Extraction Methods of Watermarks from Linearly-Distorted Images to Maximize Signal-to-Noise Ratio

By

Brandon Migdal

Advisors:

Carl Salvaggio

Chris Honsinger

A senior project submitted in partial fulfillment of
the requirements for the degree of Bachelor of Science
in the Chester F. Carlson Center for Imaging Science
of the College of Science at the Rochester Institute of Technology

April 29th 2004

1.0 Abstract

This project succeeded in evaluating two methods for the extraction of watermarks from linearly distorted images. The evaluation was in terms of the signal-to-noise ratio of the extracted message. The evaluated methods were those of windowing and subsetting. Two windowing methods were evaluated, Hanning and Hamming windows. Two subsetting cases were also evaluated they were 4 subset and 16 subset cases. The project shows that as the linear distortion increases significant gains in the signal-to-noise ratio are made with the windowed extractions. This project shows that the process of subsetting produces only marginal increases in the signal-to-noise ratio.

2.0 Background

2.1 Algorithm Background

The Eastman Kodak watermarking algorithm is a method that uses phase dispersion to accomplish the goal of data embedding. The data embedding process is performed and a message containing up to 64-bits can be embedded into the image.(Rabbani, 2000).

The first item that is needed for the embedding process is a random phase carrier. This random phase carrier is a 128 x 128 pixel array consisting of random Fourier phase values and uniform magnitude in the frequency domain. The transformation of this information back into the spatial domain yields an image with floating point values.

Despite the stated randomness of the carrier there are certain considerations that must be made that directly affects the robustness, visual transparency, and the quality of extracted information of the watermark. To increase the visual transparency of the watermark, most of the information in the carrier should be focused at the higher frequencies. Since the human visual system is not as sensitive to these higher frequencies, the watermark will be less visible. To increase the quality of the extracted information, measures must be taken to insure that the carrier will be resistant to attacks. Putting the bulk of the information at high frequency is problematic since the watermark becomes vulnerable to attacks such as low-pass filtering which, if not accounted for, could destroy the watermark. All of these considerations were taken into account by Eastman Kodak when developing the carrier.

The phase for the random carrier is created using a pseudo-random number generator with a specified key. The key to the random number generator is specified by the user. The amplitude of the coefficient of frequency is set to zero, and then the amplitude of the coefficient of frequency increased. Scaling the frequency coefficients of the carrier makes the carrier more robust to attacks, the scaling of frequency coefficients then also serves to increase the quality of extracted watermarks. The increase of the amplitude of the coefficient of frequency is continued until the value reaches 20% of Nyquist (Nyquist is defined to be the highest frequency at which a signal can contain energy and remain compatible with the Sampling Theorem.). When the frequencies are greater than 20% of Nyquist those frequencies are generated from the Contrast Sensitivity Functions (CSF). The CSF is a set of functions based on observations made by observers under various conditions that relate how they perceive contrast at a given spatial frequency. The principle of CSF works by allowing a set of viewing conditions to be

chosen and then applying statistics to determine the observer's given likeliness to see a frequency. By doing this it allows the message artifacts to be kept just below the perceptible limits of most observers. By properly designing this carrier it allows for the information that is embedded to achieve the goals that were laid out above, to increase robustness, increase visual transparency, and increase the quality of extracted data. Figure 1 shows an example of the carrier in the spatial domain. The image in the frequency domain appears similar to this representation, however the values in the frequency domain span a scale from -1 to +1. As was described above the carrier contains random values. The image of the carrier is a floating point image with digital counts ranging from -400 to +400 (carrier values are directly affected by the key to the random phase generation and as such will change with the key) and the image has merely been scaled for display.

**Figure 1:** Random Phase Carrier (Scaled for Display)
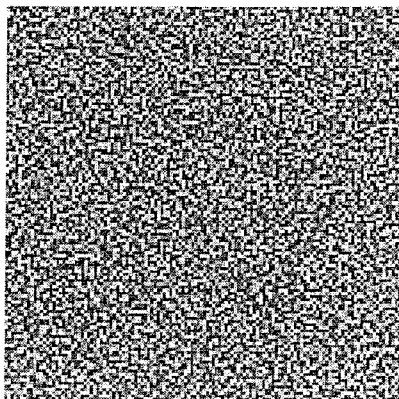
The next component to the embedding process is the message template. The message template is a 128 x 128 pixel array that contains 64 locations for embedding

binary information. This means that there are 64 places in the array to put a 1 or 0 which will be the message that is later extracted. There are a large number of designs for this message template. Basically any 128 x 128 array consisting of the digital count 0 in every pixel, except for 64 unique locations that are used for embedding. However, there are considerations that must be taken into account when designing the message template.

The first consideration in making a message template is that it should maximize the quality of the watermark extracted. To achieve this maximization it is necessary to place the 64 unique locations as far apart from each other as possible. What is gained by increasing this distance is prevention of interference between points during the extraction process. The second consideration that is taken into effect when designing a message template is that the watermark should be robust against attack. The two most common attacks are cropping and rotation, so it would be ideal to have a message template that can be removed from an image after that image has been cropped, rotated, or even both. By selecting a message template with unique and well placed embedding locations these attacks can be negated by cyclically shifting the message template during the extraction process. The data locations are important in the shifting because one wants to make certain that they are correlating the correct point. Figure 2 shows three possible designs for message templates. The template in Figure 2(a) would not be optimal since the data locations are too close together and the information might overlap upon extraction. The point overlap in extraction would be a result of ringing that occurs in the frequency domain when performing the cyclic correlation. Though the template in Figure 2(b) maximizes the distance between points it is not optimal. Since when using this template on a shifted or cropped image it becomes impossible to be certain that the correct point is

being extracted (due to overlap). The aforementioned false positives results in extraction that occur when using the template from Figure 2(b) are the result of the cyclic correlation. (See Figure 4(a), which shows how the cyclic nature of the shifts would cause false positives.) The template in Figure 2(c) is the optimal template as it has an increased distance between data locations, and insures that a cyclically shifted version of it will not overlap the wrong points for extraction. (See Figure 4(b) for an example of the cyclic shift.) The message template 2(c) was chosen to be the used template even though it contains 74 bits instead of 64.

