A Spatial and Frequency Domain Analysis of the Effect of Removal Attacks on Digital Image Watermarks‖

3 authors:

Chunlin Song
Jiangnan University
2 PUBLICATIONS **47** CITATIONS

SEE PROFILE

Sud Sudirman
Liverpool John Moores University
54 PUBLICATIONS **400** CITATIONS

SEE PROFILE

Madjid Merabti
University of Sharjah
414 PUBLICATIONS **3,627** CITATIONS

SEE PROFILE

A Spatial and Frequency Domain Analysis of the Effect of Removal Attacks on Digital Image Watermarks

Chunlin Song, Sud Sudirman, Madjid Merabti
School of Computing and Mathematical Science, Liverpool John Moores University, UK
C.L.Song@2004.ljmu.ac.uk, {S.Sudirman, M.Merabti}@ljmu.ac.uk

Abstract- **Digital image watermarking is one of the most widely used techniques for protection of ownership rights of digital images. Its commercial applications range from copyright protection to digital rights management. However, although there have been numerous new proposed digital image watermarking techniques, most are not sufficiently robust against watermark attacks. With the main objective of finding a novel solution to the robustness challenge in digital image watermarking, this paper presents the results of experiments and analysis on the effect of different watermark attacks. The analysis was carried out using two image analysis tools namely Image Histogram and Fourier Spectrum in both spatial and frequency domain respectively. The results identified some common similarities between different types of watermark attack, a property which could be exploited when designing a new solution for a more robust digital image watermarking technique.**

I. INTRODUCTION

Digital watermarking technology has been drawing the attention of researchers and practitioners as a viable method of protecting copyrights for digital data. The technique works by embedding a subliminal signal, called a watermark, into the data without significantly affecting the visual appearance or visual quality of the data. The root of watermarking as an information hiding technique can be traced to ancient Greece as steganography [1], however the science of watermarking is a modern subject and it was only developed in recent years. Recently there are more than one hundred institutes around the world [2] which deal with the issue. The application of watermarking ranges from copyright protection, file tracking and monitoring.

The success of a watermarking technology used in a copyright protection or digital rights management system relies heavily on its robustness to withstand attacks. Watermark attacks are aimed at removing or destroying any watermark signals in the host data. There are four categories based on the classification: removal attack, geometric attack, cryptographic attack and protocol attack [3].

Removal attack aims at the complete removal of the watermark information from the watermarked data without breaking the security of the watermarking algorithm. Most of image processing methods, such as image smoothing and Gaussian noise, belong to removal attack category.

Geometric attack is different from removal attack. Instead of removing the watermark signals out from the watermarked data, geometric attack intends to distort the watermark detector synchronization with the embedding information. Rotation, scaling and translation attacks are the most famous algorithm in geometric attack

The aim of cryptographic attacks is to break the security of watermarking schemes and thus find a way to procedurally remove the embedded watermark information or to embed misleading watermarks. One of the techniques in this category is the brute-force search method. This technique extensively attempts to identify the used watermark algorithm by using a large number of known possible measures and extract the watermark.

Protocol attack is a different type of watermark attack in a sense that it targets the entire concept of using watermarking techniques as a solution to copyright protection rather than the watermark itself. Whereas the other types of attacks aim at destroying, distorting or extracting the watermark signal, protocol attacks aim at producing ambiguities on the true ownership of the data in question. An example of a protocol attack is the copy attack; instead of destroying the watermark, the copy attack estimates a watermark from watermarked data and copies it to some other data with the purpose of attacking the credibility of such watermark in claiming ownership.

In this paper we will look at the effect of removal attack on digital image watermarks. The next section will provide the setup of the experiment we conducted together with the results we have obtained. Thorough analyses of these results together with our findings are reported in section 3. The experiment results and the analysis are presented using Fourier Spectrum and Image Histogram tools to better understand the effect of each watermark attack in frequency and spatial domain respectively.

