# SCC.366 Alternative Assessment Coursework

Matthew Kenely (**38834561**)

m.kenely@lancaster.ac.uk

**Abstract**

*This project focuses on the development of a private DCT-based watermarking algorithm which strikes a balance between watermark robustness and imperceptibility of changes in the host image through variability of the watermark strength, and the complexity of carrying out such development. An overview of previous work done on the task of image watermarking is presented. Experimentation on the developed model is done in MATLAB and performance against JPEG compression, gaussian blurring and gaussian noise is gauged using PSNR and watermark recognisability as metrics. The watermark embedded using a low strength value is found to be resistant to light JPEG compression, relatively resistant to gaussian blurring and weak to gaussian noise. The watermark embedded using a high strength value is found to be resistant to most attacks, with the exception of a large degree of gaussian noise. The results of the experimentation are discussed and potential improvements for the model are proposed.*

## 1 Introduction

The aim of this project is to carry out research in the field of frequency-based image watermarking (the process of embedding information into an image, usually with the intent of proving copyright ownership [1]) through experimentation.

There are two major approaches to image watermarking: spatial-based image watermarking and frequency-based image watermarking.

**Spatial-based image watermarking** involves embedding information directly into the pixel values of an image, called the host image. This is typically done by extracting the image's bit planes (groups of bits in specific positions for each pixel representation) and modifying the bit planes with low significance such that the watermark is encoded into the image with little discrepancy between the original host image and its watermarked counterpart [1].

**Frequency-based image watermarking** involves converting an image into a frequency domain based representation and encoding the watermark into components of the host's frequency domain in such a way that the watermark is robust against attacks (examples of these will be shown in Section 4) while leaving the host image relatively unaffected [1].

This project will focus on the performance of a single approach to frequency-based image watermarking using DCT (Discrete Cosine Transform). It will show the complexity of finding a private/non-blind watermarking algorithm (a watermarking algorithm where the original, unwatermarked image is used to decode the watermarked image) which strikes a balance between watermark robustness and imperceptibility of changes in the host image.

## 2 Background

Given the exponentially growing amount of images being uploaded to the internet every year, it is important, now more than ever, for robust image watermarking techniques to be developed so that photographers, videographers, illustrators and artists alike are able to reliably claim and prove ownership of the media they publish. The following are brief overviews of examples of research papers on spatial-based and frequency-based image watermarking:

### 2.1 Spatial-based image watermarking

1. R. B. Wolfgang and E. J. Delp, "A watermark for digital images" [2]

   + Proposes spatial domain based algorithms which can accommodate for JPEG compression.

2. I. Pitas, "A method for signature casting on digital images" [3]

   + Watermarks are robust against subsampling and JPEG compression with ratios up to 4:1.

3. N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain" [4]

   + Watermarks are robust against JPEG compression and lowpass filtering. Presents variations immune to geometric transformations.

4. M. Schneider and S.-F. Chang, "A robust content based digital signature for image authentication" [5]

   + Utilises a signature system created based on the host image which allows for JPEG compression but prevents other types of manipulation.

5. M. M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification" [6]

   + Utilisation of watermark stamping allows the visual localisation of regions of the host image which have been altered after extraction.

The downsides to the use of spatial-based image watermarking techniques are that they are susceptible to geometric distortions, more easily detectable and can be incompatible with JPEG compression if implemented naively (e.g. simply modifiying the )

## 2.2 Frequency-based image watermarking

1. P. Tao and A. M. Eskicioglu, "A robust multiple watermarking scheme in the discrete wavelet transform domain" [7]

   + Watermark can be encoded into the LL band (lowest frequencies) as, contrary to most DWT-based watermarking approaches, embedding a binary pattern into this band does not cause watermark transparency to be lost and provides significant resistance against JPEG compression, blurring and gaussian noise.
   − Watermark insertions into high/low frequencies are robust against different types of attacks, however mutually exclusively.

2. A. Lumini and D. Maio, "A Wavelet-based Image Watermarking Scheme" [8]

   + Watermark strength is variable, adjusted based on the host image in order to maximise change imperceptibility.
   − Requires the original, unwatermarked host image to extract the watermark from the watermarked image (private method).