Figure 2: Three Possible Message Templates for Data Embedding

(a)   (b)   (c)
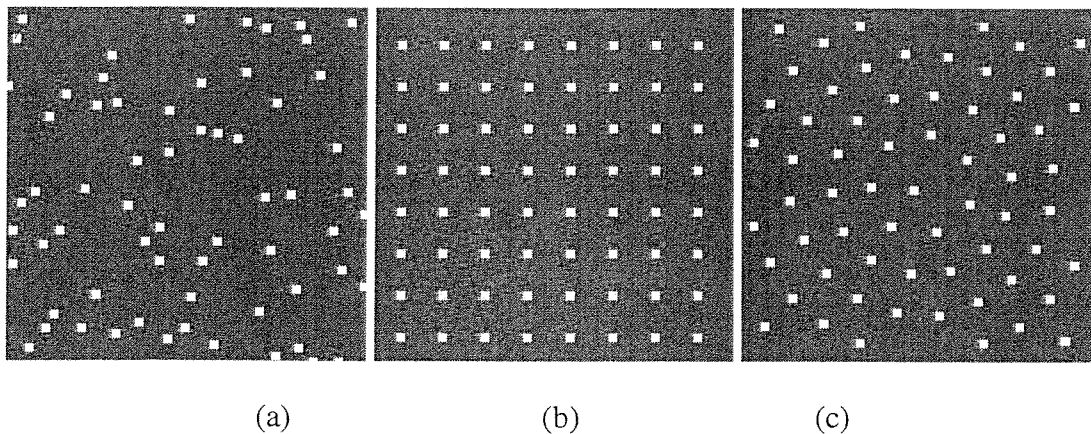
With the carrier and message templates correctly designed the watermark is ready to be embedded in the image. This process is fundamentally simple compared to the above processes of creating the message template and carrier. The embedding process begins with the determination of the message that is to be embedded. That binary message of 1s and 0s is turned into a message of +1s and -1s respectively. This

information is embedded at the locations given in the message template. Once the data is embedded in the message template it will appear as shown in Figure 3.

Figure 3: Embedded Message in Message Template (Scaled for Display)
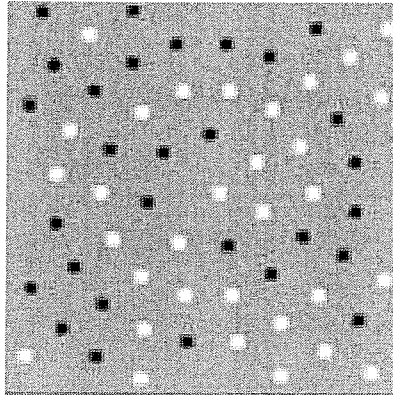
The message template with the information embedded in it is then convolved with the carrier signal to produce the watermark that will be embedded in the image. This convolution, however, is different from a typical convolution in that it is a cyclical convolution. The cyclic convolution involves a wrapping the edge information back onto the image for the convolution instead of truncating that edge information when edges are reached. Figure 4 shows three examples of the cyclical shift that is taking place, the same trend that occurs in the vertical sense that is shown in the horizontally below. The cyclic convolution of the carrier and the message template produces another 128 x 128 pixel array with floating point values. This convolution can be done in the spatial domain as a convolution, or the convolution can be done in the frequency domain as a multiplication. The array that is produced is the watermark, and is going to be added to each 128 x 128

block of the image in the spatial domain. However, before the watermark can be added to the image the strength of the watermark must be scaled so that the digital count values will not wash out the image information. The factor by which the watermark digital count values are scaled directly controls the visual transparency and the robustness of the watermark. The scaled watermark digital count values can then be added to the image. Figure 5 shows the original image and the same image with the watermark information embedded.

Figure 4: Cyclically Shifted Versions of Template 4(b) (a), Template 4(c) (b) and Lena (c)
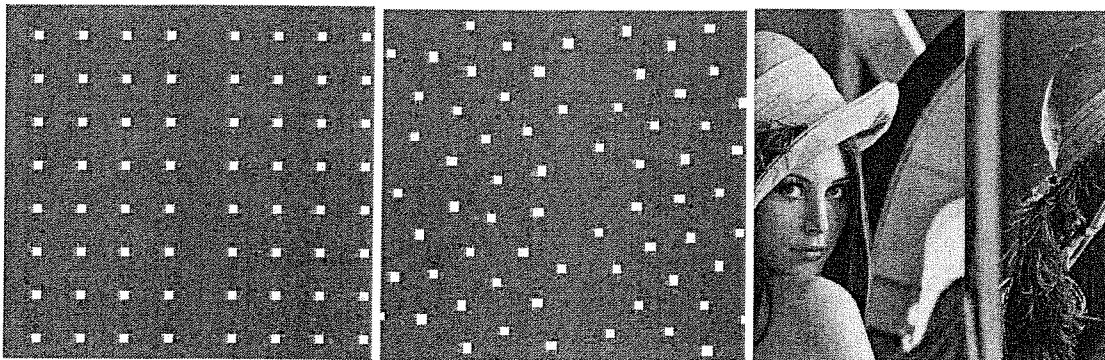
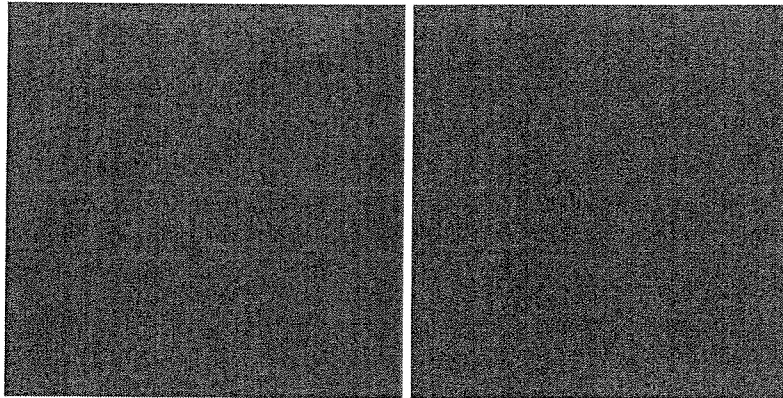(A)   (B)   (C)

Figure 5: Original (A) and Embedded (B) Images

(A)   (B)

(A)   (B)

