Exposed to AWGN Channel.

4.4.4.2 Noise Caused by Lossy Image Compression. A stegoimage may also be made resistant to noise caused by low levels of lossy image compression. JPEG compression was used with various Q-factors. Q-factor is indicative of the amount of compression performed and the quality of the compressed image. For instance, JPEG compression with a high Q-factor would result in a good quality image that has little compression. As the Q-factor is reduced, the amount of compression increases while the quality of the compressed image decreases. Using the same Ulm80 stegoimage, we applied JPEG compression with a Q-factor of 95, resulting in a 4.43-bpp compressed image; the decompressed stegoimage is displayed in Figure 39. After decompression at the decoder, the embedded signal BER is 0.1671, an increase in BER from 0.1515 without compression noise, which is within the acceptable BER for the rate 1/6 convolutional ECC. By decreasing the Q-factor to 90, a 3.30-bpp compressed image is produced, and the embedded signal BER from the decompressed stegoimage increases to 0.2057 — still within the acceptable level for this ECC. When the Q-factor is again decreased to 85, yielding a 2.64-bpp compressed image, the resulting embedded signal BER is 0.2583 (above the BER capability of the rate 1/6 code). This is reflected in the 0.0686 message BER produced by the SSIS decoder. Finally, when the Q-factor is decreased once more to 80, providing a 2.21-bpp compressed image, the embedded signal BER increases to 0.3001 and is well beyond the capabilities of this ECC. In this case, the (2040,32) binary expansion of the RS code could be used since it can correct a BER of approximately 0.35 at the expense of payload. Alternatively, the use of other decoders with soft-decision decoding and/or side information or an increase in stegosignal power must be investigated to combat this level of noise.

Figure 39. Decompressed Stegoimage.

As an aside, it should be noted that the sender of the stegoimage selects the format in which the stegoimage is transferred to the recipient. Therefore, if a compressed image is necessary so as not to arouse suspicion from observers, the compression parameters could be chosen in such a way that acceptable SSIS performance is guaranteed.

4.4.5 Comparison with Capacity

In section 3, we introduce a generic theory of additive steganographic capacity. The steganographic capacity is a function of the power of the noise resulting from the channel, the entropy of the channel, and the power of the stegosignal. For image steganography, the entropy of the channel is estimated using the state-of-the-art lossless image compression technique, CALIC.

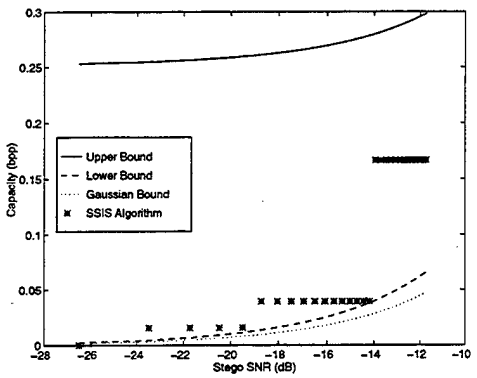
For each of our six test images, the steganographic capacity is compared to the performance of the SSIS algorithm using hard-decision decoding without side information. Figure 40 shows the relationship of the upper, lower, and Gaussian bounds on capacity to the performance of the SSIS system. Notice that the SSIS performance occurs in discrete steps. These discrete steps occur because of the fixed number of codes used to provide error correction. As a wider range of codes is added to the system, the discrete characteristic of the SSIS performance will diminish. For most of our images, the performance of the SSIS system falls between the upper and lower capacity bounds, thereby establishing a new lower capacity bound for these images. In the graph for the Eiger cover image, notice that the SSIS performance is close to the upper capacity bound for values of high Stego-SNR.