II. EXPERIMENT SETUP AND RESULTS

A. *Experiment Setup*

The objectives of the experiment are to show the effects of removal attacks on both the host image and watermark data. By understanding these effects, one could a) identify

ISBN: 978-1-902560-24-3 © 2010 PGNet

similarities in effects between different watermark attacks and b) devise a solution to alleviate the effects based on existing technique to similar phenomenon.

The watermarking process used in this experiment used two images namely cover image and watermark image. The watermarking process will attempt to insert the watermark image into the cover image to produce watermarked image using a variant of Discrete Wavelet Transform watermarking technique as described in [4]. Figure 1 shows the Lena and Hat images that are used as the cover and watermark images respectively in this experiment. Both images are identical in dimension (512x512 pixels) and depth (256 grey levels).

(a) (b)

Fig. 1. (a) Cover image – Lena and (b) Watermark image – Hat.

The resulting watermarked image is then subjected to a number of removal attacks including Gaussian smoothing attack, Gaussian noise attack, salt and pepper noise attack, median filtering attack, histogram equalization attack, sharpen attack and JPEG compression attack. The effect of these attacks on the watermarked image is then analyzed by comparing the original watermarked image and the attacked watermark images. To provide a full picture of the effect, we analyze the effect in both spatial and frequency domain using image histogram and Fourier spectrum respectively. It is understood that the two domains are complimentary to each other to provide a complete description on images or any changes to them.

In an image processing context, the histogram of an image normally refers to a histogram of the pixel intensity values. This histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. For an 8-bit grayscale image there are 256 different possible intensities, and so the histogram will graphically display 256 numbers showing the distribution of pixels amongst those grayscale values. The horizontal axis of the graph represents intensity, while the vertical axis represents the number of pixels in that intensity. The histogram of the original watermarked image is shown in Figure 2a.

Image transform converts image data into alternative representations that are more amenable for certain types of analysis. The most commonly used image transform is Fourier Transform which takes spatial data and transforms it into its frequency components or spectrum. The Fourier spectrum of an image is a representation of that image in the frequency domain. Fourier spectrum can be calculated in several ways, but the fast Fourier transform (FFT) method is the most frequently used algorithm. The Fourier spectrum of an image is often visualized as a grayscale image whose intensity corresponds to the strength or magnitude of spectrum. The coordinate of these pixels correspond to the frequency in x-y directions. The low frequency is located at the centre of the image and the high frequency is located around the edges. However for clarity in visual inspection, Fourier spectrum is often visualized as a 3D graph as illustrated in Figure 2b as it is easier to notice any changes in the spectrum by observing the it in this view. Fourier spectrum has an important characteristic in which its total energy is preserved. This means both low and high frequency components are complementary to each other. In other words, if the high frequency component of an image increases, the low frequency component of that image subsequently decreases. Therefore, any shift in frequency band can be easily detected as the change in spectrum strength in both low and high frequency regions.

B. Experimental Result

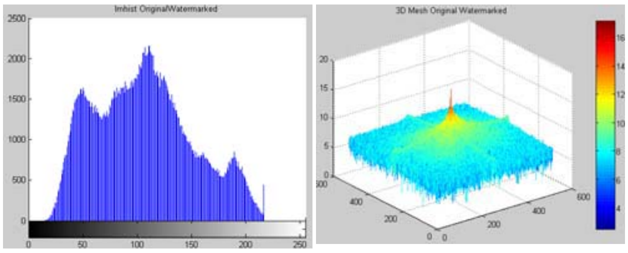
There are seven types of removal attack that will be analyzed in this paper. In paper [5], there are some conclusions have already come out, based on experimental results, we can argue that existing techniques have different sensitivity and robustness to different attacks and the experiment also shows that both LSB and DWT techniques do not possess a complete advantage over each other in terms of robustness to these attacks. Although arguably the DWT techniques are more robust to more attacks than the LSB technique. In this paper, some further research on these attacks by finding out the properties of these attacks.