With the embedding process complete, the last step to discuss is the extraction

process. The extraction process is a three-level process. The first level of the process

involves breaking the image up into 128 x 128 blocks. The second level of the process

involves determining the amount of image information in each of the 128 x 128 image

blocks. The third level is based on the signal-to-noise ratio of the image blocks. All of the

image blocks are added together using the information from level two to weight each

block based on the amount of image information in it. The more non-uniform information

that is contained in the image block the less weight that that block will hold. The weight of the individual blocks becomes important since there is an average taken of all the blocks. The benefit of computing this average is that it should eliminate the image information (the law of large numbers dictates that each pixel in this 128 x 128 block should have the same average digital count) and keep only the watermark information. This averaging process serves also to strengthen the watermark, this is due to the consistent location of the embedded messages that add during the averaging process. The watermark information can then be extracted via two methods, one in the spatial domain, and the other in the frequency domain. The spatial domain extraction is done by correlating the original carrier with the average block. The correlation that is done is similar to the convolution in that they are both cyclical. The reason for the cyclic correlation is that it will account for cropping and rotation, as well as any possible shifting. The reason that these attacks can be corrected for is that the templates were properly designed to eliminate false positives in the cyclic correlation. The process of cyclically correlating the carrier with the image will return the cropped, shifted or rotated message template. The correlation should leave the message template with noise introduced. The message is then extracted by isolating data locations where information exists. The process for isolating the data locations is done by finding peak values that are greater than some number "X" standard deviations from the mean. The process for spatially extracting watermark was not used due to its costly nature (computational time). The process used was the frequency domain method. The advantage to extracting in the frequency domain is that the number of calculations that are needed is greatly reduced. In the frequency domain extraction method a cross correlation is preformed between the

image block and the carrier in the frequency domain. The cross correlation that is used to produce the correlation image is described in Equation 1.

$$Message = |I(u,v)||C(u,v)|e^{i(\phi_I(u,v)-\phi_C(u,v))} \qquad (1)$$

In the above equation I is used to represent the image, and C is used to represent the carrier. The resultant image is then back transformed through the fast Fourier transform (FFT) into the spatial domain for processing. The process for back transforming the image is the same as above. Values are isolated that are above a number "X" standard deviations, and those are message locations. The message locations are then padded by a number of pixels in each direction to make certain that all of the signal information is gathered. The remaining values are all determined to be noise.

3.0 Project Methods

3.1 Goals

The initial goal of the research was to determine a method for extracting the embedded watermark information from Stirmark random-bend attacked images. The goal of the research has changed and progressed over time away from the extraction from random bend to extraction methods from linearly distorted images. The focus became to find the optimal method that maximizes the signal-to-noise ratio of the extraction.

3.2 Approach

The beginning of the research was done to create an understanding of the Eastman Kodak watermarking algorithm. In order to develop this understanding a significant amount of time was invested into developing code. The code preformed the same functions as the Eastman Kodak algorithm. The code was able to embed a watermark into an image, and extract that message out of the image. The code was written to perform these functions as described above in the spatial domain. There were glitches encountered with the extraction process of the algorithm, which ultimately was due to the fact it only preformed a level 1 extraction. The lack of extraction levels was due to the lack of knowledge as to their existence. This recreation of the algorithm, although taking a significant period of time, resulted in a much deeper understanding of the overall process. One time consuming snag that developed during this process, other than the lack of information about the other levels of extraction, was a lack of information about message location extraction order. Knowing the binary message that should have been extracted from the image, the actual extracted message did not match. Upon closer examination, however, the message templates that were extracted did match. Therefore it became evident that there was a pattern in the order that the binary information was read out from the message template. This piece of information was not defined in the publication describing the Eastman Kodak algorithm. After a period it was determined that the extraction of the binary message was not crucial to the success of the research, and thus was abandoned.

From the implementation of the Eastman Kodak algorithm in the spatial domain, it became quickly evident that the process must be implemented in the frequency domain.

This need developed due to the long amount of time that the cyclical convolutions and correlations took to run in the spatial domain. With each pixel shift in the spatial domain the carrier array had to be re-calculated. This convolution and correlation that required thousands of calculations in the spatial domain could be calculated with one step in the frequency domain. Therefore using the fast Fourier transform (FFT) the carrier and image information were brought into the frequency domain to perform calculations in a fraction of the time. In the case of simple embedding and extraction that was developed to mimic the Eastman Kodak algorithm the frequency domain calculation were orders of magnitude quicker than those preformed in the spatial domain.

The images supplied at the beginning of the research were those of multilayer TIFF images. These images considered of three channels. A problem arose with the determination of whether the extraction would have to be run on each of the three channels and then reassembled, or if one channel would be sufficient. The determination of this process was made by examining the information that was contained in each channel. Upon close investigation it was found that each of the three channels contained the same information and the channels were layered presumably to create the TIFF image. Later in the research these, TIFF images were replaced and interchanged with Bitmap images, and the results held as to the channel information

When the project began, the goal of the research was to determine a way to locally extract the embedded watermark information from a random-bend attacked image. The random-bend attacked image is a product of a watermarking algorithm benchmark program called Stirmark. The Stirmark program produces this random-bend

attack by randomly distorting the locations of pixels in the image. An example of the

random-bend attack can be seen below in Figure 6.

Figure 6: Stirmark Random Bend Attacked Image (A) and Original Grid Image (B)
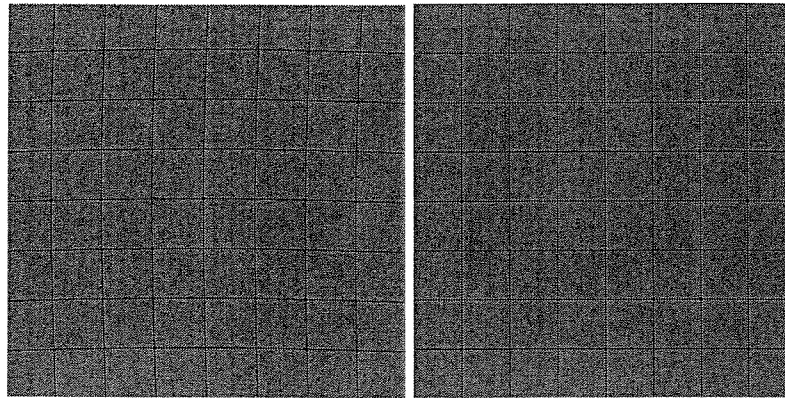
(A)   (B)

The Stirmark produced images above contain small random distortions. Problems arose

in the research when the Stirmark code used to create additional random-bend attacked

images yielded only severe distortions. These distortions made the extraction of the

embedded information far too difficult for the scope of this project. The support could not

be found to sufficiently understand the Stirmark program to alter the overall distortion to

a level that would be conducive to the extraction of valid results. This lead to the change

from extraction from random-bend attacked images to linearly-distorted images.

When the project began, the goal was to extract an image that contained a 64-bit

message. At some point during the course of the research, a carrier image was lost

therefore making the initial watermarked image set useless. This initial set was replaced

by a new set; however, the new set contained only a 24-bit message. The bit information was not the only thing that changed. By changing the number of the bits in the message it in fact changed the pattern of message locations. This new message also contained the message at a higher frequency so that it would be easier to extract. This change in the message locations required changes to be made to the early algorithm, although message location is no longer important in the code.

The decision was made to abandon the extraction from random-bend attacked images. The decision was made to work on linearly-distorted images. The linearly-distorted images were created using a program that was developed by Chris Honsinger. The program allowed the user to input X and Y values that created a vector to perform the linear distortion. An extreme case of a linearly distorted image can be seen below in Figure 7. The extreme case is show to demonstrate the distortion that is used, however, distortions not as severe were used in this study.