3. D. Kundur and D. Hatzinakos, "Digital watermarking using multiresolution wavelet decomposition" [9]

   + Multiresolutional wavelet-based embedding approach provides superior performance compared to previous techniques of the same class. Public method, i.e. does not require the original host image to extract the embedded watermark.

4. J. Ó Ruanaidh, W. Dowling, and F. Boland,"Watermarking digital images for copyright protection" [10]

   + Presents transform based methods which allow watermark bits to be placed adaptively (matching the characteristics of the host image).

5. E. Koch and J. Zhao,"Towards robust and hidden image copyright labeling" [11]

   + The developed RSPPMC method based on the JPEG model allow a copyright label code (sequence of low and high pulses) to be robust against common image-processing attacks.
   − Only utilises mid-range frequency bands to encode pulses.

# 3 Methodology

Figure 1: `lena512.png`, a 512 × 512 8-bit image, to be used as a host image

Figure 2: `unt64.png`, a 64 × 64 8-bit image, to be used as a watermark

## 3.1 Encoding the watermark into the host image and generating a watermarked image

The host image $H$ is resized such that its width $H_w$ and height $H_h$ are equal (i.e. the image is a square) and its dimensions are a power of 2. This is done by setting $H_w$ and $H_h$ equal to $2^{\text{ceil}(log_2(\max(H_w,H_h)))}$, e.g. if the host image has dimensions 100 × 240, it is resized such that is has dimensions 256 × 256.

The watermark $W$ is resized such that it has dimensions $\frac{H_w}{8} \times \frac{H_h}{8}$, i.e. $W_w = \frac{H_w}{8}$ and $W_h = \frac{H_h}{8}$.

Both $H$ and $W$ are divided into 8 × 8 blocks, and discrete cosine transform is performed on these blocks, taking the DCT of $H$ as $H_{DCT}$ and the DCT of $W$ as $W_{DCT}$.

$W_{DCT}$ is normalised by dividing all its values by the largest magnitude value in $W_{DCT}$.

$W_{DCT}$ is added to $H_{DCT}$ by allocating a single $8\times8$ block from $H_{DCT}$ to each pixel in $W_{DCT}$ ($H_{DCT}$ contains $\frac{H_w}{8}\times\frac{H_h}{8}$ $8\times8$ DCT blocks $= W_w\times W_h$ $8\times8$ DCT blocks, i.e. exactly enough $8\times8$ DCT blocks to allocate one to each pixel in $W_{DCT}$, hence the resizing done previously).

Frequency coefficients to be modified in each $8\times8$ DCT block in $H_{DCT}$ are highlighted in blue below:

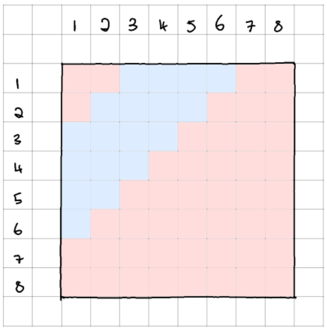


Figure 3: $8\times8$ DCT block coefficients to be modified

Coefficients in the low-mid frequency bands are selected. The lowest frequencies are avoided to prevent perceptible change to the host image, and the mid-high frequencies are avoided to increase the watermark's robustness to JPEG compression.

Each pixel, $W_{DCT_p}$, in $W_{DCT}$ is embedded into its corresponding $8\times8$ DCT block in $H_{DCT}$, $H_{DCT_P}$, by adding $\alpha\frac{W_{DCT_p}}{18}$ to each of the selected coefficients in $H_{DCT_P}$, where $\alpha$ is a weighting factor passed as an argument to specify the strength of the watermark [12]. $W_{DCT_p}$ is divided by 18 as 18 DCT coefficients are being modified (see Figure 3). This many DCT coefficients are selected for modification to create a "spreading out" effect – none of the coefficients are greatly affected (hence leaving the host image less drastically altered) and the watermark can theoretically be recovered by taking the summation of the difference between the 18 modified coefficients and their corresponding original coefficients.

The watermarked host image, $H_W$, is generated by performing inverse discrete cosine transform on each of the $8\times8$ DCT blocks in $H_{DCT}$. $H_W$ is resized to the original dimensions of the host image.

