A NEW PSYCHO VISUAL MASKING MODEL FOR ROBUST IMAGE WATERMARKING

FERDA ERNAWAN

RESEARCH VOTE NO:

RDU 160364

FACULTY OF COMPUTER SYSTEMS &
SOFTWARE ENGINEERING,
UNIVERSITI MALAYSIA PAHANG

RDU160364

A NEW PSYCHO VISUAL MASKING MODEL FOR ROBUST IMAGE WATERMARKING

FERDA ERNAWAN

ZURIANI MUSTAFFA

MOHAMAD FADLI

LUHUR BAYUAJI

Report submitted in fulfilment of the requirements

Short Grant RDU160364, Universiti Malaysia Pahang

Faculty of Computer Systems & Software Engineering

UNIVERSITI MALAYSIA PAHANG

MARCH 2018

ACKNOWLEDGEMENTS

In preparing this report, I would like to express my full gratitude to Jabatan Penyelidikan & Inovasi, Universiti Malaysia Pahang for financial support of this research. I am also very thankful to my members, Dr. Mohamad Fadli, Dr. Zuriani Mustaffa and Dr. Luhur Bayuaji for their encouragement and continued support during research study. Without their continued support and interest, this report would not have been possible.

I am indebted to Universiti Malaysia Pahang, Malaysia for giving me a funding during the course of this research through the UMP Grant (RDU160364). Besides, I am also indebted to Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang (UMP) for providing me the assistance to endure this project. My sincere appreciation is also extended to all colleges at the Faculty of Computer Systems & Software Engineering, UMP who have assisted me at various occasions during this project.

Finally, I would like to express my personal gratitude to my parents and my wife for their continuous encouragement. Thank you and may Allah bless you all.

ABSTRACT

Demand for access and sharing of the digital image through the internet, the digital image can easily be copied, edited and reused in communication nowadays. Digital image watermarking is an approach to protect and manage the intellectual property of digital images. A natural embedding watermark based on the human eye properties can be utilized to effectively hide a watermark image. Human visual system typically ignores visual redundant part of color image. This research proposed psychovisual threshold in image watermarking and image compression. In this research, the principle of psycho acoustic model has referred to develop a psychovisual threshold for robust image watermarking. Psychovisual threshold is able to determine the redundant area of the image for watermark embedding. The redundant information can be used to embed copyright of digital image. This project develop a robust image watermarking based on the psychovisual threshold and spatially along the edge. The psychovisual threshold has also been used to assign bit allocation in image compression. It can be used to determine an optimal bit for each pixel. This research will also investigate the sensitivity of the changes in DCT coefficients to the error reconstruction in order to generate a set of bit allocation. The experimental results show that the proposed embedding watermark scheme produces a strong resistance under different types of attack. The experimental result that watermarked image does not introduce visually imperceptible distortion to the host images. We also has applied psychovisual threshold to determine an optimal bits budget in image compression, the experimental results show that the proposed bit allocation can improve the image compression performance in terms of image quality and compression rate.

TABLE OF CONTENTS

CONTENT **Page**

UMP

LIST OF TABLES

LIST OF PUBLICATIONS

1. ISI JOURNAL

Ferda Ernawan, Muhammad Nomani Kabir, Jasni Mohamad Zain. (2017). Bit allocation strategy based on Psychovisual threshold in image compression. Multimedia Tools and Applications, pp. 1-24 (Science Citation Index, Q2, IF: 1.530)

2. SCOPUS JOURNAL

Ferda Ernawan, Mritha Ramalingam, Ali Safa Sadiq, Zuriani Mustaffa. (2017). An Improved Imperceptibility and Robustness of 4x4 DCT-SVD Image Watermarking Using Modified Entropy. Journal of Telecommunication, Electronic and Computer Engineering, vol. 9, no. 2-7, pp. 111-116.

Ferda Ernawan. (2016). Robust Image Watermarking Based on Psychovisual Threshold. Journal of ICT Research and Applications, vol. 10, no. 3, pp. 228-242.

3. INDEX PROCEEDING

Ferda Ernawan, Muhammad Nomani Kabir, Mohamad Fadli, Zuriani Mustaffa. (2016). Block-based Tchebichef image watermarking scheme using psychovisual threshold. International Conference on Science and Technology-Computer (ICST), 27-28 Oct. 2016, pp. 6-10.

4. NON-INDEX PROCEEDING

Ferda Ernawan. (2016). A New Psychovisual Threshold Technique in Image Processing Applications. In: Proceeding of International Competition and Exhibition on Computing Innovation 2016, 6-8 December 2016, Universiti Malaysia Pahang. pp. 9-20. ISBN 978-967-2054-04-7.

LIST OF ABBREVIATIONS

ARE	Absolute Reconstruction Error
CR	Data Compression Ratio
DC	Direct Component
DCT	Discrete Cosine Transform
DPCM	Differential Pulse-Code Modulation
DWT	Discrete Wavelet Transform
JND	Just Noticeable Difference
JPEG	Joint Photographic Experts Group
MSE	Mean Square Error
PSNR	Peak Signal to Noise Ratio
SSIM	Structural Similarity Index
TMT	Tchebichef Moment Transform
YUV / YC_bC_r	Luminance (Y), Chrominance – Blue (U / C_b), Red (V / C_r)

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Nowadays, multimedia data like images are easy to manipulate digital image contents. The intellectual properties of digital images are facing serious challenges such as piracy, illegal redistribution, forgery and theft (Halder et al., 2010). These problems make digital image watermarking to be important to protect against unauthorized duplication of the digital image. Image watermarking is to embed a watermark without degrading of the perceptual image quality and making it difficult to be removed simultaneously (Cui and Li, 2011). Image watermarking is able to comply with imperceptibility, robustness and security. The watermark image allows for embedding and extracting only by the authorized user. In modern digital image watermarking, the watermark embedding process exploits the characteristic of the Human Visual System (HVS). The embedding process inserts more bits of the watermark image at less sensitive of the human visual system (Yang et al., 2009), especially in highly textured area and significantly changing region of an image or the image edges. It means that the image pixels are highly correlated.

Multimedia communication required high bandwidth and data transfer rate to transfer data information. In order to fasten up the data transfer rate, image compression is unavoidable. Image compression is used to reduce the storage and transmission. Image compression is designed to reduce the data storage and maintain the image quality. An image contains subjective redundancy which is determined by the human visual system. The visual image redundancy relatively is not perceived in perception of the human visual system. The human visual system perceives the significant dominant image signals (Pappas et. al., 2010). The human eye presents some tolerance to the distortion, depending upon the image content and viewing conditions. The human visual system does not perceive the difference between adjacent image pixels. The characteristic of the human visual system has been applied in image watermarking applications (Niu et al., 2011; Parthasarathy, 2007) and image compression. The image watermarking schemes are typically designed based on a visual model to colour stimuli (Barni et al., 2012; Tsui et al., 2008) and derived from image compression to increase its robustness (Lu et al., 2000). The embedding watermarks based on HVS properties are able to improve the robustness while still maintaining the watermark imperceptibility. The watermark scheme (Kutter and Wingkler, 2002) utilizes the spatial masking principle of HVS to improve the watermark strength. The human visual system sensitivity is utilized to improve the robustness watermark for authentication (Wei and Ngan, 2009) and protection (Lu and Liao, 2001). Various watermarking applications exploit a model of the human visual system (HVS). A major element of HVS models describes the masking-effect, which is typically parameterized by psycho-visual experiments.

This project ideally ambitious nature aims to implement the biological structure of the human visual system in the computer vision. Alternatively, the practical approach aims to discover how to apply human visual sensitivity perception in the images. The principle of audio masking technique shall be applied to develop a new psycho visual masking in watermarking. In this project, a specific location for embedding watermark based on the psychovisual masking shall be investigated where a just-noticeable distortion (JND) or minimally noticeable distortion (MND) profile is employed to quantify the perceptual redundancy. The JND provides each signal being masked with a visibility threshold of distortion, below which reconstruction errors are rendered visually unnoticed. Based on a perceptual model that incorporates the threshold sensitivities due to background luminance and texture masking effect, the JND profile is estimated from analyzing local properties of image signals. According to the sensitivity of human visual perception to spatial frequencies, the full-band JND/MND is decomposed into components of different frequency sub-bands (Chou and Li, 1995). A new psycho visual masking shall be utilized to determine bits budget, location and strength of the watermark.

This project mainly discusses psychovisual threshold in digital watermarking and image compression. The proposed psychovisual threshold can be utilized for hiding data in digital watermarking. The psychovisual threshold is also used to measure an optimal bit allocation in image compression.

1.2 PROBLEM STATEMENTS

The choice of a specific masking model for an image processing application should be motivated by the model performance in terms of prediction quality and stability. Unfortunately, an explicit masking model on natural scenery images is not well established yet. The masking phenomena are of great relevance for image watermarking. Before these masking effects can be exploited in image watermarking applications, they need to be measured and modeled (Nadenau et al., 2002) Visual masking has always been a phenomenon deemed worthy of study in its own right (Breitmeyer, 2007). In general, a psychovisual has been designed based on our understanding of brain theory and neuroscience (Zhai et al., 2012). Unlike psycho visual, the field of psychoacoustics has made significant progress toward characterizing human auditory perception and particularly the time-frequency analysis capabilities of the inner ear (Painter and Spanias, 2000). The psycho visual masking in the watermarking has not been understood and developed to the level of psychoacoustic model.

1.3 RESEARCH QUESTION

There are several research questions that can be consider base on the previous part:

1. What is the best digital watermarking to protect the copyright?
2. How to adopt the psychovisual threshold in digital watermarking and improve the robustness of watermark image?
3. How to adopt the psychovisual threshold in image compression and improve the quality of reconstructed image at minimum bit rate in image compression?

There are many digital watermarking techniques to protect the copyright and they provide high imperceptibility and robustness. First research question is asking about the current digital watermarking techniques for copyright protection. Second question discusses the way how to apply the psychovisual threshold in digital watermarking. Third question discusses how to adopt the psychovisual threshold in image compression in order to improve the quality of reconstructed image at minimum bit rate.

1.4 RESEARCH OBJECTIVE

This study embarks on the following objectives:

1. To review the existing digital watermarking techniques for copyright protection.
2. To design and experiment on the psychovisual threshold in the block-based digital watermarking and transformed domain.
3. To implement the psychovisual threshold in image compression.
4. To evaluate the performance of the psychovisual threshold in the digital watermarking and image compression.

1.5 RESEARCH SCOPE

The scope for this research is listed below:

1. This experiment uses 80 sample images (40 graphic images and 40 natural images, 24-bit RGB image of size 512×512 pixels).
2. A watermark image size is a binary watermark image of size 32×32 pixels.
3. The imperceptibility of watermarked image is measured by Structural SIMilarity (SSIM) index, Absolute Reconstruction Error (ARE), Peak-Signal-Noise-Ratio (PSNR).
4. The robustness of watermark image is evaluated by Normalize cross Correlation (NC), Bit Error Rate (BER)

5. The proposed watermarking scheme is tested by different types of attack, e.g. JPEG compression, Gaussian white noise, salt and pepper noise, median filter, sharpening, cropping, etc.

1.6 SIGNIFICANT OF STUDY

A psychovisual masking development for image watermarking shall produces the best locations and an optimal bits budget for embedding watermark without introduce visually imperceptible distortion to the original images. The proposed watermarking scheme has potential to produce better imperceptibility and robustness.

The psychovisual threshold in image compression avoids blocking effect on the image reconstruction. The proposed scheme shall provide higher quality of the image at minimum bit rate than existing benchmark image compression technique.

1.7 RESEARCH ORGANIZATION

Chapter 1 provides an introduction of the challenges the psychovisual masking in image watermarking. This chapter covers the background of the psychovisual masking in digital watermarking and image compression. This chapter also discusses the problem area of this project, research objectives, research scope, contribution of this study and report structure.

Chapter 2 discusses block-based transform digital watermarking using the psychovisual threshold. This chapter includes human visual characteristic in order to determine the embedding regions. This chapter also discusses the feasibility study on the embedding and extracting based on the psychovisual threshold.

Chapter 3 discusses on the bit allocation strategy based on the psychovisual threshold in image compression. It explains the useful of the psychovisual threshold in image compression. Psychovisual threshold is not only can be applied in image watermarking, it is also can be used to assign bit allocation in image compression. This chapter also describes a step-by-step procedure on the generation of bit allocation based on the psychovisual threshold in image compression.

Chapter 4 explains robust image watermarking based on the psychovisual threshold along the edge. This chapter also provides the details of the watermark embedding and extracting process along the edge based on the psychovisual threshold. Besides, Chapter 4 will show the watermark recovery under different types of attacks.

Chapter 5 presents an improved imperceptibility and robustness of the image watermarking on smaller block-size. This chapter analyzes the techniques on how to embed watermark based on the trade-off between imperceptibility and robustness. Then, it explores on the concept of the optimal threshold for watermark embedding in digital watermarking. This chapter also provides the experimental results and comparison to the existing benchmark in digital watermarking.

Chapter 6 concludes the overall research objectives which are stated in the first chapter. This chapter emphasizes on the research accomplishments, research contributions, limitations and future research that can be conducted as further enhancement of this study.

CHAPTER 2

Block-based Tchebichef Image Watermarking Scheme using Psychovisual Threshold

Ferda Ernawan, Muhammad Nomani Kabir, Mohamad Fadli, Zuriani Mustaffa
Faculty of Computer Systems & Software Engineering
Universiti Malaysia Pahang
Lebuhraya Tun Razak 26300 Gambang Kuantan, Pahang, Malaysia
ferda1902@gmail.com, ferda@ump.edu.my

***Abstract*—**Digital multimedia has drastically increased the production and distribution of digital data in the recent years. Unauthorized manipulation and ownership of digital image have become a serious issue. In this paper, we propose a watermarking scheme which uses block-based Tchebichef moments considering psychovisual threshold. The psychovisual threshold is used to prescribe the potential location of embedded watermark. The proposed watermarking scheme considers minimum modified entropy values to determine the embedded blocks. The lowest psychovisual error threshold on each selected block are chosen as the best location to insert the watermark image. Experimental results demonstrate that the embedding watermark into the lowest Tchebichef psychovisual threshold can produce a good level of imperceptibility. The watermark recovery is strongly robust against JPEG compression.