Figure 7: Linearly Distorted (A) and Original (B) Grids
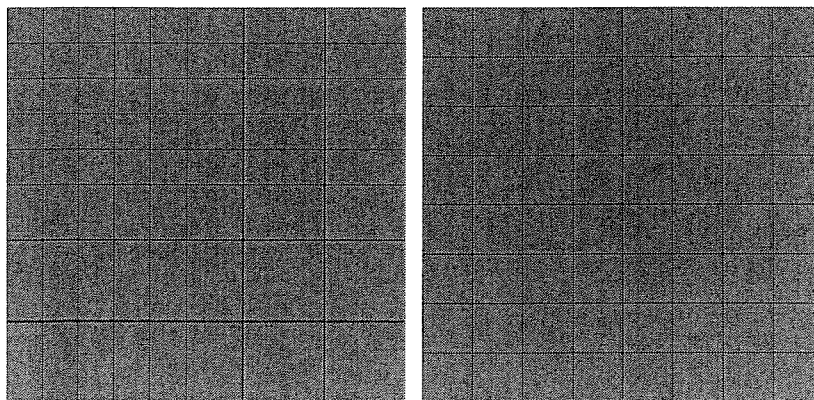
(A)   (B)

As it can be seen in the above images the distortion originates from the upper left corner (origin) of the image. This distortion then gradually increases as a function of the vector such that the distortion in the middle of the image is fundamentally zero. The distortion then continues to the edge, and reaches its full size. While developing these distorted images, the algorithm had to be tested to determine maximum range of distortion from which a viable message could be extracted. The process that was used to determine this range was merely to find a limiting case. The way to determine this limiting case was to attempt extractions from images of varying linearly distorted images. The way the limiting occurred was to look at the correlation images and determine whether peaks could be isolated from the noise. After running a number of distorted images through the algorithm, the limiting case was determined. The limiting case was determined as the linear distortion vector of 10 units in the X, and 10 units in the Y direction. With the limiting case determined, it set out a range of distortion values over which the windowing and subsetting techniques could be tested.

Once the range of the distortion was determined, a task was suggested to determine if the distortion amount could be determined from an extracted message. The method by which this problem was initially approached was to look at the extraction from an average image block. This approach would show the maximum distortion value in the image, however, since the distortion was not constant throughout the entire image the data that was determined became useless. The next approach to this problem, that was suggested, was to create a series of distortion vectors. One vector in each image block would show the localized distortion in each block. At the time this task was initially attempted a solution could not be found that would accurately and effectively estimate

the distortion on a localized level. The approach taken was to generate the correlation image for the distorted-image block, and then to correlate the message location template to the correlation image. In order to generate a vector map, this correlation would have to be done in the spatial domain using a cyclical correlation. The cyclical correlation that was used required an enormous amount of computational time to run. The results of the correlation did not provide the desired distortion vectors. Furthermore, even if the distortion vectors had been generated, the result had no apparent application at the time. As discussed below, in the future work area, there may be a way to subset, and estimate the distortion locally that could be of significant benefit to implement.

There were two types of windows that were chosen to be used to window the carrier for extraction. The windows that were used were the Hanning and Hamming windows. Below in Figure 8(A) and 8(B) the Hanning and Hamming windows can be seen respectively.

Figure 8: Hanning (A) and Hamming (B) Windows

(A)   (B)
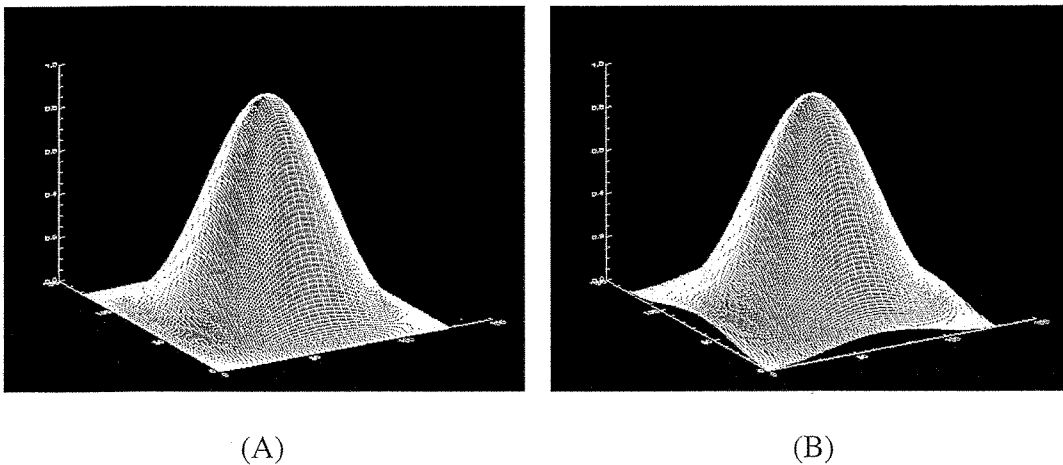
The Hanning and Hamming windows are functions that are applied to the carrier image in the frequency domain. The application of these windowing techniques serves to maintain the high frequency information while tailing off the low frequency information. Windows serve to reduce edge effects, increase signal, and decrease noise. Therefore, fundamentally, one would expect the results to show that the extracted messages using windowed carriers would have a higher signal-to-noise ratio, compared to those carriers which had not been windowed. Figure 9 shows the process of windowing.

Figure 9: The Windowing Process
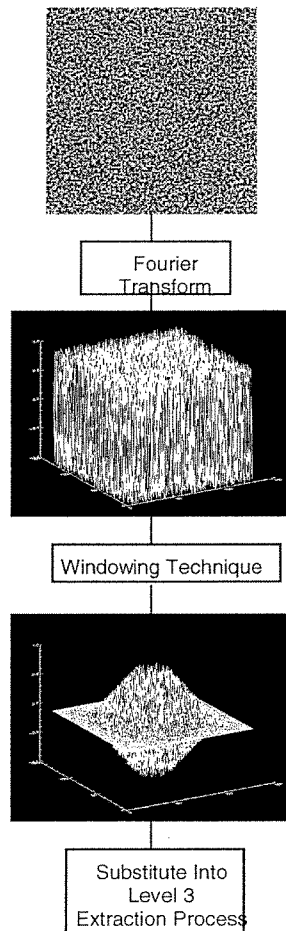
The problem of subsetting was approached within the limitations of computational power. The subsets that were extracted involved breaking the 128 x 128 pixel carrier into four 64 x 64 pixel subsets, and into sixteen 32 x 32 pixel subsets. Although there was a desire to increase the number of subsets that were used, doing so and going to sixty-four 16 x 16 pixel subsets would have quadrupled the amount of computational time. This computational time was already pushed to the limit due to the fact that there was 16 individual cross correlations preformed in the case with no subsetting. Therefore, that translates into 64 computations, 256 computations, and 1024 computations in the 4, 16, and 64 subset cases, respectively. The actual process for subsetting involved cross correlating each piece of the subset with the entire image block. After each piece was correlated all of the pieces were added together. The addition of the subsets returned the completed extraction, which was then processed to extract the signal and noise information. This process was repeated for each image block, in the case of the 512 x 512 pixel image that was used in the present case, there were 16 image blocks. Figure 10 shows an illustration demonstrating how the subsetting process works for the case of 4 subsets.