4.4.6 Summary

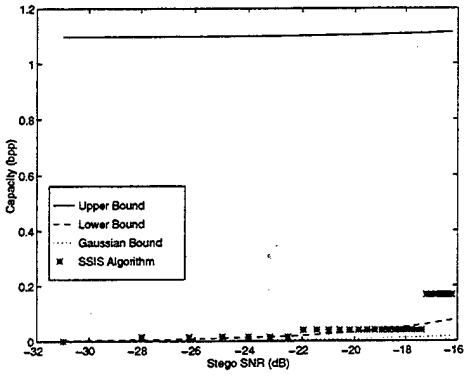
In this subsection, we have demonstrated the performance of the SSIS algorithm using six different test images. The relationship between BER and stegosignal power has been established along with the aspects of varying image quality for a fixed payload. Additionally, clipping and transmission channel noise have been addressed as issues that affect performance. Finally, bounds for image steganographic capacity were calculated and compared to SSIS performance. For all but one of our test images, SSIS performance fell within the upper and lower bounds of theoretical capacity estimates.

4.5 Summary of Section

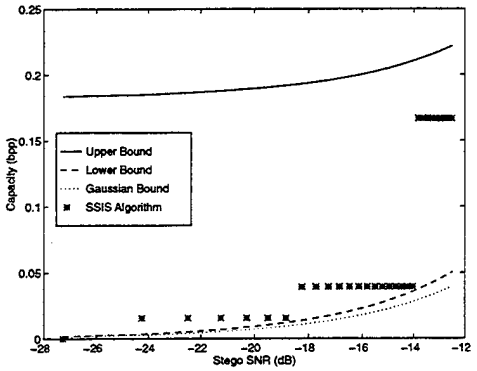
In this section, we have presented a novel technique to embed information within digital images that provides a high probability of error-free recovery. We have presented a piecewise linear modulation technique that provides increased detection performance of the hidden data at the receiver. Additionally, we have presented results that support our conjecture that the technique maximizes the minimum distance between modulation points, while adhering to the stochastic constraints of producing a stegosignal with a Gaussian distribution for message concealment. Channel estimation via image restoration has been used to facilitate blind embedded signal recovery, and two specific methods have been compared. Error control, which is vital to the functionality of the our system, has been discussed, and several types of codes have been presented. The use of channel side information to improve decoder performance has also been demonstrated. Finally, to provide an example of the tradeoffs among system parameters, the performance of the SSIS system has been demonstrated for several images.



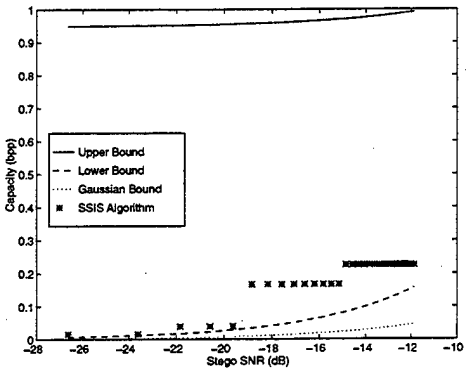
Barbara



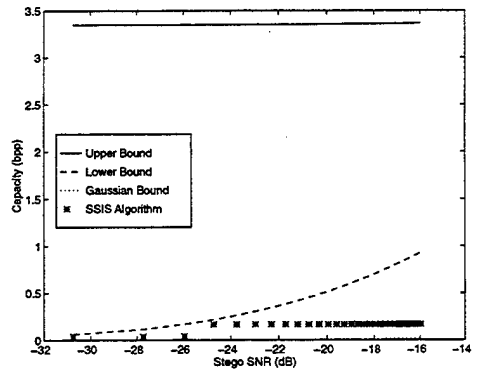
Castle



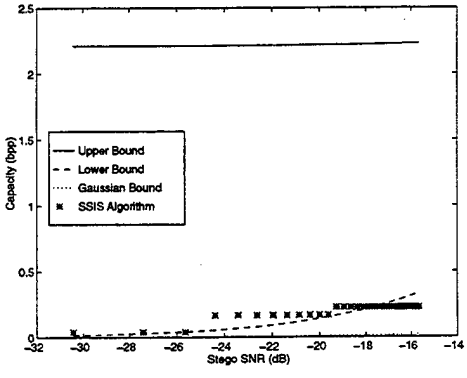
Eiger



Lena



Tank



Ulm

Figure 40. SSIS Performance Compared to Steganographic Capacity Bounds.