Keywords*—*watermarking; psychovisual threshold; tchebichef moments; robustness; imperceptibility

I. INTRODUCTION

The wide availability of the internet has drastically increased the distribution multimedia data. This leads to the chances of illegal and unauthorized manipulation of digital image. To maintain full security for digital image ownership while depending on the internet resources has become an issue. To solve this issue, digital image watermarking can be used to protect a copyright of the digital image. The image watermarking embeds an authorized mark information in the digital image to protect the ownership of digital image. The ideal watermarking is to embed such amount of information into host image that provides tradeoff between robustness and imperceptibility by human visual system.

In order to achieve the tradeoff between robustness and imperceptibility in the watermarked images, researchers investigated watermarking scheme that adopted human visual characteristics. Researchers investigated the human visual system characteristics from the entropy in embedding watermark [1]-[3]. The watermark embedding based on edge entropy of image can improve the watermark imperceptibility and robustness [4]. The entropy of image presents less sensitive image information to the human visual system [1]. The entropy is utilized to identify embedding region which produces redundancy of the image pixels to human visual system [5].

In this paper, we investigate the robustness and imperceptibility of watermark insertion based on psychovisual threshold with a modified entropy. The block regions of watermark insertion are determined by modified entropy of image pixels. The psychovisual threshold is utilized to identify the watermark embedding location on each selected block region. The embedding watermark based on psychovisual threshold is able to produce a good level of imperceptibility and the robustness.

Previous works have been carried out to investigate the effect of psychovisual threshold in image compression [6]-[12]. The psychovisual threshold prescribes the quantization tables in image compression. Furthermore, it plays the role in that quantization tables for reducing the redundancy of image signals on each diagonal orders. Referring to psychovisual threshold, the lower threshold has no significant effect to the reconstruction error [13]. The first author assumes that the lower bound of psychovisual threshold is a suitable location to insert the watermark information. The watermark insertion into the less bound of Tchebichef psychovisual threshold has not been investigated by other researchers.

The paper is organized as follows. A short description on Tchebichef psychovisual is provided in Section 2. Section 3 presents a watermarking scheme. Section 4 shows experimental results of watermarking scheme. Finally, section 5 concludes this paper.

II. TCHEBICHEF PSYCHOVISUAL

Tchebichef psychovisual system can be used for image compression. The psychovisual effect is measured by average absolute reconstruction errors on each diagonal order. The lowest reconstruction errors indicate high imperceptibility of image pixels [13]. So it is necessary to investigate the imperceptibility of embedding watermark using Tchebichef psychovisual threshold. Tchebichef moment has been widely used in image processing using matrices [14]. The Tchebichef moment of order $m + n$ is defined:

$$T_{mn} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \frac{t_m(x)}{\rho(m,M)} f(x,y) \frac{t_n(y)}{\rho(n,N)} \tag{1}$$

where $m, n = 0, 1, 2, ..., N\text{-}1$ and $f(x, y)$ implies the pixel value of the original image. The set $\{t_n(x)\}$ can be evaluated by the recursive relation in [14]. The original image size is $M\times N$ of a grayscale image.

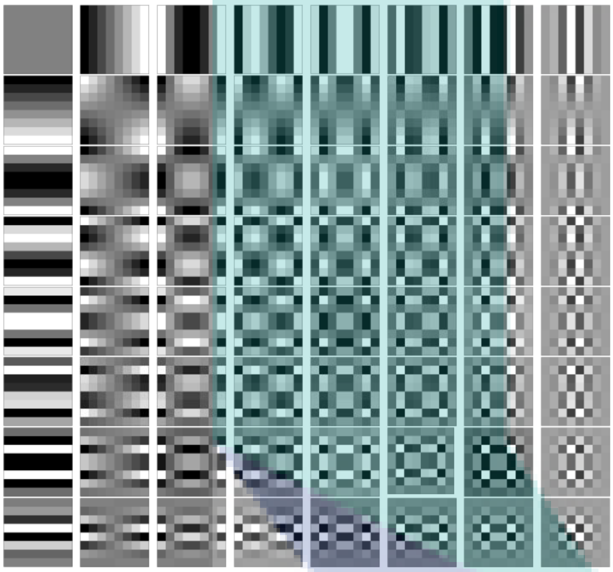


Fig. 1. 8×8 Tchebichef moments

The inverse of Tchebichef moments is given as follows:

$$\tilde{f}(x,y) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \frac{t_m(x)}{\rho(m,M)} T_{mn} \frac{t_n(y)}{\rho(n,N)} \tag{2}$$

where M,N represents the moments order being used and $\tilde{f}(x,y)$ is an reconstructed image. The kernel of orthogonal Tchebichef polynomials can be computed by:

$$K_x = \frac{t_m(x)}{\rho(m,M)}, K_y = \frac{t_n(y)}{\rho(n,N)} \tag{3}$$

The Tchebichef moments of the kernel matrix is defined as:

$$T = K^T F K \tag{4}$$

where F is the image block matrix denoting the intensity of the image pixels. The inverse of Tchebichef moments can be reconstructed by:

$$G = KTK^T \tag{5}$$

where G represents an image reconstruction matrix.

III. WATERMARKING SCHEME

Watermarking scheme includes embedding process and watermark extraction. In order to improve the robustness and imperceptibility of watermarked images, the suitable locations for watermark embedding in Tchebichef moments are investigated. The Tchebichef psychovisual threshold is utilized to determine the location of watermark embedding. Previous research [7] has shown that the psychovisual threshold improved the image compression performance. Based on the Fig. 2, the Tchebichef psychovisual threshold indicates that the second frequency order produces significantly less average reconstruction error. Therefore, second frequency order is the suitable location to embed the watermark in Tchebichef moment watermarking scheme. The embedding watermark into second frequency order of Tchebichef moments has a potential to achieve high imperceptibility and robustness.

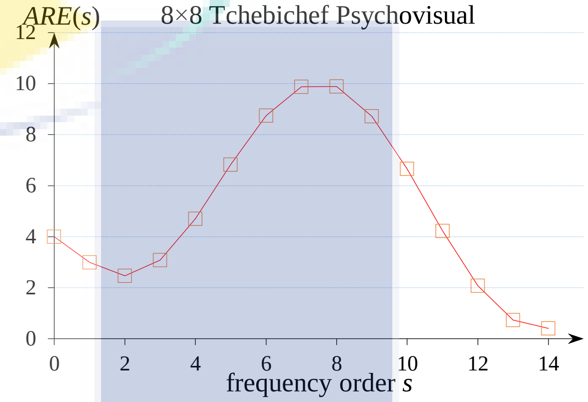


Fig. 2. The psychovisual error threshold on 8×8 Tchebichef moments

A. Embedding Process

Step 1: The host image is divided into 8×8 pixels of non-overlapping blocks.

Step 2: Calculate the modified entropy for each image block. The modified entropy is defined by:

$$E = -\sum_{i=1}^{n} p_i \log_2(p_i) + p_i \exp^{1-p_i} \tag{6}$$

p_i denotes the occurrence probability of i-th pixel with $0 \le p_i \le 1$ and $\sum_{i-1}^{n} p_i = 1$. The values obtained of modified entropy are sorted in ascending order. Choose the blocks that have the lowest value of modified entropy as the block for embedding watermark.

Step 3: The selected blocks are transformed by two-dimensional Tchebichef moment transform.

Step 4: On each selected block, the implementation of frequency order by $C_{(S\times S)}$ with S=8 is given as follows:

$$C = \begin{bmatrix} C(0,0) & C(0,1) & \cdots & C(0,S-1) \\ C(1,0) & C(1,1) & \cdots & C(1,S-1) \\ C(2,0) & C(2,1) & \cdots & C(2,S-1) \\ \vdots & \vdots & \ddots & \vdots \\ C(S-1,0) & C(S-1,1) & \cdots & C(S-1,S-1) \end{bmatrix} \tag{7}$$

The second order of image signals is located by $C_{(2,0)}$, $C_{(1,1)}$ and $C_{(0,2)}$ coordinates. The watermark is embedded by adding the moment value with watermark weight. The watermark weight that will be embedded into the host image is subjected to second diagonal order of TMT quantization values [15], W_Q=5. The watermark weight to be embedded depends upon the threshold value T as follows:

$$Q(i) = T \cdot W_Q \tag{8}$$

If the embedded watermark bit is 1, then the moment is added by Q. Otherwise, the moment is subtracted by Q.

Step 5: Obtain random numbers using mersenne twister technique with a secret key. The locations of embedding watermark into second order of the matrix moment are randomized based on a private key.

Step 6: Inverse of Tchebichef moment to the selected blocks is computed to produce the watermarked image.

B. *Extraction Process*

Step 1: Watermarked image is divided by 8×8 pixels of non-overlapping blocks.

Step 2: Ascending values of modified entropy are used to identify the selected block of embedded watermark.

Step 3: Apply Tchbichef moments transform to obtain moment frequency for each selected block.

Step 4: Obtain pseudorandom numbers using the same private key. These random numbers are utilized to determine the location of watermark in the second order $C_{(2,0)}$, $C_{(1,1)}$ and $C_{(0,2)}$ coordinates of the matrix moment.

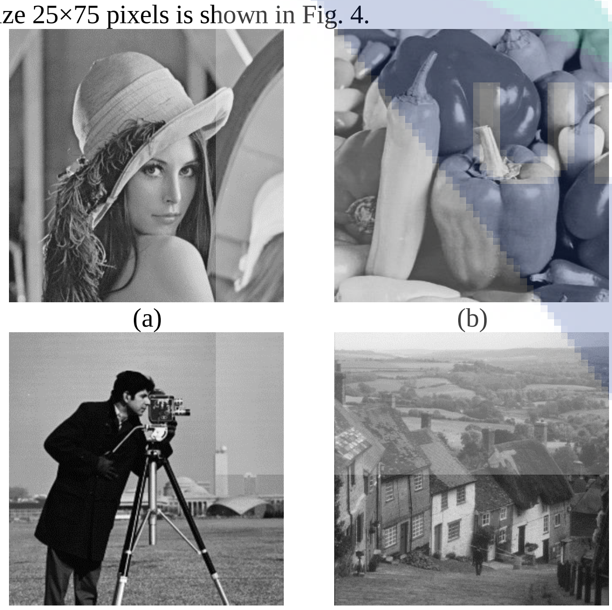
Step 5: Calculate the correlation coefficient using

$$\rho = X \cdot X^* \tag{9}$$

where X denotes the watermarked image and X^* represents the extracted sequence of watermark. If the results of the correlation is larger than a threshold, we determine the watermark is 1, otherwise the watermark is 0.

IV. Experimental Results

The imperceptibility of watermarked image is evaluated by Absolute Reconstruction Error (ARE) and Peak Signal to Noise Ratio (PSNR). Robustness of watermark recovery is estimated by Normalized Correlation (NC). In this experiment, all grayscale host images are obtained from CVG image database [16]. Four grayscale images are selected to demonstrate the proposed watermarking scheme (i.e. "Lena", "Pepper", "Cameraman" and "Gold Hill" with 512×512 pixels as depicted in Fig. 3). The binary watermark W "UMP" with size 25×75 pixels is shown in Fig. 4.



(a) (b) (c) (d)

Fig. 3. Original Lena (a), Pepper (b), Cameraman (c) and Gold hill (d) images.

Fig. 4. Original watermark image of 25×75 pixels.

The watermark image is embedded into four host images with threshold values 0.25, 0.5, 0.75 and 1 of the weighted watermark image. Table I and Table II list the PSNR and Full Error values of all watermarked images which indicate the imperceptibility between different thresholds of weighted watermark images. Table III shows the comparison of robustness of the proposed method. For proving the characteristics of robustness of our watermarking scheme, the watermarked images were tested using different attacks: noise, denoising, compression, image processing and geometrical as shown in Fig. 5. The visual watermark recovery is depicted in Fig. 6.

TABLE I. PSNR UNDER DIFFERENT THRESHOLD VALUES

Threshold	Lena	Pepper	Cameraman	Gold hill
1	45.645	45.600	45.602	45.604
0.75	48.144	48.099	48.101	48.102
0.5	51.666	51.621	51.623	51.624
0.25	57.686	57.642	57.644	57.645

TABLE II. FULL ERROR UNDER DIFFERENT THRESHOLD VALUES

Threshold	Lena	Pepper	Cameraman	Gold hill
1	0.586	0.590	0.598	0.588
0.75	0.439	0.443	0.449	0.441
0.5	0.293	0.295	0.299	0.294
0.25	0.146	0.147	0.149	0.147

The results in Table I reveal that the watermarked images produce high PSNR values, because only second frequency order of Tchebichef moments are modified. Since the watermark is embedded into second frequency order of Tchebichef moments, the watermarked image quality is not generally affected and relatively high imperceptibility is obtained. The results of numerical simulation in Table II indicate that the larger threshold of weighted watermark image produces lower quality of the watermarked image. According to the previous works [2][3][5], the second frequency order of Tchebichef psychovisual threshold produces lower absolute reconstruction error than other frequency orders. Therefore, we assume that the modified coefficient values on the second frequency order of Tchebichef moments do not significantly affect the quality of the watermarked image. We find that embedding in the second frequency order of Tchebichef moments based on psychovisual threshold can achieve robustness of watermark. The selected block based on a modified entropy can provide high imperceptibility. The present finding also supports the author study which concluded that the embedding watermark image into the second frequency order of Tchebichef moments can achieve higher robustness against image compression.