Figure 10: The Subsetting Process
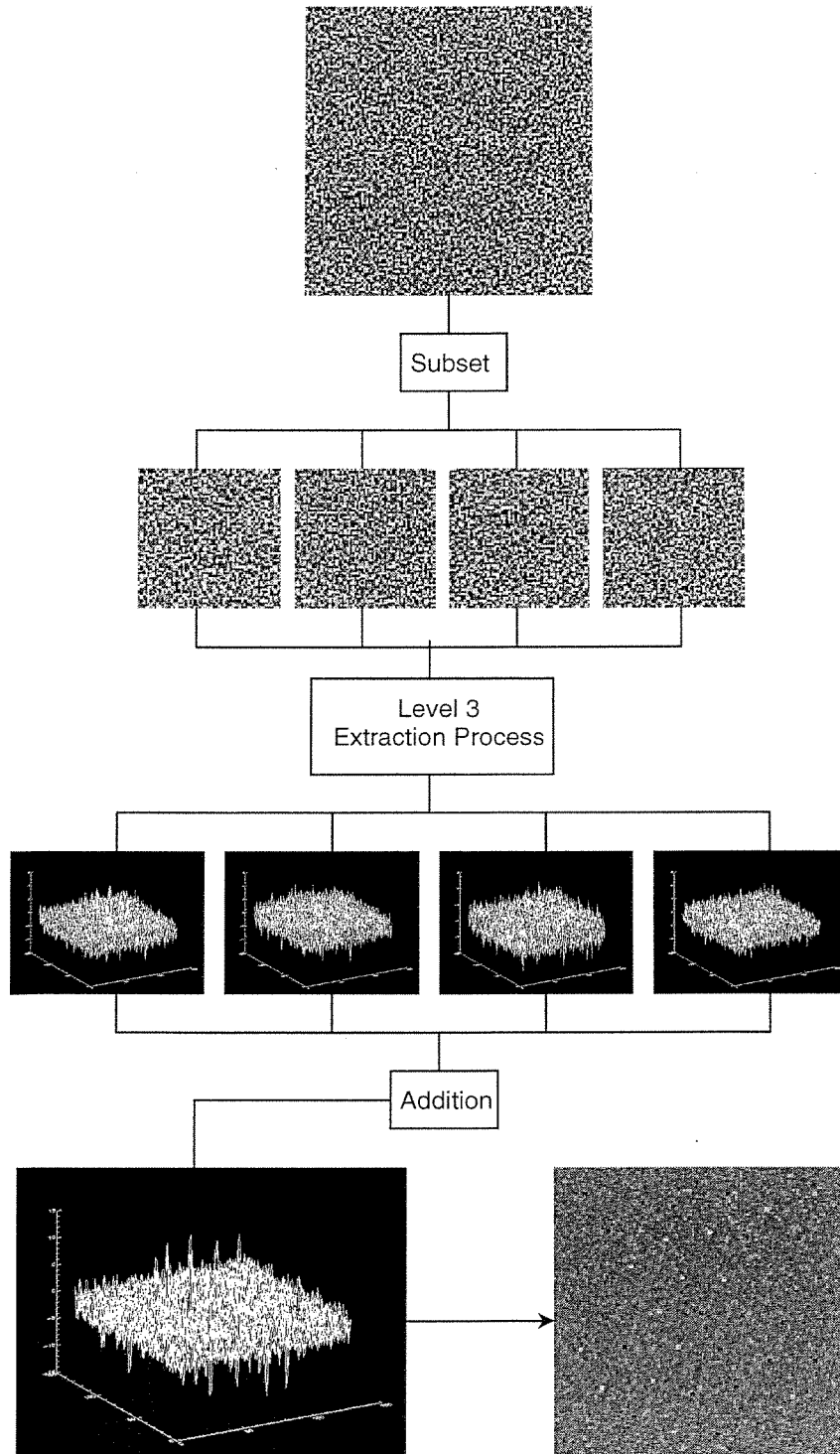
The process that is shown in Figure 10, is the same process that would be used in the 16, and 64 subset cases. The fundamental thought behind subsetting as a method to increase signal-to-noise ratio, is that as the amount of distortion increases, by breaking up the carrier into smaller and smaller pieces that one will have a large likelihood of correlating with at least some portion of the carrier. There is a problem with this, that however, is the subset of the carrier must remain large enough so that it does not return correlations that contain false positives.

The computation of the signal-to-noise ratio provided another problem whose solution needed to be determined. There are numerous methods that can be used to evaluate the signal-to-noise of a system. In this study, there were two methods used before the final method was determined. The first method used took the standard deviation of the signal and divide that number by the standard deviation of the noise. While the result of this did give a signal-to-noise ratio, the accuracy was highly questionable based on the deviation of the signal information due to padding. A second method that was used involved taking the variance of the signal divided by the variance of the noise. This method produced similar concerns with the amount of variance being increased due to signal area padding. The final and chosen method for evaluating the signal-to-noise ratio of the extracted data involves taking the highest signal peak, and dividing that value by the standard deviation of the noise information. This method corrects for the problem of padding that is introduced through the other two methods. This method only uses the highest peak in the correlated area to determine the signal-to-noise relationship, therefore, there is a small chance that the signal-to-noise could be skewed. This method for determining signal-to-noise is maintained throughout the entire

process. Since it is only necessary to compare the results to each other, by maintaining the same method for calculation it maintains the basis for comparison.

At this point tests were run on the algorithm, and data was generated. It became necessary to determine if the results that were acquired were valid. These tests were necessary because, as was stated above, the project would have no basis for comparison. In order to test the results a simple theoretical case was constructed. The case involved an image with random noise added, and the image without noise. The image with random noise was taken to be the carrier, and the other was the image. The carrier and image were cross correlated in the manner with which the message was to be extracted. The theoretical validation pointed out an error that might have gone unnoticed otherwise. It became quickly evident through the theoretical validation that there was a shift occurring in the data that was unaccounted for. This shift was causing the message locations to be moved from the locations where they should be located. The solution to the shift was simple, in that it only required a shift of the points back around the center of the image. The result of the correction changed the results of the entire experiment. The theoretical model benefited the research in two ways, first it corrected the results, and second it validated the unknown answer.