5. Conclusions

Steganography's primary objectives — imperceptibility, removal resistance, and capacity — are in conflict. It is not possible to achieve optimal levels of all three simultaneously. Consequently, tradeoffs between them must be dictated by the application. For data hiding, the effort is concentrated on achieving high levels of imperceptibility and capacity while sacrificing resistance.

Since much of the current communication occurs digitally, new data-hiding methods for digital cover media are being investigated. Applications for data hiding are numerous. They include in-band captioning, forward/backward compatibility, authentication, and, of course, hidden communication. For many of the new techniques, quantitative measurements of success do not exist, particularly in terms of capacity and imperceptibility. We have developed such a metric for a class of data-hiding methods, along with a novel data-hiding technique that can provide perfect recovery of the hidden data.

5.1 Contributions

In this report, we have approached the problem of steganographic communication as we would a typical communication problem comprising a transmitter, a receiver, and a channel. Modulation techniques are needed to convey the information; detection techniques are required at the receiver for signal recovery; and, if the channel is noisy and the signal recovered in error, error control is essential. Finally, metrics are necessary to quantify performance.

To this end, we have derived, from Shannon's information theory, a measure of additive steganographic channel capacity. The capacity can be used to gauge the performance of the class of steganographic techniques that embed information by adding it to the cover signal. Previous to this work, the capacity metric generally accepted by the steganographic community was that of the Gaussian channel, a measure that significantly understates a steganographic channel's true capacity. We have also presented a process to estimate the capacity of an image steganography system by computing capacity bounds using an estimate of the image entropy.

We have developed a complete image steganographic strategy for the purpose of data hiding. The primary objective of this system was to embed as much data into an image as possible and provide its reliable recovery without requiring the receiver to have the cover information. This method combined spread spectrum techniques, channel estimation, and error control to communicate hidden information using an image while concealing its existence from the human visual system and computer analysis.

Steganographic principles dictate that the information signal must have low power in comparison to the communication channel to promote stealthy communication. Also, without perfect knowledge of the channel, the detection of the low power signal carrying the hidden data is subject to distortion. A new modulation technique for obscure communication, called piecewise linear modulation, was created to provide a low power signal that is perceived as a type of noise that commonly exists in digital images. The use of this modu-

lation technique provides a minimum distance that varies proportionally to the stegosignal power, which can be adjusted to achieve better detection without sacrificing payload.

We have investigated several image processing techniques to better estimate our channel and have selected one that promotes reduction of the recovered embedded signal error. To combat the residual error in the hidden data, we have employed error-control codes to correct them. Procedures have been sought to improve error-control performance because its incorporation reduced the payload. In doing so, we have discovered that channel side information could be extracted from the stegoimage and used to assist the decoder with correcting errors. The performance of the SSIS system has been exhibited for six test images and compared to the theoretical channel capacity bounds.

Although piecewise linear modulation was developed as a component of our steganography system, we have found that it could function independently as a general purpose method of hidden communication. The goals of piecewise linear modulation were to modulate a binary signal to produce an output that mimics low-power, thermal, Gaussian noise. This output could then be transmitted to a recipient where the binary signal could be recovered. Since the output signal may incur distortion in transit, modulation values must have a large minimum distance while adhering to the Gaussian stochastic constraint to achieve accurate detection. We have used basic probability theory to reduce this formidable problem to one that could be easily perceived by exploiting the relationship between the uniform and Gaussian distributions. In section 4.1.2 of section 4, it has been shown that the piecewise linear modulation was optimal in maximizing the minimum distance in the uniform domain. Empirical data to support our conjecture that this relationship held in the Gaussian domain have also been presented.