V. Conclussion

This paper proposes Tchebichef watermarking scheme with psychovisual threshold to protect digital watermark image. The selected embedding blocks are considered using a modified entropy. The watermark image is embedded into second frequency order of Tchebichef moments on each selected region block of the host image. Watermarked images are tested by different attacks which include cropping, noise, denoising, compression and image processing. Our proposed watermarking scheme achieves the robustness in watermark recovery against JPEG compression. The resisting watermark against rotation will be further investigated in a future research.

Acknowledgement

The authors would like to express very special thanks to Universiti Malaysia Pahang, Malaysia for providing financial support of this research project by UMP Research Grant Scheme (RDU160364).

References

[1] S.P. Maity, M.K. Kundu, "DHT Domain Digital Watermarking with Low Loss in Image Informations," International Journal of Electronics and Communications, vol. 64, no. 3, pp. 243-257, 2010.

[2] V.P. Reddy, S. Varadarajan, "Human Visual System Sentient Imperceptible and Efficient Wavelet-Based Watermarking Scheme for Copyright Protection of Digital Images," International Journal of Computer Science and Network Security, vol. 9, no. 4, pp. 255–264, 2009.

[3] Y.T. Lin, C.Y. Huang, G. Lee, "Rotation, Scaling, and Translation Resilient Watermarking for Images," IET Image Processing, vol. 5, no. 4, pp. 328–340, 2011.

[4] C.C. Lai, "An Improved SVD-based Watermarking Scheme Using Human Visual Characteristics," Optics Communications, vol. 284, no. 4, pp. 938–944, 2011.

[5] N.M. Makbol, B.E. Khoo and T.H. Rassem, "Block-Based Discrete Wavelet Transform-Singular Value Decomposition Image Watermarking Scheme Using Human Visual System Characteristics," IET Image Processing, vol. 10, no. 1, pp. 034–052, 2016.

[6] F. Ernawan, S.H. Nugraini, "The Optimal Quantization Matrices for JPEG Image Compression From Psychovisual Threshold," Journal of Theoretical and Applied Information Technology, vol. 70, no. 3, pp. 566-572, 2014.

[7] F. Ernawan, N.A. Abu, N. Suryana, "An Optimal Tchebichef Moment Quantization Using Psychovisual Threshold for Image Compression," Advanced Science Letters, vol. 20, no. 1, pp. 70-74, 2014.

[8] F. Ernawan, N.A. Abu, N. Suryana, "Adaptive Tchebichef Moment Transform Image Compression Using Psychovisual Model," Journal of Computer Science, vol. 9, no. 6, pp. 716-725, 2013.

[9] F. Ernawan, N.A. Abu, N. Suryana, "An Adaptive JPEG Image Compression Using Psychovisual Model," Advanced Science Letters, vol. 20, no. 1, pp. 26-31, 2014.

[10] N.A. Abu, F. Ernawan, "Psychovisual Threshold on Large Tchebichef Moment for Image Compression," Applied Mathematical Sciences, vol. 8, no. 140, pp. 6951-6961, 2014.

[11] N.A. Abu, F. Ernawan, "A Novel Psychovisual Threshold on Large DCT for Image Compression," The Scientific World Journal, vol. 2015, no. 2015, pp. 001-011, 2015.

[12] N.A. Abu, F. Ernawan and N. Suryana, "A Generic Psychovisual Error Threshold for the quantization table generation on JPEG Image Compression," 9th International Colloquium on Signal Processing and its Applications, pp. 39-43, 2013.

[13] N.A. Abu, F. Ernawan, N. Suryana, S. Sahib, "Image Watermarking Using Psychovisual Threshold over the Edge," Information and Communication Technology, vol. 7804, pp. 519-527, 2013.

[14] N.A. Abu, F. Ernawan and S. Sahib, "Psychovisual Model on Discrete Orthonormal Transform," International Conference on Mathematical Sciences and Statistics, pp.309-314, 2013.

[15] F. Ernawan, N.A. Abu and N. Suryana, "TMT Quantization Table Generation Based on Psychovisual Threshold for Image Compression," International Conference of Information and Communication Technology, pp. 202-207, 2013.

[16] University of Granada. Computer Vision Group. CVG-UGR Image Database. [2012-10-22]. http://decsai.ugr.es/cvg/dbimagenes/c512.php

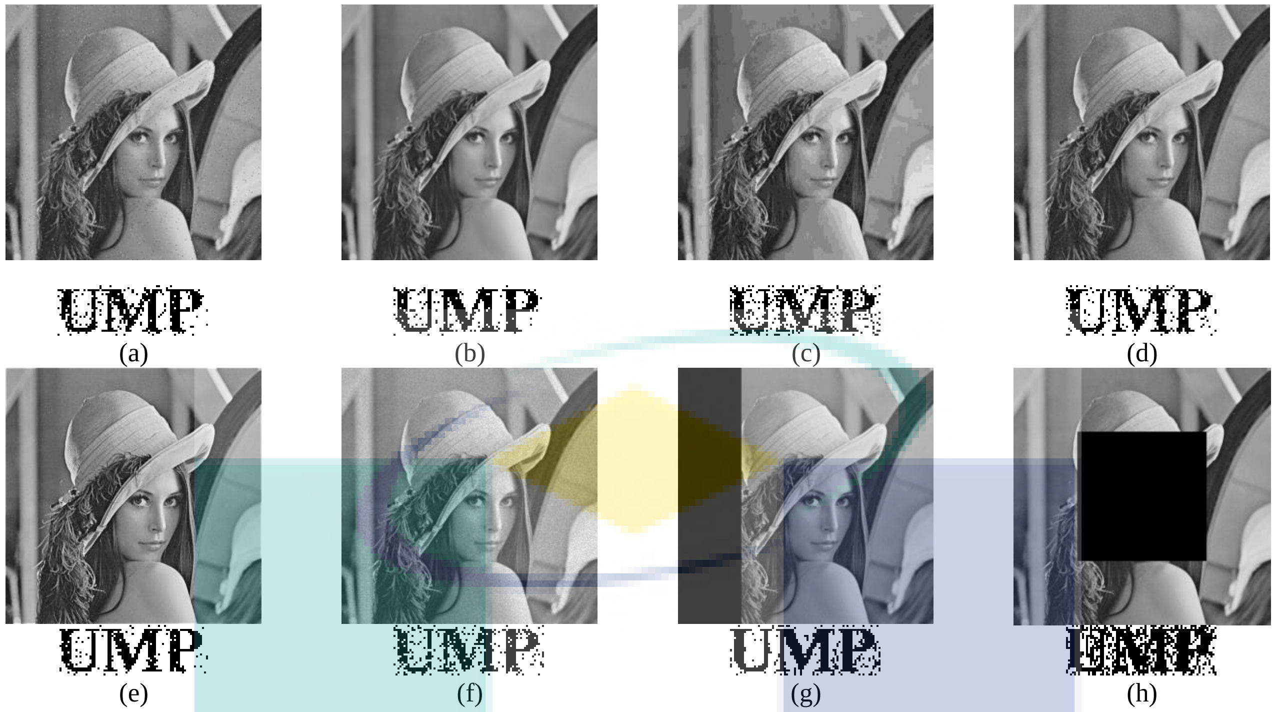


Fig. 5. Watermark recovery from (a) Salt and Peppers noise (0.01), NC=0.931; (b) Median filtering (3×3), NC=0.953; (c) JPEG compression (Q=30), NC=0.885; (d) Gaussian Noise (0.001), NC=0.926; (e) Sharpening (2×2), NC=0.958; (f) Speckle noise (0.01), NC=0.906; (g) Cropping column off (25%), NC=0.925; (h) Centred Cropping (25%), NC=0.760.

TABLE III. NC AFTER DIFFERENT TYPES OF ATTACKS ON WATERMARKED IMAGE UNDER DIFFERENT THRESHOLD VALUES

Attacks	*T*=0.25		*T*=0.5		*T*=0.75		*T*=1	
	Lena	Pepper	Lena	Pepper	Lena	Pepper	Lena	Pepper
JPEG compression Q=10	0.550	0.474	0.545	0.489	0.570	0.519	0.600	0.545
JPEG compression Q=30	0.841	0.815	0.851	0.848	0.885	0.869	0.913	0.899
JPEG compression Q=50	0.912	0.899	0.922	0.918	0.936	0.929	0.951	0.950
JPEG compression Q=70	0.948	0.950	0.961	0.965	0.968	0.969	0.975	0.979
Gaussian low pass filter [3 3]	0.957	0.947	0.954	0.938	0.955	0.944	0.948	0.944
Gaussian Noise 0.001	0.897	0.898	0.915	0.901	0.926	0.926	0.946	0.939
Gaussian Noise 0.05	0.720	0.701	0.718	0.693	0.729	0.698	0.741	0.722
Salt & Pepper 0.005	0.955	0.940	0.963	0.943	0.968	0.951	0.971	0.960
Salt & Pepper 0.01	0.915	0.912	0.922	0.917	0.931	0.921	0.935	0.934
Median filter [3 3]	0.948	0.916	0.953	0.919	0.953	0.928	0.943	0.928
Median filter [5 5]	0.916	0.874	0.907	0.864	0.909	0.868	0.897	0.863
Speckle noise 0.01	0.871	0.869	0.888	0.876	0.906	0.891	0.924	0.911
Speckle noise 0.04	0.811	0.791	0.829	0.795	0.837	0.815	0.854	0.839
Poisson noise	0.880	0.871	0.893	0.876	0.908	0.897	0.922	0.920
Sharpening [2 2]	0.948	0.946	0.955	0.954	0.958	0.953	0.963	0.957
Cropping row off 25%	0.924	0.906	0.925	0.906	0.925	0.907	0.924	0.906
Cropping column off 25%	0.924	0.882	0.924	0.883	0.925	0.883	0.925	0.884
Centred Cropping 12.5%	0.946	0.958	0.947	0.958	0.947	0.959	0.947	0.959
Centred Cropping 25%	0.760	0.855	0.759	0.855	0.760	0.856	0.760	0.856

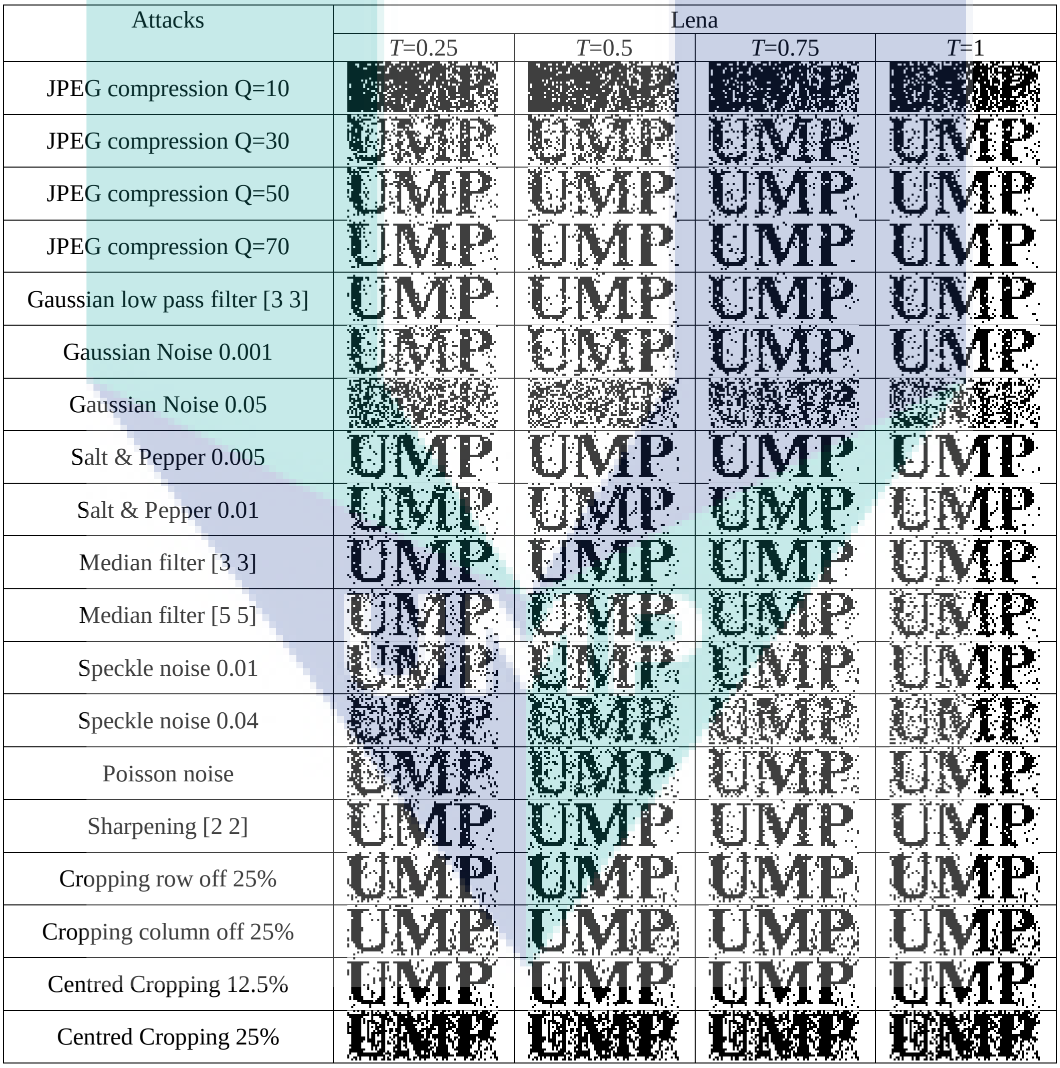


Fig. 6. The extracted watermark from different types of the attacked watermarked image.