### 3.2 Decoding the watermarked image and extracting the watermark

The original host image $H$ and watermarked host image $H_W$ are resized such that their widths, $H_w$ and $H_{W_w}$, and heights, $H_h$ and $H_{W_h}$, are equal (i.e. the image is a square) and their dimensions are a power of 2. This is done by setting $H_w$, $H_{W_w}$, $H_h$ and $H_{W_h}$ equal to $2^{\text{ceil}(log_2(\max(H_w,H_h)))}$.

Both $H$ and $H_W$ are divided into $8\times8$ blocks, and discrete cosine transform is performed on these blocks, taking the DCT of $H$ as $H_{DCT}$ and the DCT of $H_W$ as $H_{W_{DCT}}$.

The extracted watermark DCT, $W'_{DCT}$, is obtained by setting each of its pixel values equal to the summation of the difference between the 18 modified coefficients in its corresponding $8\times8$ DCT blocks in $H_{W_{DCT}}$ and their corresponding original coefficients in $H_{DCT}$.

The extracted watermark, $W'$, is generated by performing inverse discrete cosine transform on each of the $8\times8$ DCT blocks in $W'_{DCT}$.

## 4 Results

Two separate experiments are carried out to analyse the difference in watermarked host image and extracted watermark quality with regards to the strength, $\alpha$, of the watermark embedded into the original host image. Specifically, in Experiment 1 a relatively low value of $\alpha$ (200) is used, and in Experiment 2 a relatively high value of $\alpha$ (500) is used. These values were found to be "relatively low" and "relatively high" through observation of the trade-off between quality and imperceptibility by trial and error.

PSNR (Peak signal-to-noise ratio) is used to measure the quality of both the watermarked host image and the extracted watermark.

$PSNR = 10\log_{10}\left(\frac{MAX^2}{MSE}\right)$, with $MAX$ being 255 as the experiments are performed on 8-bit images, and $MSE$ being the mean squared error between the pixel values of the original host image and the watermarked host image. Larger PSNR values correspond with higher image quality.