5.2 Future Work

All of the concepts presented in this report could, with relative ease, be extended to construct steganographic systems for a variety of cover signals that possess high correlation among neighboring samples. Audio, video, and color imagery signals are a few of the numerous possibilities that come immediately to mind. The SSIS system, as presented here, is ripe for implementation and could function as a general purpose data-hiding technique to embed information within digital grayscale imagery.

The comparison of our system performance with our established theoretical steganographic bounds demonstrates that there may be more throughput to be obtained for some cover images. Modulation, channel estimation, and coding theory should be further exploited so users of steganography may gain from this potential increase in payload. Additionally, it would be beneficial to improve the estimate of image entropy to more accurately measure capacity.

Of course, it is important to consider the opposing side to steganography — the detection, attack, and possible decoding of the hidden information. This new and challenging area of research is referred to by the "stegocommunity" as *stegoanalysis*. The relationship between steganographer and stegoanalyzer is reminiscent of the that between the cryptographer and

the code breaker. The steganographer strives to perfect methods to conceal the existence of information, and the stegoanalyzer endeavors to destroy them or, at the least, render them ineffective. Each side is encouraged to further the technology by the pursuit of the other.

6. References

[1] Electronic Messaging Association.
Numbers.
Time, page 23, January 1999.

[2] B. W. Kroeger and P. J. Peyla.
Robust in-band on-channel digital audio broadcasting AM and FM technology for digital audio broadcasting.
A white paper produced by USA Digital Radio, 1998.

[3] D. Kahn.
The Codebreakers - The story of secret writing.
Scribner, New York, NY, 1967.

[4] B. Pfitzmann.
Trials of traced traitors.
In R. Anderson, editor, *Information Hiding, First International Workshop*, volume 1174 of *Lecture Notes in Computer Science*, pages 49–64. Springer-Verlag, Berlin, 1996.

[5] J. R. Smith and B. O. Comisky.
Modulation and information hiding in images.
In R. Anderson, editor, *Information Hiding, First International Workshop*, volume 1174 of *Lecture Notes in Computer Science*, pages 207–226. Springer-Verlag, Berlin, 1996.

[6] R. Van Schyndel, A. Tirkel, and C. Osborne.
A digital watermark.
In *Proceedings of the IEEE International Conference on Image Processing*, volume 2, pages 86–90, 1994.

[7] R. B. Wolfgang and E. J. Delp.
A watermark for digital images.
In *Proceedings of the IEEE International Conference on Image Processing*, volume III, pages 219–222, Lausanne, Switzerland, September 1996.

[8] R. Machado.
Stego, http://www.fqa.com/romana/romanasoft/stego.html, 1997.

[9] E. Milbrandt.
Steganography info and archive. http://members/iquest.net/~mrmil/stego.html, October 1997.

[10] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon.
Secure spread spectrum watermarking for images, audio and video.
In *Proceedings of the IEEE International Conference on Image Processing*, volume III, pages 243–246, Lausanne, Switzerland, September 1996.

[11] C. I. Podilchuk and W. Zeng.
Digital image watermarking using visual models.
In B.E. Rogowitz and T.N. Pappas, editors, *Human Vision and Electronic Imaging II*, volume 3016, pages 100–111. SPIE, February 1997.

[12] M. D. Swanson, B. Zhu, and A. H. Tewfik.
Transparent robust image watermarking.
In *Proceedings of the IEEE International Conference on Image Processing*, volume III, pages 211–214, Lausanne, Switzerland, September 1996.

[13] W. Bender, D. Gruhl, N. Morimoto, and A. Lu.
Techniques for data hiding.
IBM Systems Journal, 35(3 & 4), 1996.

[14] K. Tanaka, Y. Nakamura, and K. Matsui.
Embedding secret information into a dithered multi-level image.
In *Proceedings of the IEEE Military Communications Conference*, pages 216–220, Monterey, CA, 1990.