CHAPTER 3

Bit Allocation Strategy based on Psychovisual Threshold in Image Compression

Ferda Ernawan - Muhammad Nomani Kabir - Jasni Mohamad Zain

Abstract: Image compression leads to minimize the storage-requirement of an image by reducing the size of the image. This paper presents a bit allocation strategy based on psychovisual threshold in image compression considering a similar idea of audio coding. In the audio coding, a dynamic bit allocation to each signal is related to the concept of variable block coding and bit allocation is performed on either a short block or long block of sample signals. Similarity, in our technique, more bits are assigned to a local block with visually-significant low frequency order, and fewer, with visually-insignificant high frequency order. This paper presents a bit allocation strategy based on psychovisual threshold in image compression. A psychovisual threshold is developed by minimizing the visual impact on the image quality degradation in image frequency coding. This paper investigates the error generated by the discrete cosine transform and sets the maximum acceptable error as a psychovisual threshold. The average reconstruction error per pixel on frequency order is utilized to prescribe a set of bit allocations which provide a significant improvement on the quality of image reconstruction at relatively low bit rates. The experimental results show that our dynamic bit-allocation technique in image compression manages to overcome artifact images in the image output. The proposed bit allocation strategy improves the quality of image reconstruction by about 20% compared to JPEG compression. This bit allocation strategy is designed to replace the traditional role of the quantization process in image compression.

Keywords bit allocation, psychovisual threshold, image compression, discrete cosine transform

F. Ernawan, M.N. Kabir, J.M. Zain

Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang,

Lebuhraya Tun Razak 26300 Gambang, Kuantan Pahang Darul Makmur, Malaysia

e-mail: ferda1902@gmail.com

1 Introduction

Currently, image compression is an important area of research due to growing number of high quality digital images which need to be compressed for size reduction and efficient transmission. In order to minimize the storage, many researchers have developed standard image compression algorithms [1][2][3], e.g., JPEG, JPEG 2000, JPEG XR, and JPEG XT, as well as many other non-standard compression algorithms [4][5][6] over the last two decades. Joint Photographic Expert's Group (JPEG) invented JPEG format for image compression in 1993. In 2000, JPEG introduced JPEG 2000 [2] as a standard image compression format using discrete wavelet transform. Currently, it is widely used in interactive streaming systems [7]. In 2007, Joint Photographic Expert's Group (JPEG) launched a new standardization of Microsoft's HD Photo as JPEG XR [8] and JPEG

XT [9] where the color is represented by 12 or 14 bits per pixel that provide an improvement over the original JPEG standard. Even though JPEG 2000 and JPEG XR were proposed for the HDR image coding standard, their usage has been limited in the industry mainly due to the lack of backward compatibility with the legacy of JPEG. Both formats JPEG 2000 and JPEG XR were patented which does not allow to access their formats. However, most display devices use the de facto international image coding standard, JPEG which still remains in low dynamic range (LDR) images whose color is represented by 8 bits per pixel [10]. Therefore, JPEG still dominates compression for many devices.

JPEG block-based coding [1] employs Discrete Cosine Transform (DCT) to convert the motion compensated residual data into frequency domain. An optimization of DCT quantization matrices was proposed by Watson [11] in 1993. The proposed quantization values are designed based on perceptual error levels. Implementation of this model can produce minimum bit rate for a given total perceptual error or perceptual minimum error for image reconstruction with a given bit rate. In 2002, Hontsch and Karam [12] proposed an adaptive image coding based on visual masking threshold for controlling quantization process. The proposed image coding algorithm showed improvement in terms of bit rate and distortion control. However, non-overlapping 8x8 quantization stage is used in image coding, it can exhibit visual artifact image and blocking effect when the reconstructed image is zoomed in 400%.

The major problem associated with the block-based coding is that the decoded images manifest visually objectionable artifact images such as ringing noise (dominant distortion), mosquito-noise and block noise [13]. The artifact images are smooth regions in image blocks containing a single dominant edge [14]. These occur at the transitions between image blocks and the most of the AC coefficients in high frequency order contains many zeros during quantization process. The block-based transform coding and quantization process during the encoding image bring about blocking effects [15] and visible artifact images when the decompressed image achieves high compression ratios. JPEG 2000, which is a successor of JPEG, can provide superior image quality than JPEG. However, this potential does not come without a price [16]. JPEG 2000 has a higher computational complexity than JPEG [17]. JPEG assumes a linear human contrast sensitivity function (CSF) which is a frequency-dependent quantizer considering a fixed diagonal matrix in the DCT domain. Malo et al., [18] argued that linear transforms cannot achieve minimum distortion in a reconstructed image. They proposed a non-linear image representation of CSF. The proposed technique improved both the statistical and the perceptual redundancy of the compressed image. However, the proposed non-linear method still produces artifact image and blocking effect in reconstructed images.

The main role of bit allocation in image compression is to reach the target bits or the quality level of image output [19]. A number of techniques have been proposed to reduce such bit redundancy in lossy image compression and video coding. The bit allocation strategies have been implemented extensively in region of interest (ROI) based coding [20], multi-view image coding [21]-[25], lossy image set compression [26] and video compression [27][28].

The inability of the human auditory system to hear occurs whenever a strong audio signal makes a spectral neighbourhood of weaker audio signals imperceptible [29]. In the noise-free environment, the human ear audibility requires different loudness across various frequency orders. The sound loudness that the human audibility can hear is called the absolute hearing threshold [30] as depicted in Fig. 1.

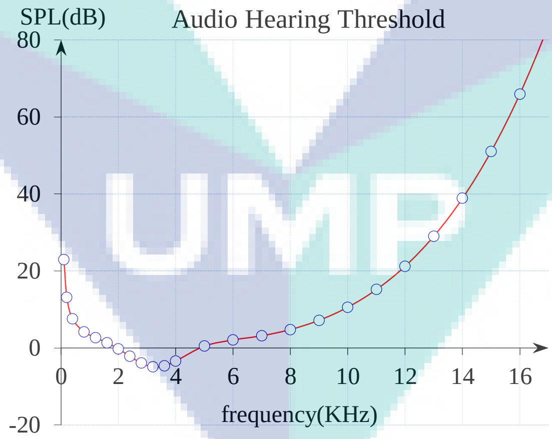


Fig. 1. Absolute threshold of hearing under quiet condition.

Referring to the psychoacoustic model, the human audibility does not perceive the sound of all frequencies equally but it detects some frequencies louder than others. The visual masking refers to the inability of human visual system to perceive a stimulus in the presence of another mask [31]. The masking phenomena are of a great relevance to image compression. Unfortunately, an explicit frequency masking signals on natural scenery images are not yet well established. This paper investigates the concept of frequency masking on the image signals by a bit allocation strategy in frequency signals. An experiment is conducted to assign bit allocation on the local frequency of image signals based on the psychovisual threshold in large DCT [32] used for image compression. The amount of bit allocation on each frequency order is set in a way so that it will provide balance impact on both the reconstruction error and compression rate.

The objective of this paper is to develop a bit allocation strategy for image compression. The proposed bit allocation strategy is able to achieve the minimum bit rate while it manages overall visual quality in reconstructed images. The concept of assigning bits budget from the audio masking model will be adopted in this work. Thus, implemented of the bit allocation strategy will improve visual quality of image texture and minimise the visual artifact image and the blocking effect for the reconstructed images. Bit allocation strategy intends to improve quality of reconstructed image with relatively low bit rates.

2 Psychovisual Threshold on Large Discrete Cosine Transform

A bit allocation strategy on the frequency signals is developed based on 40 raw images [33]. An input image consists of 512×512 color pixels. The RGB image components are converted to the YUV color space. An image is divided into four 256×256 blocks of image pixels. The two-dimensional DCT on 256×256 block is used to transform an image block into the frequency image signals. 256×256 DCT set $C_N(x)$ of size N=256 can be generated iteratively as follows:

$$C_0(x) = \frac{1}{\sqrt{N}}$$

$$C_1(x) = \sqrt{\frac{2}{N}} \cos \frac{(2x+1)1\pi}{2N}$$

$$C_2(x) = \sqrt{\frac{2}{N}} \cos \frac{(2x+1)2\pi}{2N}$$

$$\vdots$$

$$C_{N-1}(x) = \sqrt{\frac{2}{N}} \cos \frac{(2x+1)(N-1)\pi}{2N} \qquad (1)$$

For $x = 0, 1, 2, \ldots, N$-1. The first four $C(x)$ of one-dimensional DCT are shown in Fig. 2 for illustration.

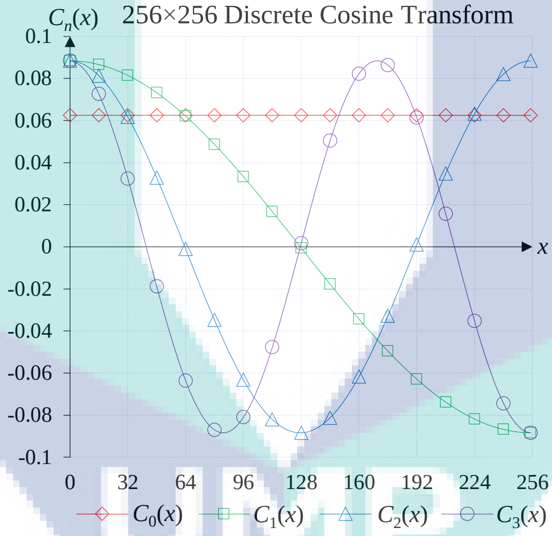


Fig. 2. First four $C_n(x)$ for n = 0, 1, 2, 3.

The kernel for the DCT [34] is defined as:

$$g(u) = \lambda(u) \cos \frac{(2x+1)u\pi}{2N} \qquad (2)$$

For $u = 0, 1, 2, \ldots, N$-1, where

$$\lambda(u) = \begin{cases} \frac{1}{\sqrt{N}}, & for \quad u = 0 \\ \sqrt{\frac{2}{N}}, & for \quad u > 0 \end{cases} \qquad (3)$$

The two-dimensional DCT B of an input image A is computed by [35]:

$$B_{pq} = \alpha_p \beta_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \qquad (4)$$

for $p = 0, 1, 2, \ldots, M–1$ and $q = 0, 1, 2, \ldots, N–1$ where

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, \; for \quad p = 0 \\ \sqrt{\frac{2}{M}}, \; for \quad p > 0 \end{cases} \qquad \beta_q = \begin{cases} \frac{1}{\sqrt{N}}, \; for \quad q = 0 \\ \sqrt{\frac{2}{N}}, \; for \quad q > 0 \end{cases} \tag{5}$$

The inverse of two-dimensional DCT is calculated using

$$A_{pq} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \alpha_p \beta_q B_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \tag{6}$$

for $p = 0, 1, 2, \ldots, M\text{-}1$ and $q = 0, 1, 2, \ldots, N\text{-}1$. In this work, 256×256 DCT is utilized to support the practical local bit allocation based on psychovisual threshold in image compression. The psychovisual threshold on large DCT has been developed based on the contribution of DCT coefficients to the absolute reconstruction error (ARE) for each frequency order [32]. ARE can be defined as:

$$ARE = \frac{1}{MNR} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \sum_{k=0}^{R-1} \left| g(i,j,k) - f(i,j,k) \right| \tag{7}$$

where $g(i,j,k)$ denotes the pixel value of the original image and $f(i,j,k)$ implies the pixel value of the compressed image. The original image size is $M \times N \times R$ with three RGB colors. The psychovisual error thresholds of 256×256 DCT coefficients for luminance and chrominance are shown in Fig. 3.

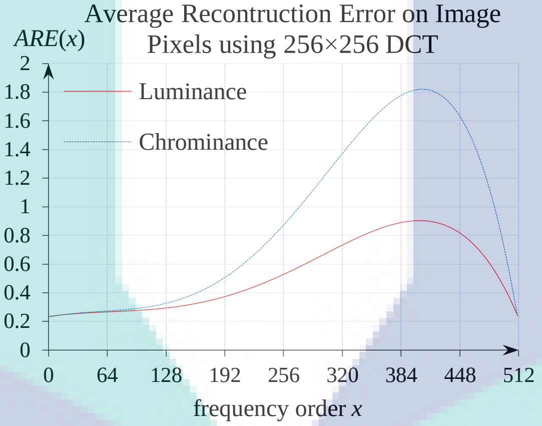


Fig. 3. Average reconstruction error on incrementing frequency order x for 40 real images.

In the figure, ideal plots of ARE for luminance and chrominance channels over the frequency order are shown by red curve and blue curve, respectively. Smooth curves of ARE can be obtained using polynomial functions that represent the psychovisual error threshold of 256×256 DCT for luminance f_{VL} and chrominance f_{VR}. The polynomial functions are given as follows:

$$f_{VL} = x^5 c_5 + x^4 c_4 + x^3 c_3 + x^2 c_2 + x^1 c_1 + c_0 \tag{8}$$

$$f_{VR} = x^5 d_5 + x^4 d_4 + x^3 d_3 + x^2 d_2 + x^1 d_1 + d_0 \tag{9}$$

Table 1: Polynomial coefficients

i	c_i	d_i
0	0.2352	0.2309
1	0.00088	0.0012
2	-0.000009	-0.0000128
3	0.000000046	0.00000006
4	0.000000000011	0.000000000156
5	-0.0000000000001435	-0.000000000000457

for the frequency order x = 0, 1, 2, ..., 512 on 256×256 DCT. The values of coefficients associated with the polynomials are listed in Table 1. These threshold functions (8)-(9) are used as a reference of the maximum acceptable error in assigning bit allocation on the local blocks for each frequency order. The psychovisual threshold on 256×256 DCT is designed to produce an optimal balance between compression rate and the image quality. Referring to psychovisual threshold above, we propose a technique of assigning bits to image signals. The bit-assignment in raw image signals across frequency orders is determined by ARE on each frequency order. The assigning bit allocation focuses mainly on the alternating current (AC) coefficients. The AC coefficients for each 256×256 image block are listed as a one dimensional array constructed by traversing the block in a zigzag pattern as shown in Fig. 4.