As we know, most of the watermark attacks methods are image processing algorithms, such as Gaussian smoothing attack, Gaussian noise attack and JPEG compression attack. And theoretically, each image is composed of high frequency components and low frequency components. Our experiment aim to add different attacks into watermarked image to get new watermarked image, hence, new watermarked images should contains the properties of different attacks. Histogram will help us to distinguish the variation of pixel values and Fourier spectrum plays very important role onto justifying and analyzing these properties. The rest of the figures (a) indicates histogram images of different new watermarked image and (b) shows 3D Fourier spectrum of different new watermarked images.

III. ANALYSIS OF RESULTS

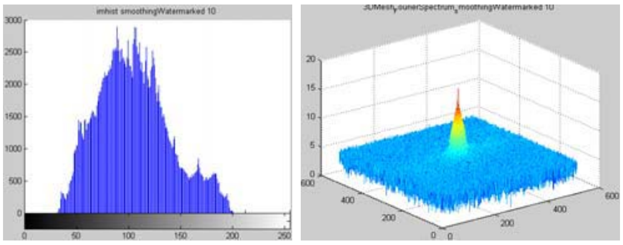
A. Gaussian smoothing

Gaussian smoothing is a process that averages the value of pixels over an area using weighting coefficients derived from a Gaussian function. This process is often used to reduce noise or to reduce pixilation in an image. Visually, the effect of Gaussian smoothing on an image is illustrated in Fig. 10. The amount of smoothing can be controlled by adjusting the width



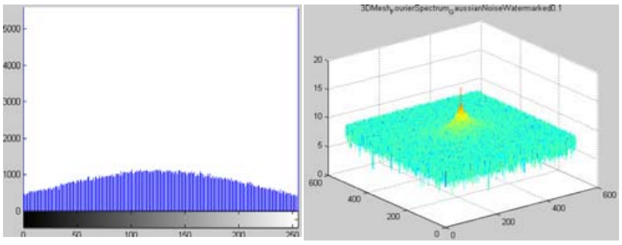
(a) (b)

Fig. 2. (a) Histogram and (b) Fourier spectrum of the original watermarked image.



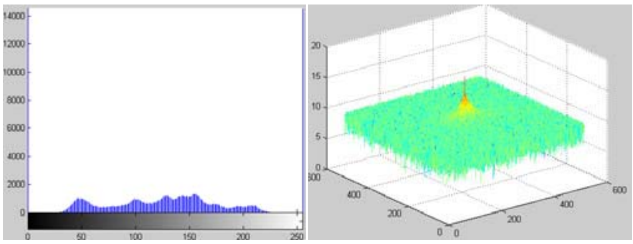
(a) (b)

Fig. 3. (a) Histogram and (b) Fourier spectrum of the Gaussian Smoothed watermarked image with σ = 10.



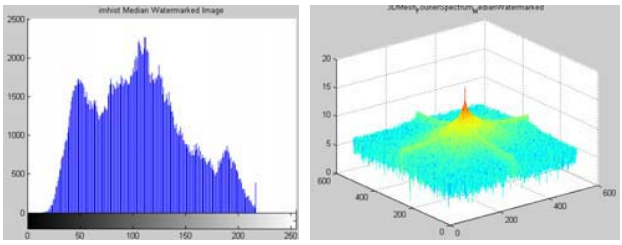
(a) (b)

Fig. 4. (a) Histogram and (b) Fourier spectrum of the Gaussian Noised watermarked image with σ = 0.1.



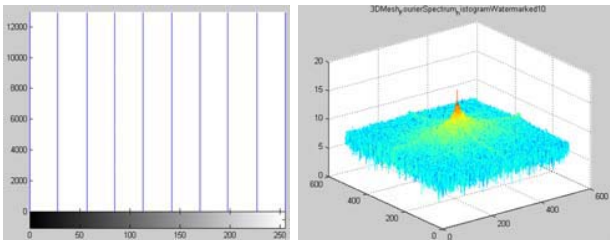
(a) (b)

Fig. 5. (a) Histogram and (b) Fourier spectrum of the Salt-Pepper Noised watermarked image with σ = 0.1.