[15] P. Davern and M. Scott.
Fractal based image steganography.
In R. Anderson, editor, *Information Hiding, First International Workshop*, volume 1174 of *Lecture Notes in Computer Science*, pages 279–294. Springer-Verlag, Berlin, 1996.

[16] M. D. Swanson, B. Zhu, and A. H. Tewfik.
Robust data hiding for images.
In *Proceedings of the IEEE Digital Signal Processing Workshop*, pages 37–40, Loen, Norway, September 1996.

[17] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon.
Secure spread spectrum watermarking for multimedia.
Technical Report 95-128, NEC Research Institute, August 1995.

[18] Chr. Neubaur, J. Herre, and K. Brandenburg.
Continuous steganographic data transmission using uncompressed audio.
In D. Aucsmith, editor, *Information Hiding, Second International Workshop*, volume 1525 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1998.

[19] A. Westfeld and G. Wolf.
Steganography in video conferencing systems.
In D. Aucsmith, editor, *Information Hiding, Second International Workshop*, volume 1525 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1998.

[20] CCITT(ITU) recomendation H.261, video codes for audiovisual services at p $\times$ 64 kbit/s, 1993.
ITU-3/93.

[21] D. Mukherjee, J. J. Chae, and S. K. Mitra.
A source and channel coding approach to data hiding with application in hiding speech in video.
In *Proceedings of the IEEE International Conference on Image Processing*, Chicago, IL, October 1998.
MA09.08.

[22] G. Strang and T. Nguyen.
Wavelets and Filter Banks.
Wellesley-Cambridge Press, Wellesley, MA, 1996.

[23] CCITT recomendation H.263, video coding for low bit rate communication, 1998. ITU-2/98.

[24] S. D. Servetto, C. I. Podilchuk, and K. Ramchandran.
Capacity issues in digital image watermarking.
In *Proceedings of the IEEE International Conference on Image Processing*, Chicago, IL, October 1998.
MA.11.05.

[25] L. Xie and G. Arce.
A watermark for digital images.
In *Proceedings of the IEEE International Conference on Image Processing*, Chicago, IL, October 1998.
TA10.09.

[26] C. E. Shannon.
Communication in the presence of noise.
Proceedings Institute of Radio Engineers, 37:10–21, 1949.

[27] C. E. Shannon.
A mathematical theory of communication.
Bell Systems Technical Journal, 27:379–423 and 623–656, July and October 1948.

[28] R. E. Blahut.
Principles and Practices of Information Theory.
Addison-Wesley Publishing Co., Reading, MA, 1987.

[29] J. Fan.
Some results on the capacity of non-Gaussian channels.
In *Proceedings of the IEEE International Symposium on Information Theory*, page 152, Kobe, Japan, June 1988.

[30] F. Hartung and B. Girod.
Fast public-key watermarking of compressed video.
In *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, CA, October 1997.

[31] N. F. Johnson and S. Jajodia.
Exploring steganography: Seeing the unseen.
IEEE Computer, pages 26–34, February 1998.

[32] L. M. Marvel, C. G. Boncelet, Jr., and C. T. Retter.
Spread spectrum image steganography.
IEEE Transactions on Image Processing, 8(8):1075–1083, August 1999.

[33] L. M. Marvel, C. G. Boncelet, Jr., and C. T. Retter.
Reliable blind information hiding for images.
In D. Aucsmith, editor, *Information Hiding, Second International Workshop*, volume 1525 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1998.

[34] L. M. Marvel, C. G. Boncelet, Jr., and C. T. Retter.
Hiding information in images.

In *Proceedings of the IEEE International Conference on Image Processing*, Chicago, IL, October 1998.

[35] L. M. Marvel and C. G. Boncelet, Jr.
Capacity of the steganographic channel.
(Submitted to the IEEE Transactions on Signal Processing).

[36] L. M. Marvel, C. G. Boncelet, Jr., and C. T. Retter.
A methodology for data hiding using images.
In *Proceedings of the IEEE Conference on Military Communication (MILCOM'98)*, Boston, MA, October 1998.