Referring to MPEG audio coding, the audio coding is structured with the concept of variable block coding and it performs on either a short block of 12 samples or long block of 36 samples [36][37]. The long block allows larger frequency resolution for audio signals with stationary characteristics, while the short block provides better time resolution for transient audio signals [38]. The spectral value of each block is separated into tonal and non-tonal components (noise). Then, the tonal and noise signals are classified for different encoding. In psychoacoustic model, the tonal component is identified based on the local spectral peaks of the audio power spectrum [29]. The surrounding of the frequency peak is identified as the non-tonal components (noise). Initially, the audio coding process removes the frequency signal component below the audibility threshold and removes the weaker tonal (noise) components surrounding of strongest tonal components. Masking on the noise component is more effective than masking on the complex frequency tones [39].

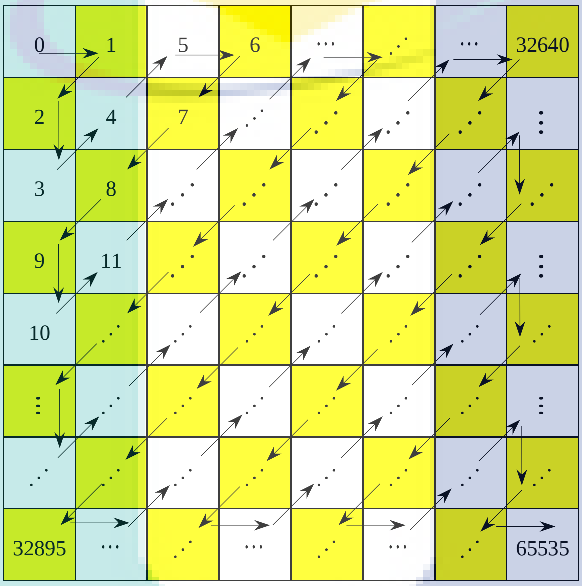


Fig. 4. A zigzag order on 256×256 block provides one-dimensional signals.

In this work, the principles of the audio coding are adopted to assign bits budget for image signals. The array of image signals under regular 256×256 discrete transform is divided into local frequency blocks. We propose three local frequency block-sizes of image signals: short blocks of 8 coefficients, medium blocks of 16 coefficients and long blocks of 32 coefficients. Each local block of AC coefficients is classified into local peak and non-peak signals. The assigning bit allocation focuses mainly on the non-peak signals of AC coefficients which have less significant contribution to the visual output of images.

3 A Bit Allocation Strategy

In our bit allocation strategy, the local peaks are managed in the original image signals by rounding to integer. The local peak signal is identified by finding the absolute maximum of local coefficients. There is only one peak signal coefficient and the rest are the non-peak signals on each local block. The peak and non-peak signals are separately encoded. Specifically, the bit allocation is used for the non-peak AC signals. The non-peak signals in the surrounding of peak signals are masked by assigning n bit allocation on each local block based on the psychovisual threshold. The masking bit on the non-peak signals $C(x)$ is proportionally encoded relative to the peak signal as follows:

$$M = round\left(2^n \frac{C(x)}{P}\right) \qquad (10)$$

where M is the masking frequency, x denotes the local frequency block, P implies the local frequency peak signal and n represents the number of bits assigned to the local coefficients. In order to determine the optimal bit n, the non-peak signals on each block are masked by incrementing bit allocation. The effect of incrementing bit allocation is measured by ARE per pixel and the compression rate. The number of bits is incremented until the average reconstruction error reaches the lower bound of the psychovisual threshold. The masking bits on each local block are assigned along the matrix diagonal order of the block from lower frequency order on the top left corner to the higher-frequency order on the bottom-right corner as shown in Fig. 4. The inverse of frequency non-peak coding can be determined by

$$C(x) = \left(\frac{M}{2^n}\right)P \tag{11}$$

where M is the masking frequency. The decoded frequencies of non-peak signals are closer to the original image signals. Finding optimal bits for the local blocks is a crucial part of this masking operation. A reduction of the number of bits assigned for image signals will significantly impact on the image quality and compression rates. For an optimal bit allocation for each block of AC coefficients, the effect of assigning bits to the average reconstruction error (ARE) must be limited to the psychovisual threshold on each frequency. The experimental results of generating bits budget on three local block-sizes are shown in Tables 2-4.

Table 2: Bit allocation for the local blocks with 8 coefficients of 256×256 DCT for luminance and chrominance

Luminance			Chrominance		
Block No	Freq order	Bit	Block No	Freq order	Bit
1	1-3	9	1	1-3	7
2-4	4-7	8	2-3	4-6	6
5-16	8-15	7	4-9	7-11	5
17-49	16-27	6	10-31	12-21	4
50-170	28-51	5	32-88	22-37	3
171-479	52-87	4	89-213	38-57	2
480-1097	88-131	3	214-638	58-100	1
1098-1831	132-170	2	639-8191	101-507	0
1832-2977	171-217	1			
2978-8191	218-507	0			

Table 3: Bit allocation for the local blocks with 16 coefficients of 256×256 DCT for luminance and chrominance

Luminance			Chrominance		
Block No	Freq order	Bit	Block No	Freq order	Bit
1	1-5	9	1	1-5	7
2-3	6-9	8	2-3	6-9	6
4-12	10-19	7	4-6	10-13	5
13-37	20-33	6	7-22	14-26	4
38-119	34-61	5	23-60	27-43	3
120-323	62-101	4	61-145	44-67	2
324-659	102-144	3	146-417	68-115	1
660-1026	145-180	2	418-4095	116-505	0
1027-1743	181-235	1			
1744-4095	236-505	0			

Table 4: Bit allocation for the local blocks with 32 coefficients of 256×256 DCT for luminance and chrominance

Luminance			Chrominance		
Block No	Freq order	Bit	Block No	Freq order	Bit
1	1-7	9	1	1-7	7
2	8-11	8	2	8-11	6
3-8	12-22	7	3-5	12-18	5
9-27	23-41	6	6-14	19-29	4
28-87	42-74	5	15-39	30-49	3
88-205	75-114	4	40-97	50-78	2
206-376	115-154	3	98-244	79-124	1
377-590	155-193	2	245-2047	125-503	0
591-970	194-248	1			
971-2047	249-503	0			

Form Tables 2-4, it can be observed that the luminance requires more bit allocations than the chrominance channel. It implies that the luminance channel of the encoded image signals on requires more bits to perceptually maintain the relevant image data. The sensitivity level in terms of the bit allocation mainly depends on the frequency order and the local block size. The local block of image signals on lower frequency order requires more bits to store the image data. In this experiment, bit allocation on a large transform coding with 65535 AC coefficients is investigated. Due to the large block size, the large transform coding provides high quality image output with an overall low computational time and low power consumption [40]. First, this experiment examines the bit allocations needed on the local block of 8 coefficients, 16 coefficients and 32 coefficients.

As discussed before, the tonal signals in the audio coding represent the most energy of the audio signals which contain the important information to human ears [41]. Thus, we assume that the peak AC coefficients of each local block provide significant contribution to the image quality. Therefore, the peak AC coefficients are maintained as in the original image signals. The bit allocation is used for the non-peak AC coefficients. The histograms of peak AC coefficients and non-peak AC coefficients after assigning the bits are shown in Fig. 5 and Fig. 6, respectively.

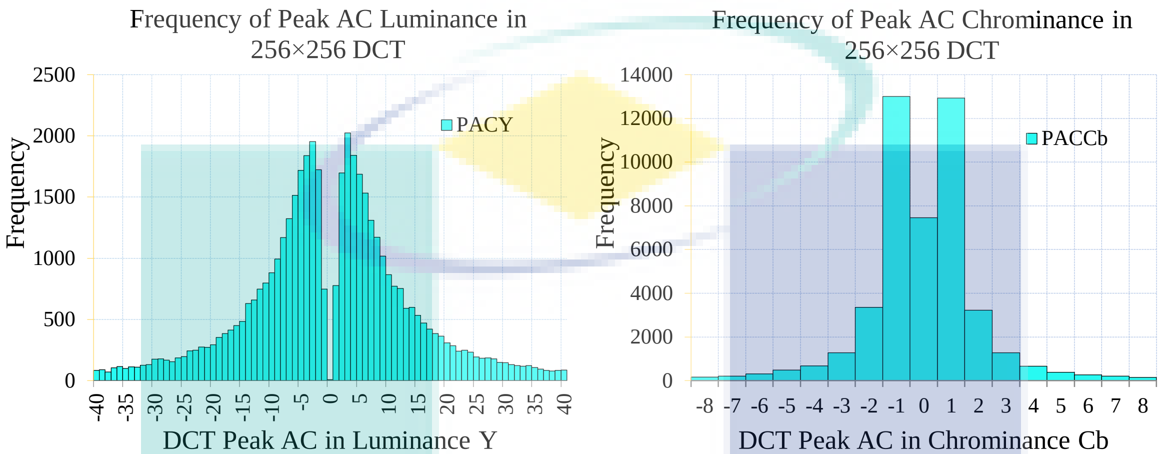


Fig. 5. Histograms of peak AC coefficients in 256×256 DCT for 40 real images.

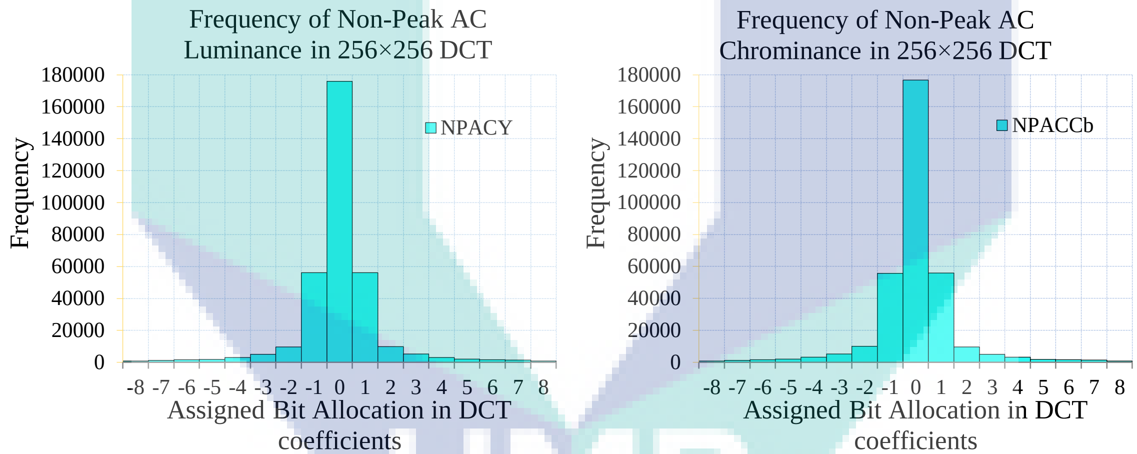


Fig. 6. Histograms of non-peak AC coefficients in 256×256 DCT for 40 real images.

Referring to Fig. 5, the histogram of peak image signals generates extreme value distribution for AC coefficients. The frequency distribution of non-peak signals produces a normal distribution as depicted in Fig. 6. We use the probability density function to generalize extreme values of peak image signals as follows [42]:

$$y = f(x|k,\mu,\sigma) = \left(\frac{1}{\sigma}\right)\exp\left(-\left(1+k\frac{(x-\mu)}{\sigma}\right)^{-\frac{1}{k}}\right)\left(1+k\frac{(x-\mu)}{\sigma}\right)^{-1-\frac{1}{k}} \tag{12}$$

$$\text{with } k \neq 0 \text{ and } 1+k\frac{(x-\mu)}{\sigma} > 0$$

where μ represents the location, σ is a scale parameter and k represents the shape parameter. The generalized extreme value distribution is often used to model the smallest or largest value among a large set [42]. According to the histogram as shown in Fig. 5, this experiment uses the generalized extreme value probability to estimate the distribution of the peak AC coefficients as shown in Fig. 7. A red curve represents the generalized extreme distribution estimation at each frequency signal. These results are used to calculate the probability density function (pdf). The probability density function on the peak AC coefficients and non-peak AC coefficients are depicted in Fig. 8 and Fig. 9, respectively. The total area under the red graphs of the probability distribution is equal to one.

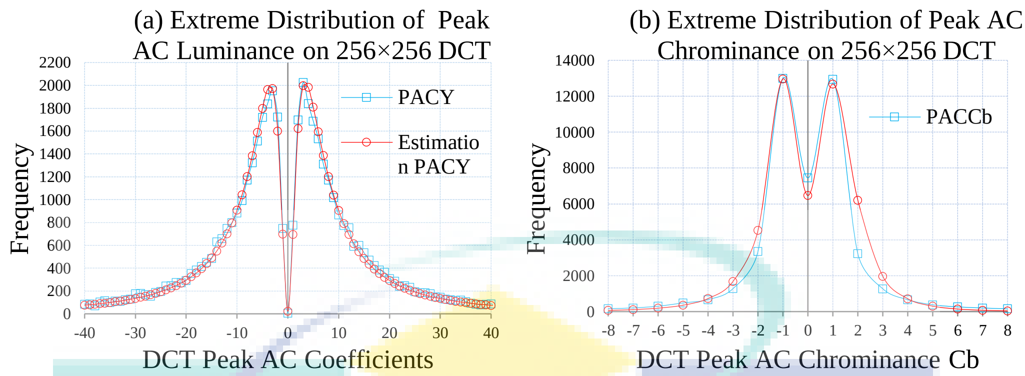


Fig. 7. The extreme distribution of peak AC coefficients in 256×256 DCT for 40 real images.