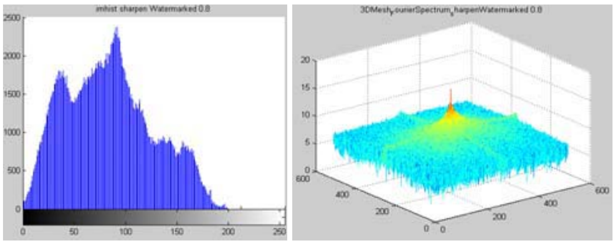
(a) (b)

Fig. 6. (a) Histogram and (b) Fourier spectrum of the Median filtered watermarked image.



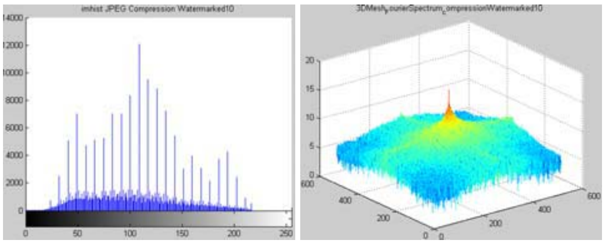
(a) (b)

Fig. 7. (a) Histogram and (b) Fourier spectrum of the Histogram Equalized watermarked image with η = 10.



(a) (b)

Fig. 8. (a) Histogram and (b) Fourier spectrum of the Sharpened watermarked image with σ = 0.8.



(a) (b)

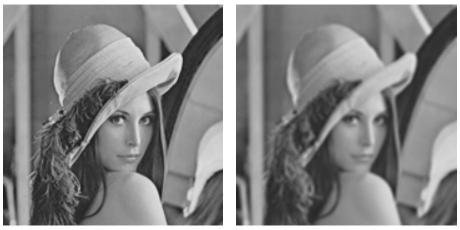
Fig. 9. (a) Histogram and (b) Fourier spectrum of the JPEG compressed watermarked image with σ = 10.

of the Gaussian function. In MATLAB this is the equivalent of adjusting the second input parameter (denoted here as σ).

The experiment uses a number of different σ values to allow better observation on the effect of Gaussian smoothing on watermarked image. The σ values used range between 10 and 200.

An observation on the image histogram of the Gaussian smoothed images shows that as the width of the histogram decreases and the peak of the histogram increases as the level of the smoothing increases. This phenomenon is an indication that smoothing reduces the variation in the image pixel values.

This argument is also backed up by the changes in Fourier spectrum of the watermarked image. As the image is smoothed further, the energy of the signal shifts towards the centre, i.e., the low frequency area of the spectrum. At σ = 10, the spectrum contains a dominant peak at the center and very low value elsewhere.



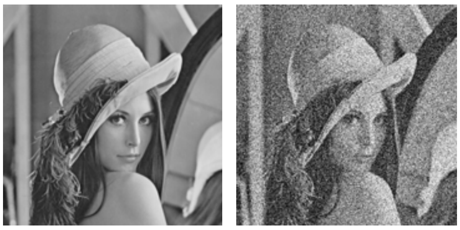
(a) (b)

Fig. 10. Effect of Gaussian smoothing on an image (a) Original watermarked image (b) Gaussian smoothed image.

This phenomenon gives a clear indication that Gaussian smoothing is essentially a low pass filter function. The process reduces the high variation, i.e., high frequency components, of the image and produces low variation pixels or low frequency signal.

B. *Gaussian noise*

Gaussian noise is statistical noise that has a probability density function of the zero mean normal distribution. The power of the noise is controlled by varying the width of the normal distribution. The wider the width the more variation the noise value takes. In MATLAB controlling the distribution width is achieved by adjusting the third input parameter (denoted here as σ) of the *imnoise* function. Visually, the effect of adding noise into an image can be seen in Fig. 11.



(a) (b)

Fig. 11. Effect of adding Gaussian noise to an image (a) Original watermarked image (b) Gaussian noised image.

The experiment uses a number of different σ values to allow better observation on the effect of adding Gaussian noise to the watermarked image. The σ values used range between 0.005 and 0.1.