[37] J. W. Woods.
Two-dimensional discrete markovian fields.
IEEE Transactions on Information Theory, pages 232–240, March 1972.

[38] A. K. Jain.
Advances in mathematical models for image processing.
Proceedings of the IEEE, pages 502–528, May 1981.

[39] R. L. Kshyap.
Characterization and estimation of two-dimensional ARMA models.
IEEE Transactions on Information Theory, pages 736–745, September 1984.

[40] R. M. Haralick and L. Watson.
A facet model for image data.
Computer Graphics and Image Processing, pages 113–129, February 1981.

[41] R. W. Hamming.
Coding and Information Theory.
Prentice-Hall, Inc., Englewood Cliffs, NJ, 1980.

[42] X. Wu, N. Memom, and K. Sayood.
A contex-based, adaptive, lossless/nearly-lossless coding scheme for continuous-tone images (CALIC).
A proposal submitted in response to the Call for Contributions for ISO/IEC JTC 1.29.12, 1995.

[43] S. Lin and D. J. Costello, Jr.
Error Control Coding: Fundaments and Applications.
Prentice-Hall, Inc., 1983.

[44] A. K. Jain.
Fundamentals of Digital Image Processing.
Prentice-Hall, Inc., Englewood Cliffs, NJ, 1989.

[45] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt.
Spread Sprectrum Communications, Volume I.
Computer Science Press, Rockville, MD, 1985.

[46] B. Schneier.
Applied Cryptography - Protocols, Algorithms, and Source Code in C, 2nd Edition.
John Wiley and Sons, Inc., New York, NY, 1996.

[47] D. E. Knuth.
The Art of Computer Programmering: Volume 2, Seminumerical Algorithms, Third Edition.
Addison-Wesley, Reading, MA, 1998.

[48] Probability theory: Transforming variables.
In Daniel Zwillinger, editor, *Standard Mathematical Tables and Formulae, 30th Ed.*, page 577. CRC Press, Inc., Boca Raton, F.L., 1996.

[49] P. G. Hoel.
Introduction to Mathematical Statistics, 2nd Ed.
John Wiley and Sons, Inc., New York, NY, 1960.

[50] R. E. Walpole, R. H. Myers, and S. L. Myers.
Probability and statistics for engineers and scientists, 6th Ed.
Prentice-Hall, Upper Saddle River, NJ, 1998.

[51] S. Ross.
A First Course in Probability, 3rd Edition.
Macmillan Publishing Co., New York, NY, 1988.

[52] J. M. H. Olmsted.
Advanced Calculus.
Appleton-Century-Crofts, Inc., New York, NY, 1961.

[53] S. S. Wilks.
Mathematical Statistics.
John Wiley and Sons, Inc., New York, NY, 1962.

[54] D. L. Donoho.
De-noising by soft-thresholding.
IEEE Transactions on Information Theory, 41(3):613–627, 1995.

[55] J. S. Lim.
Two-Dimensional Signal and Image Processing.
Prentice-Hall, Inc., Englewood Cliffs, NJ, 1990.

[56] J. S. Lee.
Digital image enhancement and noise filtering by use of local statistics.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 2:165–168, March 1980.

[57] J. B. Bednar and T. L. Watt.
Alpha-trimmed means and their relationship to the median filters.
IEEE Transactions on Acoustics, Speech, and Signal Processing, 32(1):145–153, February 1984.

[58] H. J. Trussell and B. R. Hunt.
Sectioned methods in image processing.
IEEE Transactions on Acoustics, Speech and Signal Processing, 26:157–164, April 1978.

[59] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel.
Adaptive noise smoothing filters for images with signal-dependent noise.

IEEE Transactions on Pattern Analysis and Machine Intelligence, 7:165–177, March 1985.

[60] C. T. Retter.
Binary weight distributions of low rate Reed-Solomon codes.
Technical Report ARL-TR-915, U.S. Army Research Laboratory, Aberdeen Proving Ground, MD, December 1995.