### 4.1 Experiment 1: $\alpha = 200$

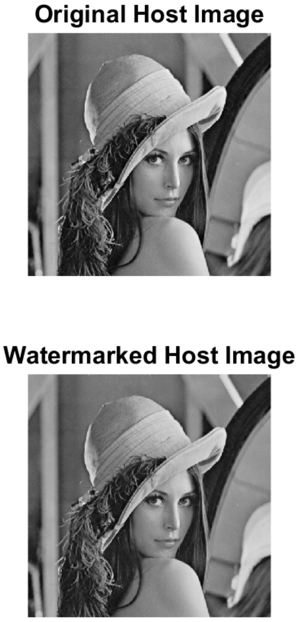


Figure 4: $\alpha = 200$ – Original Host Image v.s. Watermarked Host Image

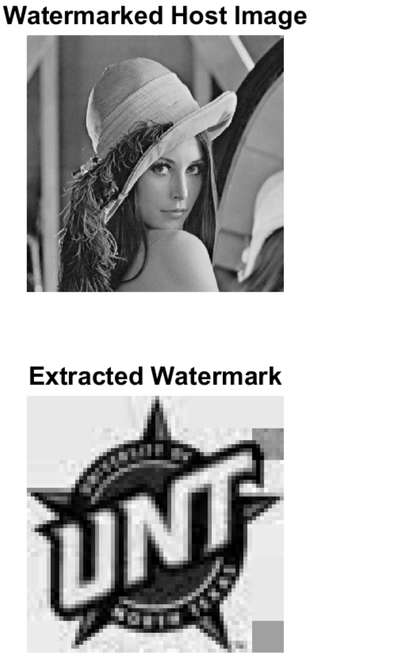


Figure 5: $\alpha = 200$ – No Manipulation

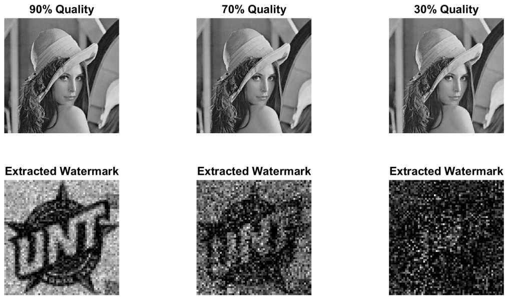


Figure 6: $\alpha = 200$ – JPEG Compression

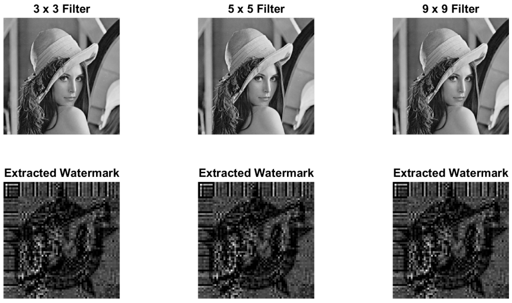


Figure 7: $\alpha = 200$ – Gaussian Blurring

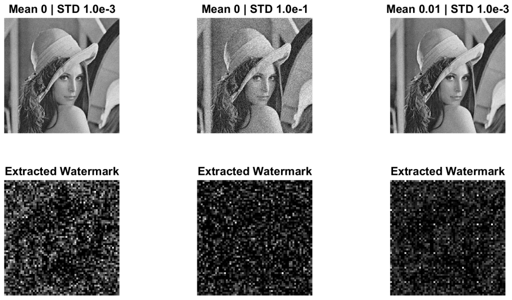
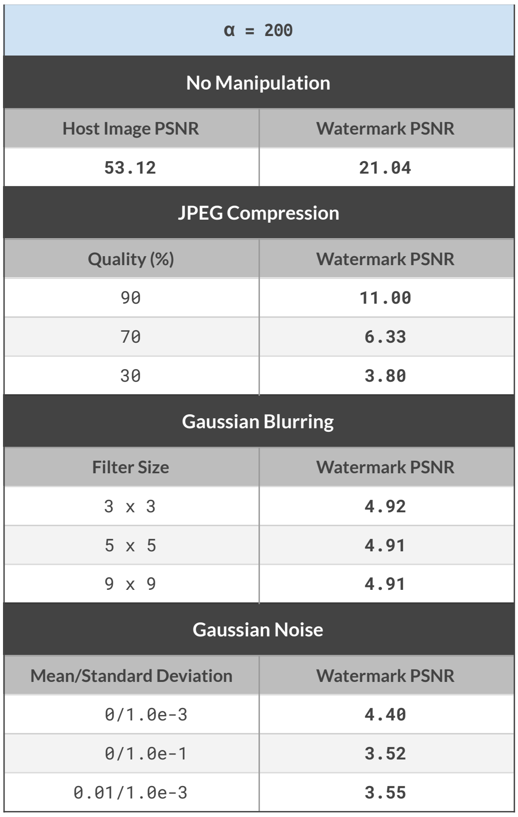


Figure 8: $\alpha = 200$ – Gaussian Noise



| α = 200 | |
|---|---|
| **No Manipulation** | |
| Host Image PSNR | Watermark PSNR |
| 53.12 | 21.04 |
| **JPEG Compression** | |
| Quality (%) | Watermark PSNR |
| 90 | 11.00 |
| 70 | 6.33 |
| 30 | 3.80 |
| **Gaussian Blurring** | |
| Filter Size | Watermark PSNR |
| 3 x 3 | 4.92 |
| 5 x 5 | 4.91 |
| 9 x 9 | 4.91 |
| **Gaussian Noise** | |
| Mean/Standard Deviation | Watermark PSNR |
| 0/1.0e-3 | 4.40 |
| 0/1.0e-1 | 3.52 |
| 0.01/1.0e-3 | 3.55 |

## 4.2 Experiment 2: $\alpha = 500$

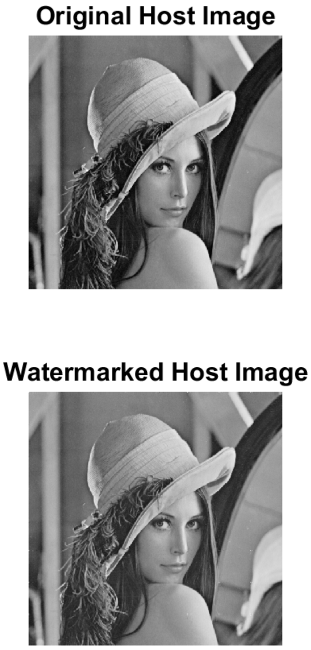
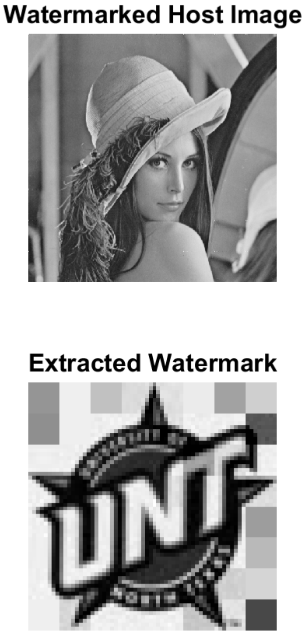