An observation of the image histogram shows that as moderate strength noise is added to the image the width of the histogram increases and the peak of the histogram decreases. At some point when the width reaches the physical limit of the pixel values (i.e., 0 and 255) more pixel having these values are accumulated. Therefore, as the strength of the noise is increased further there are a large number of black and white pixels in the image with the rest of the histogram seems to look more uniformly distributed.

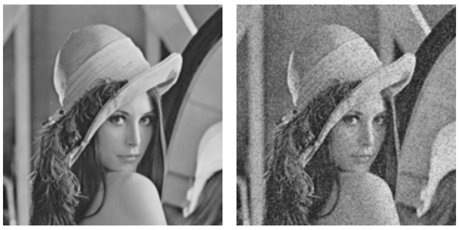
Observation of the Fourier spectrum provides another insight into this phenomenon. As stronger noise is used to attack the watermarked image, the Fourier spectrum shows an increase in strength of the high frequency components of the spectrum and conversely, a decrease in strength of the low frequency components. A direct comparison between the original watermarked image and the most severely attacked watermarked image show a significant change in the spectrum shape especially in the high frequency region.

This phenomenon gives a clear indication that Gaussian noise attack is similar to a high pass filter. The process increases the variation in pixel values to the extent that local edges start appearing in the image. This type of attacks has consequence of removing low frequency watermarks and obscuring high frequency watermarks that may be present in an image.

C. *Salt and pepper noise*

Salt and pepper noise is another example of statistical noise albeit with a very different probability density function (PDF) than Gaussian noise. Its PDF takes the form of two impulse functions at two discrete locations. In MATLAB, the impulse functions are controlled by the second input parameter (denoted here as σ) of *imnoise* function. The visual effect of adding salt & pepper noise on the watermarked image can be seen in Fig. 12.

The experiment uses a number of different σ values to allow better observation on the effect of salt and pepper noise on watermarked image. The σ values used range between 0.005 and 0.1.



(a) (b)

Fig. 12. Effect of adding salt & pepper noise to an image (a) Original watermarked image (b) Salt & pepper noised image.

An observation of the image histogram and Fourier spectrum yields similar conclusion as that of Gaussian noise. In both cases, the image histogram and Fourier spectrum shows the same indication of an increase in pixel value variation and high frequency component of the image as stronger noise is added. This gives a clear indication that salt and pepper noise attack is also essentially a high pass filter function.

D. *Median filter*

Median Filtering is an image processing technique which aims at reducing the presence of noise in an image, hence enhancing the image quality. The median filter is the best-known order-statistic filter, which is a type of non-linear filter. The visual effect of median filter on the watermarked image is shown as Fig. 13.

A comparison of histogram between the original and attacked watermarked image shows little difference between them. Although, closer inspection shows that there is a relatively small reduction in histogram width and relatively small increase in histogram height.

On the other hand, Fourier spectrum provides a more peculiar insight into this event. Our observation shows that, after the image was median filtered the spectrum shows more energy in the high frequency area.

It seems that median filter produces a rather conflicting outcome in spatial and frequency domain. In spatial domain, it has shown to produce lower variation pixel values, but in frequency domain it seems to have increased the frequency components of the image. This phenomenon is perhaps can be explained by the non-linear nature of median filter itself.

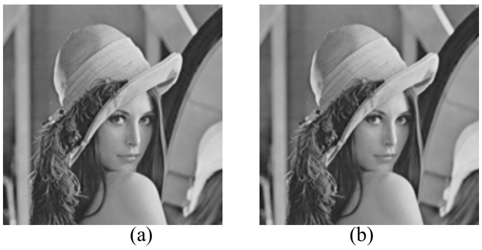

(a) (b)

Fig. 13. Effect of median filter on an image (a) Original watermarked image (b) median filtered image.