[61] C. T. Retter.
Decoding binary expansions of low-rate Reed-Solomon codes far beyond the BCH bound.
In *Proceedings of the 1995 IEEE International Symposium on Information Theory*, page 276, Whistler, British Columbia, September 1995.

[62] M. Bossert and F. Hergert.
Hard- and soft-decision decoding beyond the half minimum distance - an algorithm for linear codes.
IEEE Transactions on Information Theory, 32(5):709–714, Sep 1986.

[63] L. J. Harcke and G. E. Wood.
Laboratory and flight performance of the Mars Pathfinder (15,1/6) convolutionally encoded telemetry link.
NASA Code 624-04-00-MN-20 NASA/JPL TDA Progress Report 42-129, National Aeronautics and Space Administration, May 1997.

[64] A. J. Viterbi.
Error bounds for convolutional codes and an asymptotically optimum decoding algorithm.
IEEE Transactions on Information Theory, IT-13(2):260–269, April 1967.

[65] R. E. Blahut.
Theory and Practice of Error Control Codes.
Addison-Wesley Publishing Co., Reading, MA, 1983.

[66] G. D. Forney, Jr.
The Viterbi algorithm.
Proceedings of the IEEE, 61:268–278, March 1973.

[67] H. L. Van Trees.
Detection, Estimation and Modulation Theory.
John Wiley and Sons, Inc., New York, NY, 1968.

[68] R. C. Gonzalez and R. E. Woods.
Digital Image Processing.
Addison-Wesley Publishing Co., Reading, MA, 1993.

[69] J. Pearl.
Probilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.
Morgan Kaufmann Publishers, Inc., Palo Alto, CA, 1988.

[70] C. Berrou, A. Glavirux, and P. Thitimajshima.
Near Shannon limit error-correcting coding and decoding: Turbo-codes.

In *Proceedings of the 1993 International Conference on Comunications (ICC'93)*, pages 1064–1070, 1993.

[71] C. Berrou and A. Glavirux.
Near optimum error correcting coding and decoding: Turbo-codes.
IEEE Transactions on Communications, 44(10):1261–1271, Oct 1996.

[72] C. Heegard and S. B. Wicker.
Turbo Coding.
Kluwer Academic Publishers, Boston, MA, 1999.

[73] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv.
Optimal decoding of linear codes for minimizing system error rate.
IEEE Transactions on Information Theory, pages 284–287, March 1974.

[74] S. Benedetto, R. Garello, and G. Montorsi.
A search for good convolutional codes to be used in the construction of turbo codes.
IEEE Transactions on Communications, 46(9):1101–1105, September 1998.

INTENTIONALLY LEFT BLANK.

NO. OF COPIES	ORGANIZATION
2	DEFENSE TECHNICAL INFORMATION CENTER DTIC DDA 8725 JOHN J KINGMAN RD STE 0944 FT BELVOIR VA 22060-6218
1	HQDA DAMO FDQ D SCHMIDT 400 ARMY PENTAGON WASHINGTON DC 20310-0460
1	OSD OUSD(A&T)/ODDDR&E(R) R J TREW THE PENTAGON WASHINGTON DC 20301-7100
1	DPTY CG FOR RDA US ARMY MATERIEL CMD AMCRDA 5001 EISENHOWER AVE ALEXANDRIA VA 22333-0001
1	INST FOR ADVNCD TCHNLGY THE UNIV OF TEXAS AT AUSTIN PO BOX 202797 AUSTIN TX 78720-2797
1	DARPA B KASPAR 3701 N FAIRFAX DR ARLINGTON VA 22203-1714
1	NAVAL SURFACE WARFARE CTR CODE B07 J PENNELLA 17320 DAHLGREN RD BLDG 1470 RM 1101 DAHLGREN VA 22448-5100
1	US MILITARY ACADEMY MATH SCI CTR OF EXCELLENCE DEPT OF MATHEMATICAL SCI MADN MATH THAYER HALL WEST POINT NY 10996-1786