3.3 Process

The process that was ultimately used to determine the optimal method to maximize the signal-to-noise ratio combined a number of the above approaches. The method began with a distorted image. That distorted images was 512 x 512 pixels in size.

That distorted image was then broken down into sixteen 128 x 128 pixel image blocks. The image blocks were then evaluated on an individual basis, and the signal-to-noise ratio for each of the extractions was determined. To determine the optimal method of extraction, three techniques were implemented. The first technique that was implemented was the extraction in which the carrier image was not windowed. The second techniques was one in which the carrier was windowed using a Hanning window. And the last was a technique in which the carrier was windowed using a Hamming window. In the first case the carrier that had not been windowed was cross correlated with each of the 16 image blocks. The result of this correlation was an image of correlation values. The image was then passed through a thresholding function that isolated areas that were above 3.5 standard deviations from the mean. These isolated areas were then passed through a blurring filter and thresholded again. These enlarged areas then became the map of message locations. The rest of the correlation image then was assigned to be the noise information. This map of message locations (signal) and noise were saved to be used for each of the following techniques. The signal-to-noise ratio was then determined as described by taking the peak signal value and dividing that by the standard deviation of the noise. This yielded the 16 signal-to-noise ratios for the extraction method that used a carrier that had not been windowed. This process was then repeated for the two windowed cases, giving 16 signal-to-noise ratios for the Hanning windowed case, and 16 signal-to-noise ratios for the Hamming windowed case.

After extracting the above data, the next test was to determine if subsetting in each of the three cases would increase the signal-to-noise ratio of the extraction. To determine this result first the carriers were subset. This technique involved first breaking

the 128 x 128 pixel carrier into the 4 equal subsets. Next each piece of the carrier was cross correlated with the first image block. This produced 4 correlation images for the first image block. These 4 correlation image were then added together to produce a total correlation image for the first image block. This process was repeated for each of the 16 image blocks. The resultant total correlation images were passed through the same functions to determine the signal-to-noise ratios described prior. This subsetting process was then repeated for the Hanning and Hamming windowed cases. Upon completion of the data collecting for subsetting with 4 subsets, the process was repeated exactly, with the only difference being that the carrier images were in this case broken up into 16 equal subsets instead of 4.

With the windowing and subsetting techniques implemented, and the signal-to-noise ratios of the extractions determined, the final step was to interpolate the data.
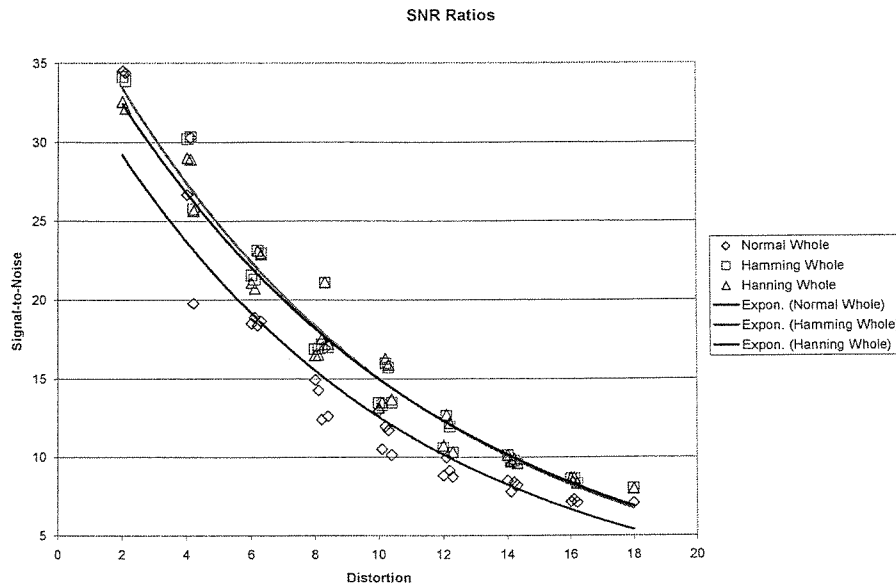
4.0 Results

4.1 Windowing

The first set of results to be examined are those that exist in the case of the windowed functions. The signal-to-noise ratios for the non-windowed, Hanning, and Hamming functions were compared as the distortion vector in the image was increased. Figure 11 is the plot of these signal-to-noise ratios.

Figure 11: Plot of Signal-to-Noise Ratios for Windowing Techniques

SNR Ratios

As can be seen above in the results of the signal-to-noise ratio plots, the best signal-to-noise ratio is yielded using a Hamming window. The data that is shown above as the result of the windowing process is intuitively correct. The intuitive nature of the above data is contained in the fact that as the amount of distortion increases, the overall signal-to-noise decrease. Furthermore, one expects that windowing will increase the signal-to-noise ratio in cases that contain distortion. As was discussed earlier the reason for the increase in the signal-to-noise ratio due to the windowing process comes from the windows ability to maintain signal, and reduce noise (High Frequency Information). One important item to note is that while the Hamming window provides the highest overall signal-to-noise ratio, it is not always higher. The results show that for small distortions,

the non-windowed carrier provides a better or equal signal-to-noise ratio, and does so at a fraction of the computational cost.

The next piece of the data that was examined was the results of subsetting the carrier, and then performing the cross correlation. To determine the benefits of subsetting the carrier, the subset signal-to-noise ratio for the subsets of 4 and 16 were compared to signal-to-noise ratio achieved when using the entire carrier of extraction. To start the interpretation of the data the results for the non-windowed carrier were first examined; those results can be seen below in Figure 12.