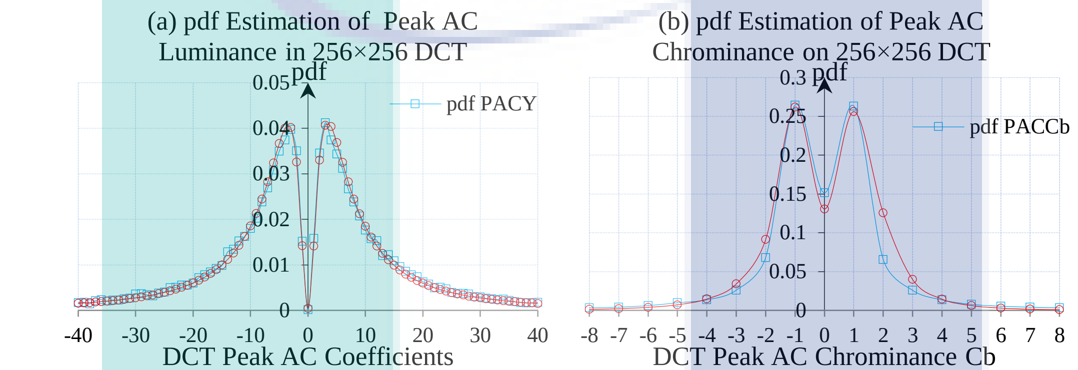


Fig. 8. The probability density function of the peak AC coefficients in 256×256 DCT for 40 real images.

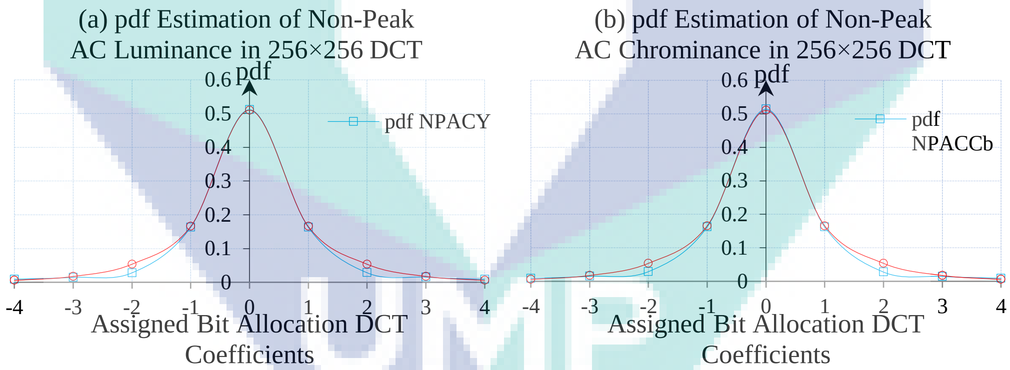


Fig. 9. Probability density function of the non-peak AC coefficients in 256×256 DCT for 40 real images.

The normal distribution is used to estimate the probability density function of non-peak AC coefficients after assigning bit allocation. The probability density of non-peak AC coefficients is depicted as a blue curve in Fig. 9. The normal distribution estimates the probability density of the non-peak AC coefficients as presented by red curve. The proposed normal distribution under the red graph has a total probability density equal to one.

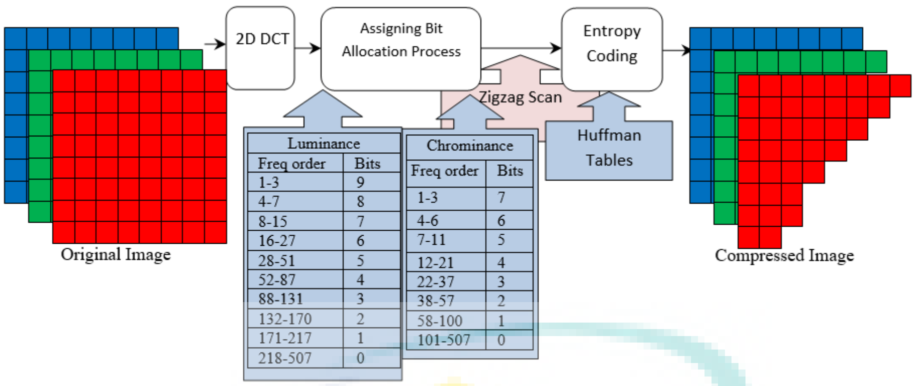


Fig. 10. Image Compression using bit allocation strategy.

4 Experimental Setup

The proposed bit allocation strategy for image compression scheme can be illustrated in Fig. 10. An experiment of assigning bit allocation for image compression has been carried out using 40 real and 40 graphical images. First, all 80 RGB images are converted to the YUV color space. An image is divided into four 256×256 blocks of image pixels and it is transformed by DCT. The four DC coefficients and AC coefficients are separately encoded. The AC coefficients are encoded based on a set of bits for each frequency order. The bit allocation has to achieve a trade-off between average compression rates and the high quality on the image output. Thus, AC coefficients are listed as a traversing array in a zigzag pattern. Next, the run-length encoding is used to shorten any repeating coefficients in the sequence of AC coefficients. The AC coefficients are represented compactly by the coefficient value and the length of its run wherever it appears. The output of run-length encoding represents the symbols and the length of occurrence of the symbols. The symbols and variable length of occurrence are used in Huffman coding to retrieve the code words and their lengths. Using these probability values, a set of Huffman code of the symbols is generated via Huffman Tree. Next, the average bit length is derived from Huffman codes for AC coefficients. The evaluation sets on image quality are ARE, means square error (MSE), peak signal to noise ratio (PSNR) and structural similarity (SSIM) index. MSE calculates the average of the squares of the errors defined by [43]:

$$MSE = \frac{1}{MNR}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\sum_{k=0}^{R-1}\left[g(i,j,k) - f(i,j,k)\right]^2 \tag{13}$$

The standard PSNR is calculated to obtain the quality of image reconstruction. A higher PSNR means that the reconstructed image is more similar to the original image [44]. PSNR is defined as:

$$PSNR = 20\log_{10}\left(\frac{Max_i}{\sqrt{MSE}}\right) = 10\log_{10}\left(\frac{255^2}{MSE}\right) \tag{14}$$

where Max_i is the maximum possible pixel value of the image. Another measurement of image quality is SSIM, a method to measure quality by capturing the similarity between original image and compressed image [45]. SSIM is computed by:

$$SSIM(x,y) = [l(x,y)]^{\alpha} \cdot [c(x,y)]^{\beta} \cdot [s(x,y)]^{\gamma} \tag{15}$$

where $\alpha>0$, $\beta>0$, $\gamma>0$, are parameters which can be adjusted to signify their relative importance. A detail description can be found in [45]. The proposed bit allocation will be tested and compared with existing quantization tables in image compression to demonstrate the strength of our technique. The quantization table used by JPEG standard [46] tends to preserve low-frequency information and discard high-frequency (HVS) details, because HVS is less sensitive to the information-loss in high-frequency bands. However, JPEG standard quantization table only considers the HVS features while the relationship between the distortions and rates has been largely ignored [47]. In [48], Ernawan and Nugraini derived a new JPEG quantization tables according to a psychovisual threshold to achieve the optimal balance between the quality of reconstructed images and compression rates. The psychovisual threshold is developed by a quantitative experiment that can automatically predict perceptual image quality.

In [49], TMT quantization tables were proposed by Rahmalan et al. with Tchebichef compression. Their proposed quantization tables can improve compression rate up to 80% than JPEG compression. TMT quantization tables are considered to preserve low-frequency information and to maintain the quality of reconstructed images. A TMT quantization table [50] is derived based on psychovisual threshold of Tchebichef moments. This quantization table is designed based on the trade-off between reconstruction error and compression rate for each frequency order. The quantization values are increased one at a time on each frequency order to measure an optimal balance between quality of reconstructed image and compression rate. However, all of these quantization tables produce large artifact images and blocking effects due to non-overlapping blocks of image compression. Therefore, the default JPEG quantization tables [46], DCT quantization tables [48], TMT quantization tables [49], TMT quantization tables [50], and the proposed bit allocation are compared to demonstrate its compression performance.

5 Experimental Results

Practically, the results of bit allocation in Tables 2, 3, and 4 are used to replace the quantization tables in image compression. The proposed bit allocation strategy is used to assign the maximum number of bits in DCT coefficients for each frequency block. Our strategy has been compared to JPEG compression [46], DCT psychovisual threshold in image compression [48], TMT compression [49] and TMT psychovisual threshold in image compression [50] in terms of average bit length of Huffman code, quality image output, compression ratio and file size. The experimental results show the average bit length of Huffman code on AC Luminance (ACY), AC Chrominance red (ACU) and AC chrominance blue (ACV) respectively for image compression as listed in Table 5.

Table 5. Average bit lengths of Huffman code in image compression for 40 real images and 40 graphical images

Method	40 Real Images			40 Graphical Images		
	ACY	ACU	ACV	ACY	ACU	ACV
Default quantization tables in JPEG compression [46]	2.868	2.095	2.184	2.965	2.505	2.515
Quantization tables generation from DCT psychovisual image compression [48]	2.796	2.143	2.072	2.900	2.464	2.406
Quantization tables in TMT image compression [49]	1.767	1.260	1.212	2.040	2.061	1.983
Quantization tables generation from TMT psychovisual image compression [50]	**1.764**	**1.191**	**1.166**	**2.027**	**1.745**	**1.625**
Bit allocation on the local blocks of 8 coefficients	2.917	2.459	2.533	2.994	2.734	2.761
Bit allocation on the local blocks of 16 coefficients	2.741	2.513	2.531	2.763	2.636	2.654
Bit allocation on the local blocks of 32 coefficients	2.661	2.558	2.584	2.605	2.588	2.578

Referring to Table 5, the proposed bit allocation strategy produces higher average bits per pixel on AC coefficients resulting in a large number of AC coefficients under regular 256×256 block. The maintaining bits on peak signals of AC coefficients as the original image signals also provide less effect to the average bit length of Huffman code. Since there are only few direct current (DC) coefficients, they are maintained as the original value without any compression. There are four DC coefficients of an image under regular 256×256 DCT. The four DC coefficients of a 512×512 image do not make significant effect on the compression bit rate of image compression, while they produce most significant impact on the quality of image reconstruction. Otherwise, JPEG image compression provides 4096 DC coefficients under regular 8×8 DCT which produce significant effect to the image quality and the average bit rate.

The average measurements of the image quality are shown in Table 6. Bit allocation on local image signals produces significantly higher statistical image quality from the reconstruction score than JPEG compression. The proposed bit allocation on AC coefficients is able to achieve high quality in image reconstruction. The experimental results show that the large local block size gives significant effect on quality of the image reconstruction and the average bit length of Huffman code. The experimental results of average compression rate and compressed image size are shown in Table 7. The amount of encoded data required to represent a digital image is equal to the size of the compressed image file. The average compression ratios of JPEG compression and the proposed bit allocation technique are about 3.324 and 3.06, respectively. However, the quality of the image reconstructed from the assigned bits is much higher than JPEG compression.

In the previous research, a generic quantization table from DCT psychovisual error threshold was proposed to replace the default JPEG quantization tables. The quantization table in [48] showed slightly better than default JPEG quantization tables in image coding. In 2014, Ernawan et al. investigated the psychovisual threshold in Tchebichef moment to generate new TMT quantization tables [50]. These TMT quantization tables exhibited improvement of compression performance in terms of quality and bit rate compared to quantization tables [49]. However, 8×8 quantization tables from psychovisual threshold still produce blocking effects due to non-overlapping blocks in image coding. In the present work, bit allocation based on psychovisual threshold is used that exhibits less artifact images of the reconstructed image. The proposed bit allocation method is able to overcome artifact images and blocking effects of reconstructed image using 8×8 quantization tables.

Table 6 Measurement of average image quality from 40 real images and 40 graphical images

Method	40 Real Images				40 Graphical Images			
	Error	MSE	PSNR	SSIM	Error	MSE	PSNR	SSIM

Default quantization tables in JPEG compression [46]	5.535	70.964	31.190	0.956	5.648	92.711	31.636	0.957
Quantization tables generation from DCT psychovisual image compression [48]	5.499	69.520	31.252	0.955	5.445	82.237	31.880	0.957
Quantization tables in TMT image compression [49]	5.258	58.159	31.372	0.947	4.840	61.036	32.274	0.952
Quantization tables generation from TMT psychovisual image compression [50]	5.246	57.448	31.379	0.946	4.767	56.745	32.363	0.951
Bit allocation on the local blocks of 8 coefficients	**3.123**	**20.593**	**37.565**	**0.983**	**3.513**	**29.874**	**38.338**	**0.980**
Bit allocation on the local blocks of 16 coefficients	3.449	25.225	36.740	0.982	3.820	35.382	37.678	0.978
Bit allocation on the local blocks of 32 coefficients	3.781	29.788	35.845	0.980	4.200	41.422	36.506	0.977

Table 7. Average compression rate score and the size of the compressed image from 40 real images and 40 graphical images

Method	Compression Rate		Bit size	
	40 Real Images	40 Graphical Images	40 Real Images	40 Graphical Images
Default quantization tables in JPEG compression [46]	3.324	2.972	230.99 Kb	258.39 Kb
Quantization tables generation from DCT psychovisual image compression [48]	3.388	3.053	226.66 Kb	251.54 Kb
Quantization tables in TMT image compression [49]	5.559	3.892	138.14 Kb	197.28 Kb
Quantization tables generation from TMT psychovisual image compression [50]	**5.713**	**4.372**	**134.41 Kb**	**175.63 Kb**
Bit allocation on the local blocks of 8 coefficients	3.033	2.826	253.13 Kb	271.69 Kb
Bit allocation on the local blocks of 16 coefficients	3.081	2.979	249.20 Kb	257.79 Kb
Bit allocation on the local blocks of 32 coefficients	3.074	3.087	249.78 Kb	248.76 Kb

The proposed bit allocation improves the visual quality of reconstructed images by about 20% $(((37.565-31.190)/31.190)\times100\%=20.43\%)$ in terms of PSNR compared to JPEG compression. However, the JPEG compression is able to produce slightly more compression rate than the proposed bit allocation. Furthermore, the compression rate of TMT is higher than all other techniques. TMT uses mathematical framework using matrix and it can perform better than DCT in compression rate. The proposed bit allocation is designed for DCT coefficients. In addition, JPEG and TMT image compressions still produce artifact and block effects due to non-overlapping blocks.