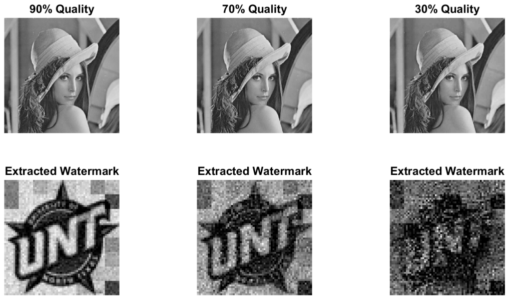
Figure 9: $\alpha = 500$ – Original Host Image v.s. Watermarked Host Image



Watermarked Host Image

Extracted Watermark

Figure 10: $\alpha = 500$ – No Manipulation



90% Quality

70% Quality

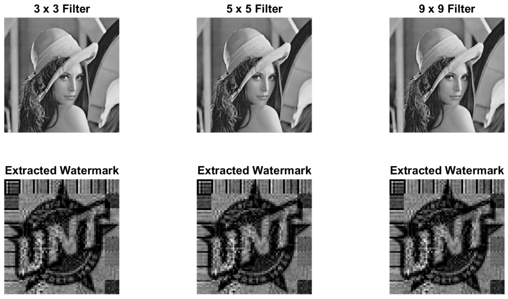
30% Quality

Extracted Watermark

Extracted Watermark

Extracted Watermark

Figure 11: $\alpha = 500$ – JPEG Compression



3 x 3 Filter

5 x 5 Filter

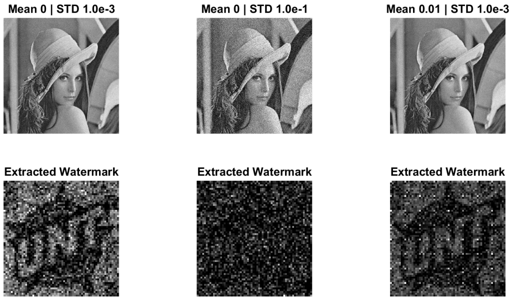
9 x 9 Filter

Extracted Watermark

Extracted Watermark

Extracted Watermark

Figure 12: $\alpha = 500$ – Gaussian Blurring



Mean 0 | STD 1.0e-3

Mean 0 | STD 1.0e-1

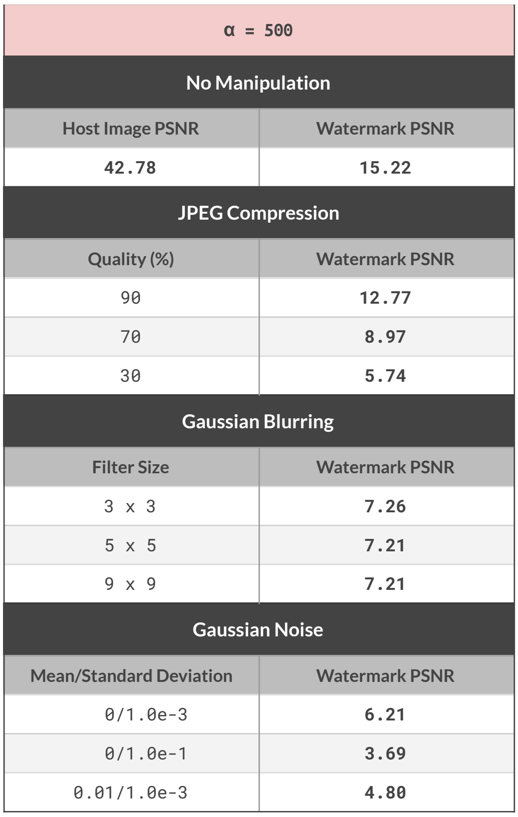
Mean 0.01 | STD 1.0e-3

Extracted Watermark

Extracted Watermark

Extracted Watermark

Figure 13: $\alpha = 500$ – Gaussian Noise



| α = 500 | |
|---|---|
| **No Manipulation** | |
| Host Image PSNR | Watermark PSNR |
| 42.78 | 15.22 |
| **JPEG Compression** | |
| Quality (%) | Watermark PSNR |
| 90 | 12.77 |
| 70 | 8.97 |
| 30 | 5.74 |
| **Gaussian Blurring** | |
| Filter Size | Watermark PSNR |
| 3 x 3 | 7.26 |
| 5 x 5 | 7.21 |
| 9 x 9 | 7.21 |
| **Gaussian Noise** | |
| Mean/Standard Deviation | Watermark PSNR |
| 0/1.0e-3 | 6.21 |
| 0/1.0e-1 | 3.69 |
| 0.01/1.0e-3 | 4.80 |

## 4.3 Analysis

### 4.3.1 Experiment 1

When the watermarked host image is not subject to any manipulation, it achieves a PSNR value of 53.12, indicative of a good level of watermark imperceptibility [13] [14] [15]. This is to be expected given the low watermark strength. Of note is the fact that despite the watermarked host image not being subject to any manipulation, the extracted watermark is not identical to the original watermark. This is due to the inverse discrete cosine transform function outputting values beyond the range of 0-255 and having to be normalised into 8-bit representation, resulting in a loss of information. The extracted watermark achieves a PSNR value of 21.04, and is visually clearly recognisable.

When the watermarked host image is subject to JPEG Compression with 90% quality, despite a noticeable degradation in quality of the extracted watermark (PSNR 11.00), it is still visually recognisable. This is to be expected given the DCT coefficients chosen to be modified are in the low-mid frequency bands (as per Figure 3) and are not greatly modified by JPEG compression at this level of quality [16]. At 70% quality, while being moderately recognisable, the extracted watermark suffers from a significant degradation in quality (PSNR 6.33). At 30% quality, the extracted watermark is practically unrecognisable. This is to be expected given the DCT coefficients in the low-mid frequency bands are greatly modified by JPEG compression at this level of quality.

When the watermarked host image is subject to Gaussian blurring, regardless of window size, the extracted watermark achieves a consistent PSNR value, averaging 4.91, and is visually clearly recognisable, with the limitation of noticeable artifacts appearing around the border of the watermark given the grid-structure approach of the encoding function and the zero-padding required to carry out gaussian blurring.