Figure 12: Signal-to-Noise Ratio Comparison for Subset Cases
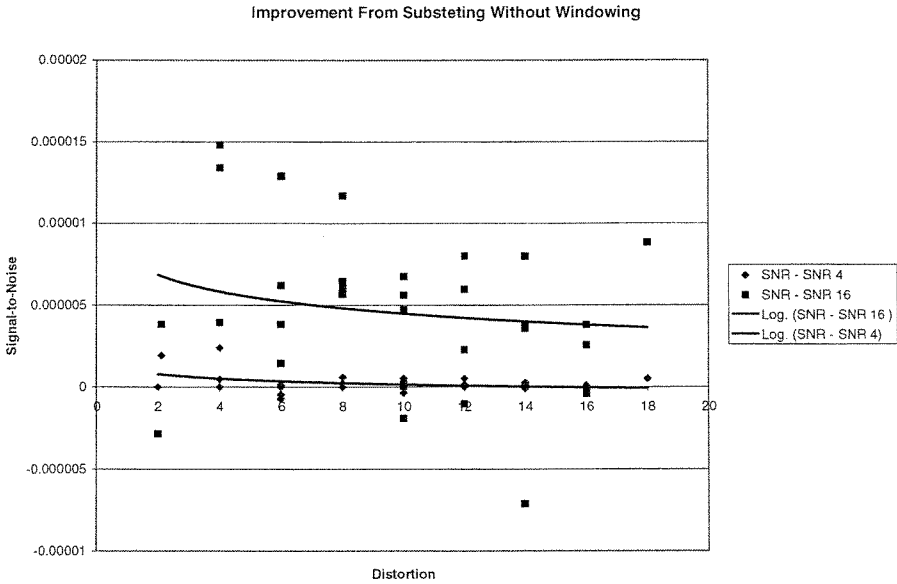

The above plot shows the difference between the signal-to-noise ratio of the whole carrier and the subset carriers. Positive values indicate the there was an improvement made through subsetting, while negative values indicate that subsetting lowered the signal-to-

noise ratio. The difficulty in these above results is that intuitively one would expect that as distortion increases then larger gains would be made through the subsetting process. The determination was made that the algorithm, as implemented in this case, begins to deteriorate in the case of distortions vectors above a length of approximately 8 units. The other problem that arises with the results of the data is that the gains that are made are marginal at best. All of the increase in signal-to-noise ratio that were made by subsetting in the non-windowed case were below 0.00001, and those were only in the 16 subset case which requires 240 more calculations than leaving the carrier whole. The problem then is whether to use the extra computational power to produce only a marginal increase in the signal-to-noise ratio, or to apply it to speeding up the extraction process.

The next subset that was examined was the Hanning windowed subsets. Similar to the above case, and as defined earlier there was three Hanning windowed carriers. One carrier remained whole, one carrier was broken into 4 subsets, and the final carrier was broken into 16 subsets. Below in Figure 13 is the plot showing improvements made by subsetting the Hanning windowed carrier.

Figure 13: Improvements from Subsetting the Hanning Windowed Carrier
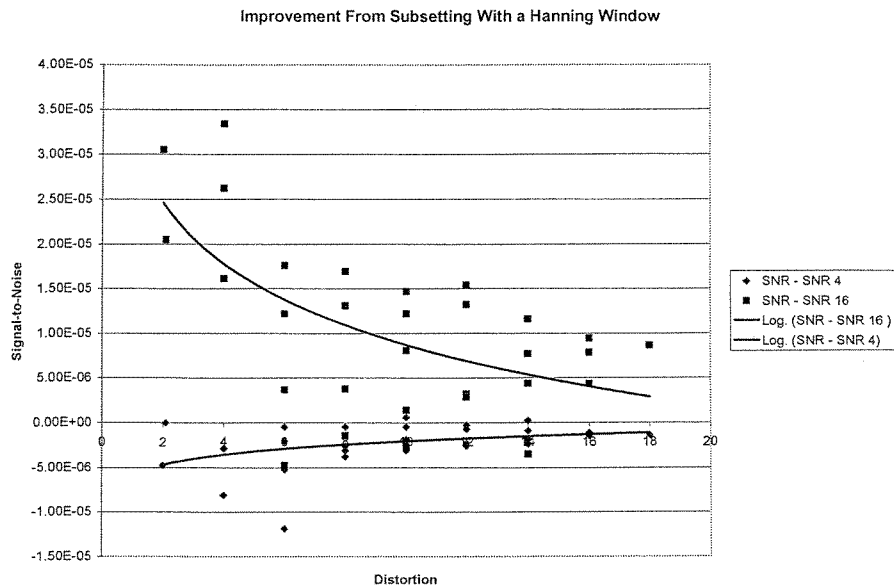
The above plot shows results similar to the case of the non-windowed carrier. Both sets of results fall off as the distortion vector gets larger. The gains achieved by subsetting in the Hanning windowed case are larger than those in the non-windowed case, however, they are still very marginal gains. These marginal gains most likely are not worth the computational power necessary to subset. There is something slightly unusual about the above plot. The anomaly can be seen in the fact that subsetting down to 4 subsets results in a decrease in the signal-to-noise ratio.

The final subsetting that was examined was that of the Hamming windowed carrier. Again like the above two cases the Hamming windowed carrier was split into subset so 4 and 16, and also left whole. The plot of the improvements made from subsetting can be seen below in Figure 14.

Figure 14: Improvements from Subsetting the Hamming Windowed Carrier

Improvements From Subsetting With a Hamming Window
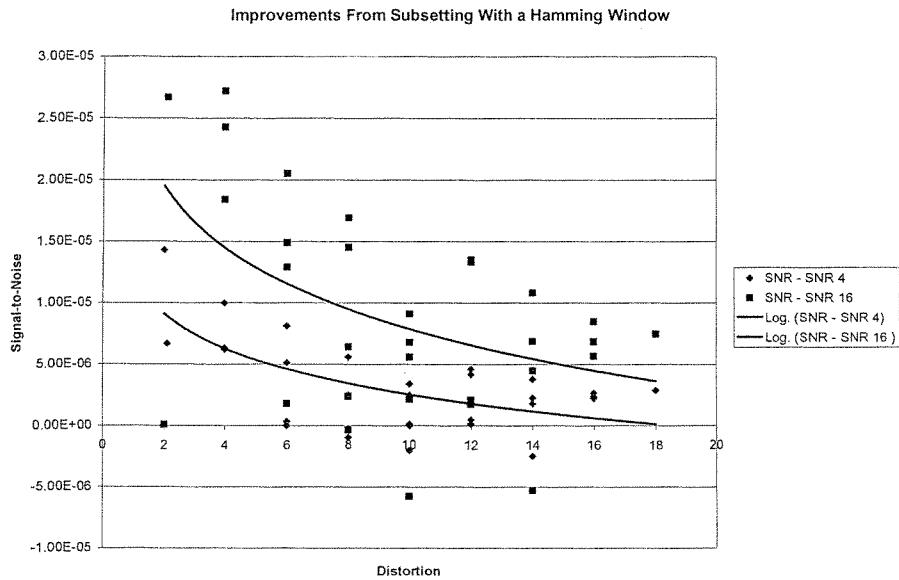
The above results are similar to the previous two cases, they shows the same type of trend in the results. The same principle concern arises in the above data. That principle concern being the cost-benefit relationship associated with spending the computational power and time to subset the carrier and cross correlate each subset piece.