We observe that the proposed bit allocation technique on local blocks of 8 coefficients in image compression produces better image quality than that on local blocks of 16 or 32 coefficients. Since there are 8191 peak signals of 65535 AC coefficients in the local blocks of 8 coefficients, bit allocation for peak signals with low frequency order significantly improves the visual quality of image texture. The peak signals contribute 12.49% of the AC coefficients, while the non-peak signals contribute 87.51% and they do not make significant effect to the image quality. Table 8 shows the comparison of the visual output among JPEG compression, TMT image compression, DCT psychovisual threshold, TMT psychovisual threshold and our technique of set of bits budget in image compression.

Quantitative measurement has been perfomed using MSE, PSNR, SSIM compression rate. In Table 8, the proposed method achieves a PSNR value of 40.569 db on Lena image. On the contrary, the same image in Table 8 generates PSNR value of 32.915 db with JPEG image compression. TMT image compression produces an ARE value of 2.419 for Tauchan image. The corresponding ARE of the proposed bit allocation achieves lower value i.e., 0.816 than the other approaches for the same image. Table 8 also shows that SSIM value of 0.990 from the proposed method is little greater than the results i.e. 0.942 and 0.932 as obtained from DCT psychovisual and TMT psychovisual image compressions on Pepper image.

In order to observe respectively the visual quality, the samples of the visual image outputs are cropped and zoomed in to 200% in the specific texture colors. Table 9, 11, 13 show comparison of the visual outputs of different images from JPEG, DCT psychovisual, TMT compression, TMT psychovisual and the proposed bit allocation strategy with 8, 16, 32 coefficients. Tauchan images obtained from Table 8 are cropped at the mouth of Tauchan as shown in Table 9. The right of lena eye in Table 8 is selected to analyse the visual quality as given in Table 11. Stalk of peppers image from Table 8 is cropped and chosen to demonstrate the visual image quality of the proposed bit allocation as depicted in Table 13.

Table 8. Comparison of the visual outputs from JPEG compression, DCT psychovisual, TMT compression, TMT psychovisual and the proposed bit allocation strategy for image compression

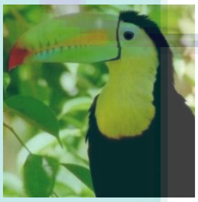
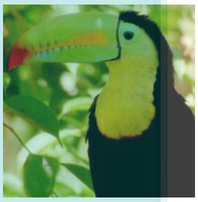
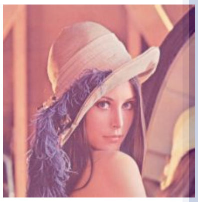
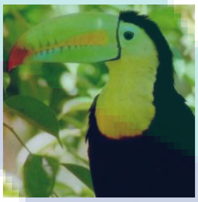
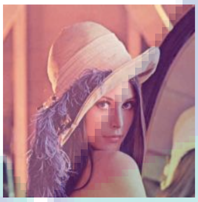
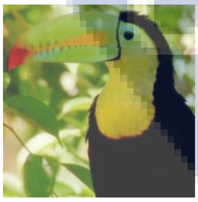
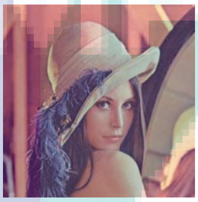
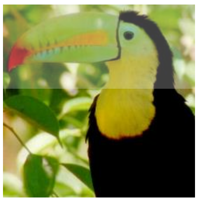
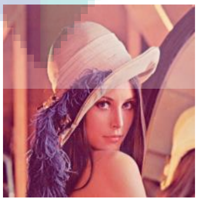
	Toucan image	Lena image	Peppers image
JPEG compression [46]	 ARE=2.461 PSNR=37.534 SSIM=0.958	 ARE=4.180 PSNR=32.915 SSIM=0.950	 ARE=5.064 PSNR=31.395 SSIM=0.943
DCT psychovisual image compression [48]	 ARE=2.423 PSNR=37.421 SSIM=0.964	 ARE=4.204 PSNR=32.882 SSIM=0.947	 ARE=5.022 PSNR=31.536 SSIM=0.942
TMT image compression [49]	 ARE=2.419 PSNR=36.865 SSIM=0.954	 ARE=4.394 PSNR=32.660 SSIM=0.939	 ARE=4.877 PSNR=31.792 SSIM=0.934
TMT psychovisual image compression [50]	 ARE=2.444 PSNR=36.763 SSIM=0.953	 ARE=4.410 PSNR=32.615 SSIM=0.938	 ARE=4.879 PSNR=31.778 SSIM=0.932
Bit Allocation on each 8 coefficients	 **ARE=0.816 PSNR= 47.875 SSIM=0.994**	 **ARE=1.845 PSNR=40.569 SSIM= 0.992**	 **ARE=2.243 PSNR=39.138 SSIM=0.990**

Table 9. Comparison of the visual output (zoomed in to 200%) of Toucan image from JPEG, DCT psychovisual, TMT compression, TMT psychovisual and the proposed bit allocation strategy with 8, 16, 32 coefficients

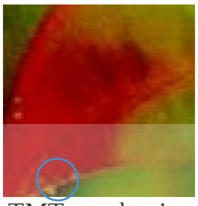
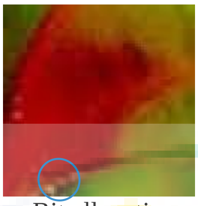
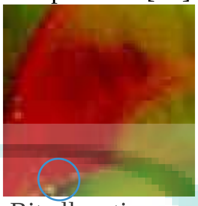
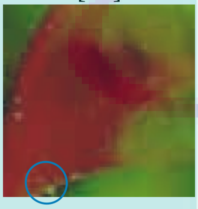
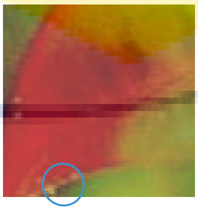
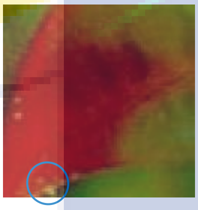
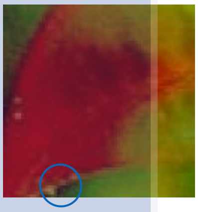
Original	Original image croped and zoomed in to 200%	JPEG compression [46]	DCT psychovisual image compression [48]	TMT image compression [49]
Mouth of toucan image				
	TMT psychovisual image compression [50]	Bit allocation on each 8 coefficients	Bit allocation on each 16 coefficients	Bit allocation on each 32 coefficients
Mouth of toucan image				

Table 10: Comparison of the visual output based on the blue markers of mouth of Toucan 10×10 pixels

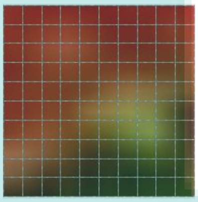
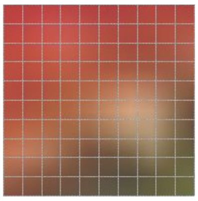
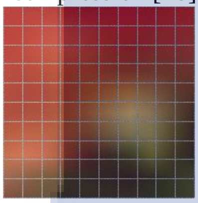
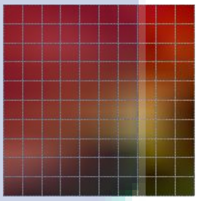
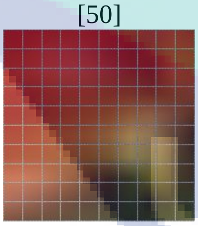
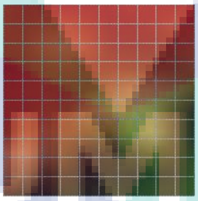
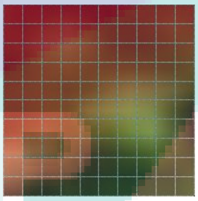
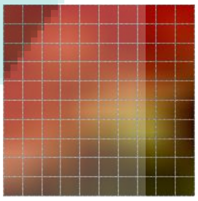
Original	Original image croped and zoomed in to 200%	JPEG compression [46]	DCT psychovisual image compression [48]	TMT image compression [49]
Mouth of toucan image				
	TMT psychovisual image compression [50]	Bit allocation on each 8 coefficients	Bit allocation on each 16 coefficients	Bit allocation on each 32 coefficients
Mouth of toucan image				

Table 11: Comparison of the visual output (zoomed in to 200%) of Lena image from JPEG, DCT psychovisual, TMT compression, TMT psychovisual and the proposed bit allocation strategy with 8, 16, 32 coefficients

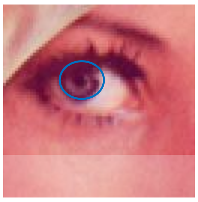
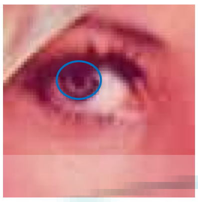
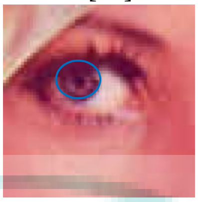
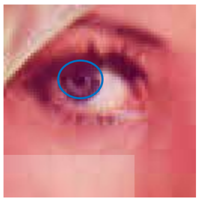
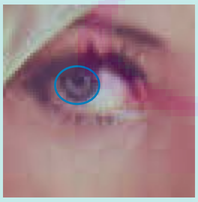
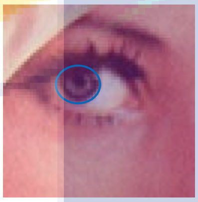
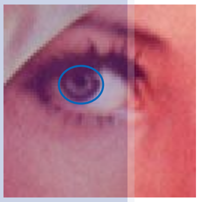
Original	Original image croped and zoomed in to 200%	JPEG compression [46]	DCT psychovisual image compression [48]	TMT image compression [49]
Rigt eye of Lena image				
	TMT psychovisual image compression [50]	Bit allocation on each 8 coefficients	Bit allocation on each 16 coefficients	Bit allocation on each 32 coefficients
Rigt eye of Lena image				

Table 12. Comparison of the visual output based on the blue markers of right eye of Lena10×10 pixels

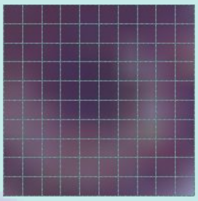
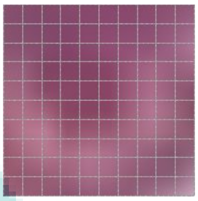
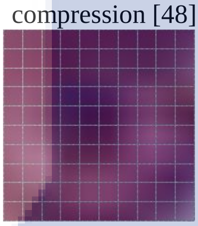
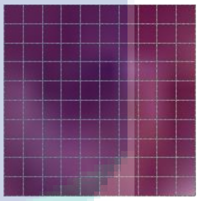
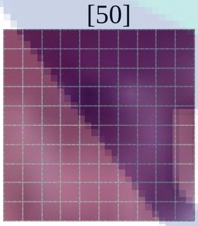
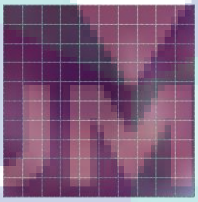
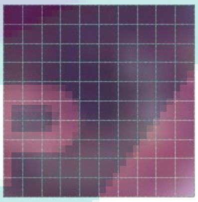
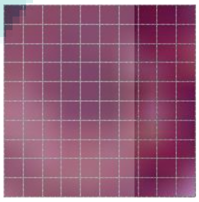
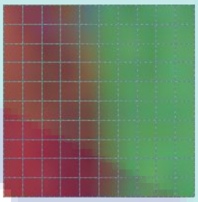
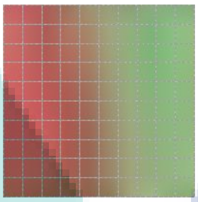
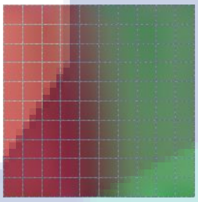
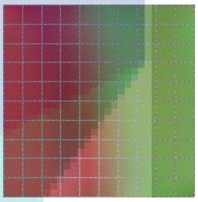
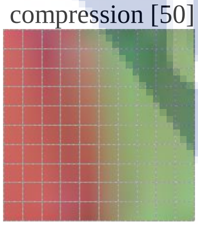
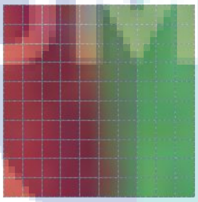
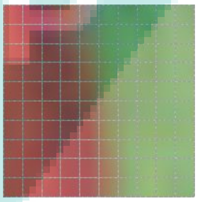
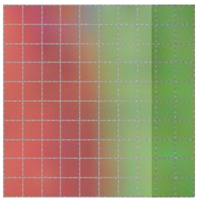
Original	Original image croped and zoomed in to 200%	JPEG compression [46]	DCT psychovisual image compression [48]	TMT image compression [49]
Rigt eye of Lena image				
	TMT psychovisual image compression [50]	Bit allocation on each 8 coefficients	Bit allocation on each 16 coefficients	Bit allocation on each 32 coefficients
Rigt eye of Lena image				