When the watermarked host image is subject to Gaussian noise, regardless of mean and standard deviation, the extracted watermark is hardly recognisable, achieving an average PSNR value of 3.82. This is to be expected as, unlike JPEG Compression, gaussian noise is applied entirely randomly and the DCT coefficients of the frequency transform watermarked host image are greatly modified regardless of the frequency band they are in.

#### 4.3.2 Experiment 2

When the watermarked host image is not subject to any manipulation, it achieves a PSNR value of 42.78, with noticeable white artifacts appearing in a grid-like manner. This is to be expected given the high watermark strength, with the white artifacts being a result of the large amount of flat regions in the watermark image, with their corresponding $8 \times 8$ DCT blocks being assigned large DCT coefficient values in the upper-left corner. The extracted watermark achieves a PSNR value of 15.22, achieving lower quality compared to the previous experiment, also a result of the inverse discrete cosine transform function outputting values beyond the range of 0-255 and having to be normalised into 8-bit representation, resulting in a loss of information.

When the watermarked host image is subject to JPEG Compression, watermark extraction provides remarkably better results when compared to Experiment 1, achieving an average PSNR value increase across the 3 levels of quality of 6.35. Of note is the fact that the extracted watermark at 30% quality, while suffering from a significant degradation in quality, is still visually recognisable unlike the watermark extracted in Experiment 1.

When the watermarked host image is subject to Gaussian blurring, regardless of window size, the extracted watermark achieves a consistent PSNR value, averaging 7.23, an increase of 2.32 compared to Experiment 1. The watermarks are visually clearly recognisable, yet still suffer from the limitation of noticeable artifacts appearing around the border of the watermark given the grid-structure approach of the encoding function and the zero-padding required to carry out gaussian blurring.

When the watermarked host image is subject to Gaussian noise, the first and third extracted watermarks, while achieving relativeving relatively low PSNR values (6.21, 4.80), are visually recognisable, unlike the watermarks extracted in Experiment 1. The second extracted watermark is still practically unrecognisable. It should be noted that the watermarked host image corresponding to the second extracted watermark has suffered from a significant loss of quality (Gaussian noise with mean 0, standard deviation 0.1), rendering it hardly comparable to the original image in terms of usability and opportunities for copyright infringement.