Based on the results of the data in all three of the above cases, it can be safely said that there is no significant gain made via subsetting. Therefore it is impossible to offset the amount of computational power that is required to generate the result. Based on these conclusion it would be advisable to skip the subsetting process or investigate more effective ways to perform the subsetting. The problem with abandoning the subsetting process all together is that fundamentally there are strong suggestions that there is an optimal small subset size. At this subset size the distortions would begin to disappear and

the correlation would increase. This increase in correlation would yield a much higher signal-to-noise ratio.

5.0 Future Work

5.1 Future Windowing

Given an increased timetable, there is more work that could be done to experiment with windowing of the carrier. Based on the success that has already been made in increasing the signal-to-noise ratio by just using two basic windowing methods the Hanning and Hamming methods, it shows this area is worth the effort to continue research. The future of the research in this area could, in the short term, be focused on implementing other existing windowing techniques to determine their effect on the signal-to-noise ratio in the distorted images. While in the long term a method could be investigated that would adapt the method in which the windowing techniques are applied. This adaptive windowing could be a method for determining an estimate for the amount of distortion in an image block, and then determine if it is worth the computational time to window the carrier, or if that computational power could be better used in another image block. This may require more computational power to estimate the amount of distortion in each image block, than it would to just window each carrier. The second case is more likely than the first; however, it is still worth investigation.

5.2 Future Subsetting

The future work to be done in the area of subsetting would be in automatic image registration. Through a technique of image registration below it becomes possible to account for rotation, translation, and scale changes in a image. The fundamental basis for subsetting is to isolate a small enough portion of the image that the distortion changes can be in terms of the above three distortion terms. The automatic image registration begins with the idea that if you have two images that vary only by a translation then Equation 2 is true. (Xie, 2003).

$$g(x, y) = f(x - x_0, y - y_0) \tag{2}$$

Then by taking the Fourier transform of the above images the resultant image are related by Equation 3.

$$G(u, v) = F(u, v)e^{-i2\pi(ux_0 + vy_0)} \tag{3}$$

At this point an equation can be used from the Fourier Shift Theorem. This equation says that two Fourier transforms can be related through the ratio of their phase differences. Equation 4 shows how to determine the ratio of phase that is the difference between the two images.

$$R = \frac{F(u, v) \cdot G^*(u, v)}{|F(u, v)||G(u, v)|} \tag{4}$$

In order to account for rotation and scale terms there are further steps that must be taken. The first step is to change the coordinate space that is being used. The change in space should be to one which is rectilinear with angle and phase. The seemingly best coordinate space to convert to would be Log-polar space. The transformation equation from Cartesian space to the Log-polar space can be found in Equation 5.

$$X = e^{\log(\rho)} \cos(\theta)$$
$$Y = e^{\log(\rho)} \sin(\theta)$$

(5)

Using the converted Log-polar space it allows the relationship between two images that vary by rotation, scale, and translation to be defined. Equation 6 is the demonstration of this relationship.

$$g(x, y) = f(\alpha(x \cos \theta_0 + y \sin \theta_0) - x_0, \alpha(-x \sin \theta_0 + y \cos \theta_0) - y_0)$$

(6)

Where in the above equation θ_0 represents the change in rotation, α represents the change in scale, and x_0, and y_0 represent the translational change. Now by taking the Fourier transform of the above, the two images can be related in magnitude by Equation 7

$$|G(u, v)| = \frac{1}{\alpha^2} \left| F(\frac{1}{\alpha}(u \cos \theta_0 + v \sin \theta_0), \frac{1}{\alpha}(-u \sin \theta_0 + v \cos \theta_0)) \right|$$

(7)

The Fourier transform of these images are thus independent of translational differences because all that matters in the Fourier case is the phase. Converting Equation 7 into Log-polar space will yield the final result of the process. The substitution and simplified equations can be seen in Equation 8 and Equation 9.

$$|G(u,v)| = \frac{1}{\alpha^2}\left|F(\frac{1}{\alpha}e^{\rho}\cos(\theta - \theta_0), \frac{1}{\alpha}e^{\rho}\sin(\theta - \theta_0))\right| \tag{8}$$

$$|G(u,v)| = \frac{1}{\alpha^2}\left|F(\rho - \ln\alpha, \theta - \theta_0)\right| \tag{9}$$

Using the above equation and knowing two important facts allows for the scale and rotation to be determined. The facts are that the scaling that occurs in the image results in a shift α along the log radius. The second fact that must be known is that the rotation results in a translational shift about the θ axis. Using these facts about the scale and rotation is become possible to use Equation 3 to determine the scale and rotational changes. This information would could ultimately make a significant impact on the subset extractions. The way to implement this method towards the goal of subset extraction would be to determine the shift, rotation, and scale changes at each subset location. After determining these changes the subset could then be changed to maximize the correlation values. By minimizing the subset size to isolate the probability that rotation, scale, and translational changes would be the only likely distortion, then using the above method to correct for those distortions the extraction signal-to-noise ratio could be maximized. There is only one foreseeable downfall to the above method, the

computational time that would be required to implement the method. To significantly

minimize the likelihood that the subset would only be distorted by scale, rotation, and

translation the number of calculations would be maximized. Then to apply the above

method to each of the subsets would further increase the computational time. In order to

determine the cost-benefit relationship of this method there would have to be significant

tests run on a number of distortion levels.

6.0 Conclusion

Based on the results of this research, it becomes evident the methods that should

be used to maximize signal-to-noise ratios of the extracted messages. From the research,

and above given data, it can be seen that windowing provides the most notable increase in

signal-to-noise ratio. This notable increase could be maximized by using a Hamming

window on each of the carriers prior to the cross correlation process. For whatever

reason, if the use of a Hamming window was not an available solution, a Hanning

window could be used with little loss in signal-to-noise. Moreover, although there is an

increase made to the signal-to-noise ratio that is associated with the subsetting of the

carrier. The increase is so small that it is not worth the amount of computational time or

power that it takes to arrive at the subset results. That is only speaking of the present

case; if the algorithm were able to better extract from more heavily distorted images then

the subsetting process could be more significant.

7.0 References

Gonzalez, R., and R. Woods, *Digital Image Processing*, Second Edition, 2002

Rabbani, M., and C. Honsinger, *Data Embedding Using Phase Dispersion*, IEEE

  Confrence Paper on Secure Images, 2000

Xie, H. N. Hicksa, G. Kellera, H. Huangb, and V. Kreinovich,
  An IDL/ENVI implementation of the FFTbased algorithm for automatic image
  registration, Computers & Geosciences, 2003, Volume 29, pp. 1045–1055

Signal-to-Noise Calculations, January 2004,

  http://emalwww.engin.umich.edu/emal/courses/SEM_lectureCW/SEM_SignalNoise.html