E. *Histogram equalization*

Histogram equalization is a method in image processing of contrast adjustment using the image's histogram. This method works by reducing the number of unique grey values in an image and reshape the histogram to approximate a uniform distribution. In effect, histogram equalization is controlled by adjusting the desired number of unique grey values. In MATLAB this is the equivalent of adjusting the second input parameter (denoted here as η) of the *histeq* function. The effect of histogram equalization is shown as Fig. 14.

The experiment uses a number of different η values to allow better observation on the effect of histogram equalization on watermarked image. The η values used range between 10 and 200.

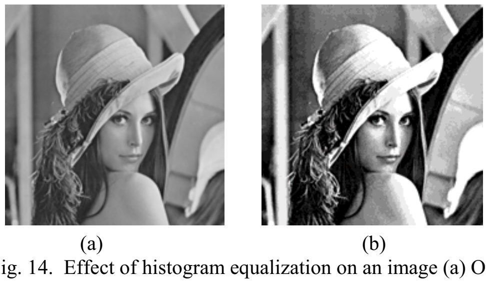

(a) (b)

Fig. 14. Effect of histogram equalization on an image (a) Original watermarked image (b) Histogram equalized image when η is 10.

Since histogram equalization is a histogram reshaping process, its effect can be seen clearly by comparing the histograms of the original and attacked watermarked image. It shows that the numbers of grey values have been decreased dramatically and the shape of the histogram has also changed drastically.

The changes are not as straightforward in the Frequency spectrum as in spatial domain. However, our observation shows a modest shift in energy from the high frequency region to the lower frequency region as the severity of attack is increased. Nonetheless, it is not sufficient to say that there are any significant changes in this domain to warrant any claim.

F. *Sharpen*

Image sharpening can be seen as the opposite of image smoothing whereas the former amplifies the presence of edges in an image. The process is achieved in the frequency domain by using a high pass filter, which intensifies the high frequency components in the Fourier spectrum. A high pass filter *H(ω)* is obtained from its low pass filter *L(ω)* counterpart and calculated using *H(ω) = 1 – L(ω)* formula in the frequency domain. The parameter that controls this filter is similar to that of the Gaussian smoothing function. In MATLAB this is the equivalent of adjusting the second input parameter (denoted here as σ) of the *fspecial* function. The visual effect of sharpen can be seen in Fig. 15.

The experiment uses a number of different σ values to allow better observation on the effect of sharpen on watermarked image. The σ values used range between 0.5 and 0.8.

The histogram of the sharpened watermark image shows that this attack shifts the histogram a few grey values to the left. This means that the overall intensity of the image is reduced. On the other hand, the shape of the histogram is more-or-less maintained with the exception of some peaks which have been eliminated after the attacks.

In the frequency domain, the effect of image sharpening is a lot more obvious. As stronger sharpening coefficient is used, there is a shift in spectrum power from the lower frequency region to the higher frequency ones. This phenomenon is in fact proves that sharpening is indeed the inverse of smoothing, i.e.. the inverse of a low pass filter.

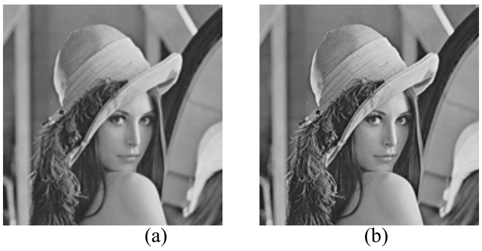
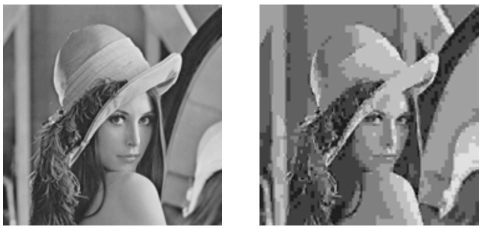

(a) (b)

Fig. 15. Effect of sharpen on an image (a) Original watermarked image (b) sharpened image.

G. *JPEG compression*

JPEG is the most popular image compression technology due to its versatility and good compression ratio. JPEG can perform both lossless and lossy compression, However it is the JPEG lossy compression technology which made it so versatile and popular. As with other lossy compression technologies, JPEG compression produces artifacts on the compressed image. These artifacts are the factor which allows JPEG to be used to attack watermarks in an image. The visual effect of JPEG compression is shown as figure 16.