## 5 Conclusion

This project has shown that, given knowledge about the fundamentals of frequency-based image representation, discrete cosine transform and JPEG compression, an image watermarking model which is relatively simple to implement yet is robust to common image processing attacks can be developed.

The greatest strengths of the model developed in this project are its resistance to JPEG compression and its variability. In situations where the original host image is not expected to be subject to large amounts of compression/image processing attacks, a light watermark can be embedded with near-imperceptibility. If it is paramount that the watermark be resistant to stronger attacks, watermark imperceptibility can be sacrificed for greater robustness. The grid-like structure of the watermark encoding function also allows for a degree of localisation of changes in the host image.

Further development can be carried out to improve the model. Watermark imperceptibility can be improved through the generation of a pseudorandom sequence of indices in the DCT transform of the original host image wherein the watermark is added, as is demonstrated by C.-T. Hsu and J.-L. Wu in [14]. This would somewhat mitigate the grid-like artifacts that appear after the watermark is embedded. The model is also limited in that it is a private model, i.e. the original host image is needed to be able to extract the watermark from the watermarked host image, limiting the potential applications of the model.

## References

[1] V. M. Potdar, S. Han, and E. Chang, "A survey of digital image watermarking techniques," in *INDIN'05. 2005 3rd IEEE International Conference on Industrial Informatics, 2005.* IEEE, 2005, pp. 709–716.

[2] R. B. Wolfgang and E. J. Delp, "A watermark for digital images," in *Proceedings of 3rd IEEE International Conference on Image Processing*, vol. 3. IEEE, 1996, pp. 219–222.

[3] I. Pitas, "A method for signature casting on digital images," in *Proceedings of 3rd IEEE International Conference on Image Processing*, vol. 3. IEEE, 1996, pp. 215–218.

[4] N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain," *Signal processing*, vol. 66, no. 3, pp. 385–403, 1998.

[5] M. Schneider and S.-F. Chang, "A robust content based digital signature for image authentication," in *Proceedings of 3rd IEEE International Conference on Image Processing*, vol. 3. IEEE, 1996, pp. 227–230.

[6] M. M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification," in *Proceedings of international conference on image processing*, vol. 2. IEEE, 1997, pp. 680–683.

[7] P. Tao and A. M. Eskicioglu, "A robust multiple watermarking scheme in the discrete wavelet transform domain," in *Internet Multimedia Management Systems V*, vol. 5601. SPIE, 2004, pp. 133–144.

[8] A. Lumini and D. Maio, "A wavelet-based image watermarking scheme," in *Proceedings International Conference on Information Technology: Coding and Computing (Cat. No. PR00540).* IEEE, 2000, pp. 122–127.

[9] D. Kundur and D. Hatzinakos, "Digital watermarking using multiresolution wavelet decomposition," in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181)*, vol. 5. IEEE, 1998, pp. 2969–2972.

[10] J. Ó Ruanaidh, W. Dowling, and F. Boland, "Watermarking digital images for copyright protection," *IEE PROCEEDINGS VISION IMAGE AND SIGNAL PROCESSING*, vol. 143, pp. 250–256, 1996.

[11] E. Koch and J. Zhao, "Towards robust and hidden image copyright labeling," in *IEEE Workshop on Nonlinear Signal and Image Processing*, vol. 1. IEEE Neos Marmaras, Greece, 1995, pp. 123–132.

[12] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE transactions on image processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[13] X. Kang, J. Huang, Y. Q. Shi, and Y. Lin, "A dwt-dft composite watermarking scheme robust to both affine transform and jpeg compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 776–786, 2003.

[14] C.-T. Hsu and J.-L. Wu, "Hidden digital watermarks in images," *IEEE Transactions on Image Processing*, vol. 8, no. 1, pp. 58–68, 1999.

[15] C.-C. Lai and C.-C. Tsai, "Digital image watermarking using discrete wavelet transform and singular value decomposition," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 11, pp. 3060–3063, 2010.

[16] G. K. Wallace, "The jpeg still picture compression standard," *Communications of the ACM*, vol. 34, no. 4, pp. 30–44, 1991.