NO. OF COPIES	ORGANIZATION
1	DIRECTOR US ARMY RESEARCH LAB AMSRL DD 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	DIRECTOR US ARMY RESEARCH LAB AMSRL CS AS (RECORDS MGMT) 2800 POWDER MILL RD ADELPHI MD 20783-1145
3	DIRECTOR US ARMY RESEARCH LAB AMSRL CI LL 2800 POWDER MILL RD ADELPHI MD 20783-1145
	ABERDEEN PROVING GROUND
4	DIR USARL AMSRL CI LP (BLDG 305)

NO. OF COPIES	ORGANIZATION
1	COMMANDER US ARMY CECOM AMSEL RD ST SP P VAN SYCKLE FT MONMOUTH NJ 07703
1	DIRECTOR US ARMY RESEARCH OFFICE AMXRO EL W SANDER PO BOX 12211 RESEARCH TRIANGLE PARK NC 22709-2211
1	COMMANDANT US MILITARY ACADEMY WEST POINT NY 10996
1	COMMANDANT US NAVAL ACADEMY ANNAPOLIS MD 21404
1	COMMANDANT US AIR FORCE ACADEMY COLORADO SPRINGS CO 80840
1	DIRECTOR NAVAL RESEARCH LABORATORY WASHINGTON DC 20375-5000
	ABERDEEN PROVING GROUND
2	DIR USARL AMSRL IS DR GANTT AMSRL IS TP DR GOWENS

NO. OF COPIES	ORGANIZATION

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 2000	Final, Jan 97 - May 99

4. TITLE AND SUBTITLE
Image Steganography for Hidden Communication

5. FUNDING NUMBERS
611104.H509FEA00

6. AUTHOR(S)
Lisa M. Marvel

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
ATTN: AMSRL-IS-CI
Aberdeen Proving Ground, MD 21005-5067

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-2200

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

10.SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Modern steganographic methods, which conceal the existence of communication, are needed to exploit contemporary modes of information exchange. Measures of performance for these methods are essential to compare specific algorithms and determine appropriate uses. This report develops a methodology for steganographic data hiding. The methodology encompasses derivation of a general theory of steganographic communication, including theoretical capacity bounds, and design of an actual data-hiding technique that used digital imagery as a cover. The technique promotes maximization of payload, allows error-free recovery of embedded date, and provides some resilience to removal while concealing the existence of the embedded information from the observer and the observer's resources (e.g., computer).

14. SUBJECT TERMS	15. NUMBER OF PAGES
steganography, capacity, hidden communication	85
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

INTENTIONALLY LEFT BLANK.

USER EVALUATION SHEET/CHANGE OF ADDRESS

This Laboratory undertakes a continuing effort to improve the quality of the reports it publishes. Your comments/answers to the items/questions below will aid us in our efforts.

1. ARL Report Number/Author ARL-TR-2200 (Marvel) Date of Report April 2000

2. Date Report Received _____

3. Does this report satisfy a need? (Comment on purpose, related project, or other area of interest for which the report will be used.) _____

4. Specifically, how is the report being used? (Information source, design data, procedure, source of ideas, etc.) _____

5. Has the information in this report led to any quantitative savings as far as man-hours or dollars saved, operating costs avoided, or efficiencies achieved, etc? If so, please elaborate. _____

6. General Comments. What do you think should be changed to improve future reports? (Indicate changes to organization, technical content, format, etc.) _____

CURRENT ADDRESS

_____
Organization

_____
Name E-mail Name

_____
Street or P.O. Box No.

_____
City, State, Zip Code

7. If indicating a Change of Address or Address Correction, please provide the Current or Correct address above and the Old or Incorrect address below.

OLD ADDRESS

_____
Organization

_____
Name

_____
Street or P.O. Box No.

_____
City, State, Zip Code

(Remove this sheet, fold as indicated, tape closed, and mail.)
(DO NOT STAPLE)