The experiment uses a number of different compression ratios to allow better observation on the effect of JPEG compression on watermarked image. The compression ratio used range between 10 and 90.



(a) (b)

Fig. 16. Effect of JPEG compression on an image (a) Original watermarked image (b) compressed image.

Lossy image compression often results in blocky patches. These patches are manifestation of regions in the image having the similar values. This can be seen from the histogram of the JPEG compressed image in Fig. 9. The histogram shows a number of pixels having the same grey values and little variations in between. As the image is compressed further, there will be fewer of these variations and higher peaks.

Although the observation in spatial domain suggests that the effect of JPEG compression is more akin to a low pas filter, in which image pixel variation is reduced, an observation of the Fourier spectrum does not give that clear cut conclusion. In fact, we noted an increase in strength on the high frequency components of the spectrum and a decrease in strength on the low frequency components.

It seems that JPEG compression produces a rather conflicting outcome in spatial and frequency domain. In spatial domain, it has shown to produce lower variation pixel values, but in frequency domain it seems to have increased the frequency components of the image. This is a similar conclusion that we gather as with median filter attack.

Table I summarize the effects in spatial and frequency domain. As can be seen from the table, a number of these attacks share similar effects in either spatial or frequency or both domains. This knowledge would provide a valuable insight into the design of a more robust digital image watermarking technique.

We have identified a number of possible solutions derived from the findings presented in this paper. These solutions take advantage of the commonalities between attack types as well as using established solutions to similar phenomena elsewhere.

IV. Summary

In this paper, we have presented a description and analysis of different types of removal attack on digital image watermarks. The analysis was carried out using two image analysis tools namely Image Histogram and Fourier Spectrum in both spatial and frequency domain respectively. The results identified some common similarities between different types of watermark attack, a property which could be exploited when designing a new solution for a more robust digital image watermarking technique.

TABLE I
Effects of different removal watermark attacks in spatial and frequency domain

Watermark attacks	Effect in spatial domain	Effect in frequency domain
Gaussian smoothing attack	Reduces the variation in image pixel values.	Acts as low pass filter
Gaussian noise attack	Increases the variation in pixel values	Similar effect to a high pass filter
Salt & pepper noise attack	Same as Gaussian noise attack	Same as Gaussian noise attack.
Median filter attack	Similar, albeit much smaller, effect as with Gaussian smoothing attack.	Similar effect to a high pass filter
Histogram equalization attack	Reduces the number of unique grayscale values and make the histogram more uniformly distributed	Similar, albeit more moderate, effect as Gaussian smoothing attack
Sharpen attack	Reduces the overall image intensity and amplifies differences around edges	Acts as high pass filter
JPEG Compression attack	Reduces the variation in image pixel values and creating blocks, or uniform regions, in the image.	Similar effect to a high pass filter.

References

[1]. Ingemar J. Cox, Matthew L. Miller, Jeffrey A. Bloom, Jessica Fridrich and T. Kalker, Eds. (2008). Digital Watermarking and Steganography. Burlington, Morgan Kaufmann.

[2]. Cl.Song, S.Sudirman and M.Merabti, "Recent Advances and Classification of Watermark Techniques in Digital Images," *Proc 10th of PostGraduate Network Symposium* 283-288, 2009.

[3]. R. Z. Wang, C. F. Lin and J. C. Lin, "Image hiding by optimal LSB substitution and genetic algorithm," *Pattern Recognition*, vol. 34671-683, 2003.

[4]. P. N. Tao and Eskicioglu, "A robust multiple watermarking scheme in the Discrete Wavelet Transform domain," *Internet Multimedia Management System Conference*, vol. 5601133-144, 2004.

[5]. C.Song, S.Sudirman, M.Merabti and D.L.Jones, “Analysis of Digital Image Watermark Attacks”, 6th IEEE International Workshop on Digital Rights Management, 2010.