Table 13: Comparison of the visual output (zoomed in to 200%) of Peppers image from JPEG, DCT psychovisual, TMT compression, TMT psychovisual and the proposed bit allocation strategy with 8, 16, 32 coefficients

Original	Original image croped and zoomed in to 200%	JPEG compression [46]	DCT psychovisual image compression [48]	TMT image compression [49]
Stalk of peppers image				
	TMT psychovisual image compression [50]	Bit allocation on each 8 coefficients	Bit allocation on each 16 coefficients	Bit allocation on each 32 coefficients
Stalk of peppers image				

Table 14: Comparison of the visual output based on the blue markers of right eye of Peppers 10×10 pixels

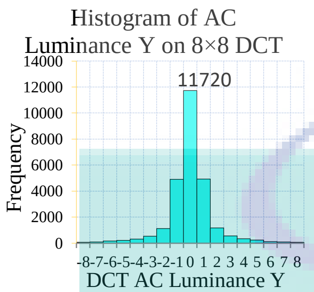
Original	Original image croped and zoomed in to 200%	JPEG compression [46]	DCT psychovisual image compression [48]	TMT image compression [49]
Stalk of peppers image				
	TMT psychovisual image compression [50]	Bit allocation on each 8 coefficients	Bit allocation on each 16 coefficients	Bit allocation on each 32 coefficients
Stalk of peppers image				

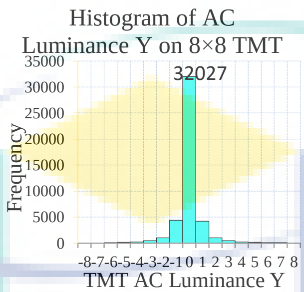
In order to analyze the texture color pixels, we select specified regions to identify the visual intensity of color pixels. Tables 10, 12, 14 present the detail of visual 10×10 pixels of the specified regions (marked by a circle) of the images in Tables 9, 11, and 13, respectively. In these images, we check that the pixel colors of our technique resemble to the original which demonstate the higher quality of the proposed technique. The JPEG compression output as depicted on the second column of Table 10 appears significantly to contain artifact images with blocking effects. The image samples for a set of bits budget for the local blocks of 8, 16 and 32 coefficients in image compression are very close to one another and visually similar textures. The image outputs from a set of bits budget produce rich texture pixels. Referring to the second row of Tables 12 and 14, the image output from a set of bits budget on the local blocks of 8 coefficients shows a significantly closer to the original image than the other image outputs. The proposed bits allocation strategy produces less artifact image and smooth interchange image pixels in the image output.

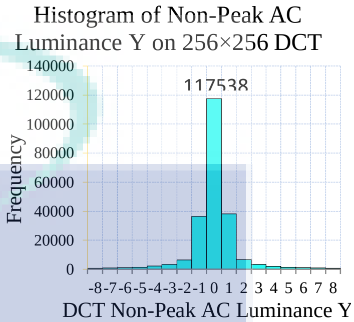
Table 15 shows the histograms of AC luminance for three different images with three different methods – DCT, TMT and the proposed bit allocation method. For all the images, it can be noticed that the proposed method has higher peak-values than DCT. However, TMT has higher peak-values than the proposed method except for Toucan image. Moreover, slenderness of the histogram for the proposed method is between that for TMT and DCT.

Table 15. Comparison of the frequency distributions among JPEG compression, TMT compression and the proposed bit allocation strategy

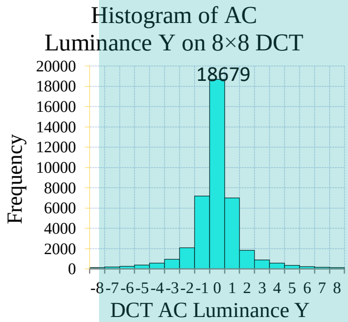
JPEG compression [46]	TMT compression [49]	Bit allocation on each 8 coefficients

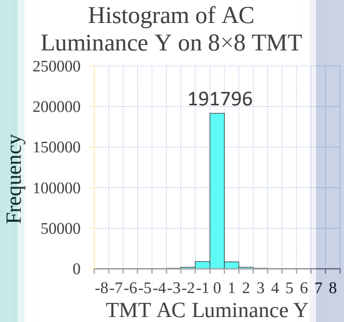


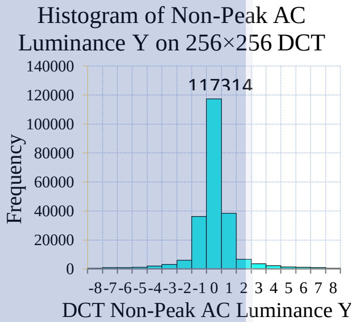




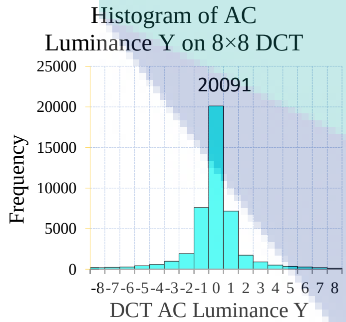
Toucan image | Toucan image | Toucan image

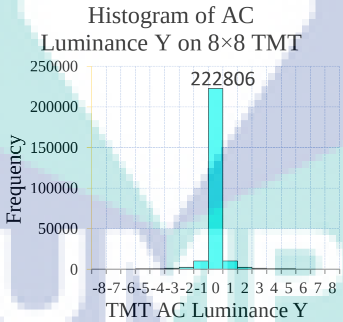


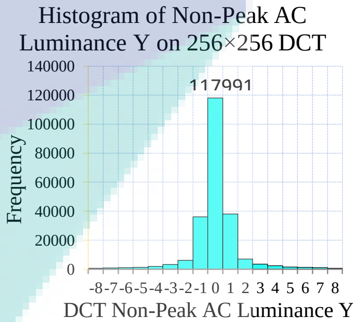




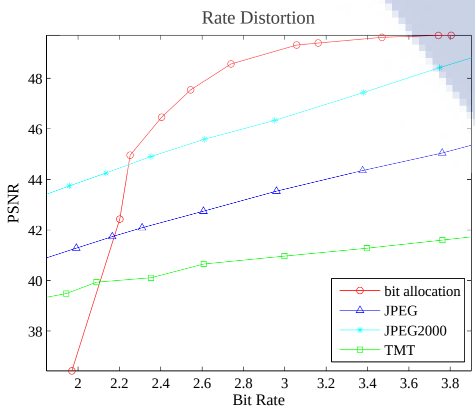
Lena image | Lena image | Lena image

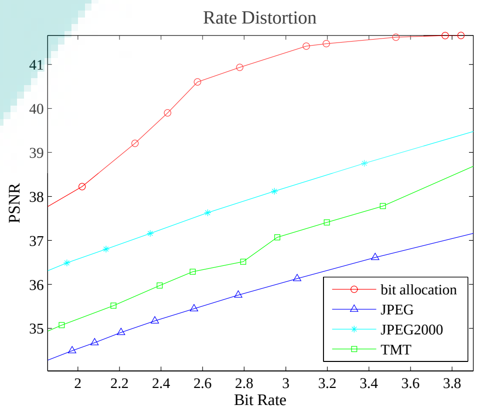




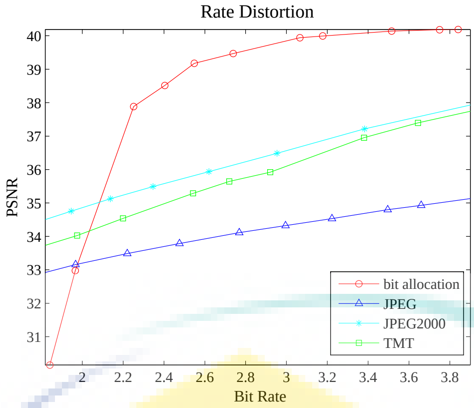


Peppers image | Peppers image | Peppers image





(a) Taucan (b) Lena



(c)Pepper

Fig. 11: Comparison of the rate distortion among the proposed bit allocation strategy, JPEG compression, JPEG2000 and TMT compression.

Comparison of the rate distortions of JPEG, JPEG2000, TMT compression and the proposed bit allocation is shown in Fig. 11. The rate distortion of JPEG, TMT compression has been calculated by different quality factors. Moreover, the rate distortion of JPEG2000 has been computed by different compression ratios. On the contrary, for the same image, the rate distortions of the proposed bit allocation have been measured by scaling the bit allocation. The proposed method seems to degrade fast at low rates (bit rates less than 2.3-2.5) except for Lena image. The proposed bit allocation achieves high quality at low compression rate on Lena image. As illustrated in Fig. 11, the proposed bit allocation exhibits higher image quality than JPEG, JPEG2000, and TMT image compression. We check that the proposed bit allocation strategy outperforms other techniques when the bit rate is greater than 2.3.

In terms of computational time, the proposed bit allocation strategy performs slightly slower than JPEG. Due to the advanced technology with faster processing power, the quality of image output is more vital and important issue. In addition, improving quality of image output should achieve a tradeoff between the quality of reconstructed image and the bit rate. JPEG2000 requires a wavelet transform to perform image compression. JPEG2000 provides higher computational complexity than our strategy. However, their usage has been limited in the industry mainly due to the lack of backward compatibility with the JPEG.

6 Conclusion

This paper presents a bit allocation strategy for image compression based on noticeable differences in average absolute reconstruction errors in frequency order. The bit allocation strategy assumes that assigning more bits on low frequency order and fewer bits on high frequency order can keep the visual quality and avoid serious degradation. Here, the peak image signals are maintained as in original frequency signals. Peak image signals that provide the important information to human visual system contain a significant contribution to the image texture quality. The proposed bit allocation technique provides an optimal image quality with minimum bit allocation to represent image signals. The technique is based on the psychovisual threshold in each frequency order has been designed to replace the main role of the quantization process in image compression. The experimental results indicate the proposed technique in image compression for short local blocks produces a rich texture on the image output which is close to the original image. Unlike a typical outcome from the quantization process in image compression, our technique generates minimum artifact images in the visual image output. Comparison with other techniques indicates that our bit allocation technique provides higher quality in image reconstruction with relatively moderate compression rates.

Acknowledgments

The authors would like to express very special thanks to Universiti Malaysia Pahang, Malaysia for providing financial support of this research project by UMP Research Grant Scheme (RDU160364).

REFERENCES

[1] Information Technology (1994) JPEG digital compression and coding of continuous-cone still image. Part I: Requirement and guidelines ISO/IEC 10918-1 and ITU-T Recommendation T.81.

[2] (2000) JPEG 2000 part I final committee draft version 1.0, ISO/IEC JTC1/SC29/WG1 N1646R.

[3] Richter T, Artusi A, Ebrahimi T (2016) JPEG XT: A new family of JPEG backward-compatible standard, IEEE MultiMedia, 23(3): 080-088.

[4] Ernawan F, Abu NA, Suryana N (2013) Adaptive tchebichef moment transform image compression using psychovisual model. Journal of Computer Science 9(6): 716-725.

[5] Ernawan F, Abu NA, Suryana N (2014) An adaptive JPEG image compression using psychovisual model. Advanced Science Letters 20(1): 026-031.

[6] Do B, Zhang M, Zhang L, Hu R, Tao D (2017) PLTD: Patch-based low-rank tensor decomposition for hyperspectral images. IEEE Transactions on Multimedia, 19(1): 067-079.

[7] Sanchez-Hernandez JJ, Garcia-Ortiz JP, Gonzalez-Ruiz V, Muller D (2015) Interactive streaming of sequences of high resolution JPEG2000 images. IEEE Transactions on Multimedia, 17(10): 1829-1838.

[8] Srinivasan S, Tu C, Zhou Z, Ray D, Regunathan S, Sullivan G (2007) An introduction to the HDPhoto technical design, JPEG document WG 1 N4183.

[9] ISO/IEC SC29 WG1: “Proposal for a New Work Item-JPEG Extensions,” JPEG document, WG1N6164, 2012.

[10] Choi S, Kwon O-J, Jang D, Choi S (2015) Performance evaluation of JPEG XT standard for high dynamic range image coding. Proceedings of the World Congress on Engineering, July, London, U.K.

[11] Watson A (1993) DCT quantization matrices visually optimized for individual images, Proc. SPIE 1913, Human Vision, Visual Processing, and Digital Display IV, 202 (September 8, 1993)

[12] Hontsch I, Karam L J (2002) Adaptive image coding with perceptual distortion control, IEEE Transactions on Image Processing, 11(3): 213-222.

[13] Oizumi M (2006) Preprocessing method for DCT based image compression. IEEE Transactions on Consumer Electronics, 52: 1021-1026.

[14] Bovik AC (2009) The essential guide to image processing. United States of America: Academic Press, pp. 429-432.

[15] Babu RV, Tom M, Wadekar P (2016) A survey on compressed domain video analysis techniques. Multimedia Tools and Applications, 75:1043-1078.

[16] Zadeh PB, Akbari AS, Buggy T, Soraghan J (2010) Multiresolution HVS and statistically based image coding scheme. Multimedia Tools and Applications, 49: 347-370.

[17] Andra K, Chakrabarti C, Acharya T (2003) A high-performance JPEG2000 architecture. IEEE Trans Circuits Syst Video Technol 13(3): 209–218.

[18] Malo J, Epifanio I, Navarro R, Simoncelli EP (2006) Nonlinear image representation for efficient perceptual coding. IEEE Transactions on Image Processing, 15(1): 068-080.

[19] Zhu Y, Yuan J (2014) A bit allocation optimization method for ROI based image compression with stable image quality. International Conference on Pattern Recognition, pp. 849-854.

[20] Subramanian P, Alamelu NR, Aramudhan M (2011) Bit rate control schemes for ROI video coding. International Journal of Computer Theory and Engineering, 3: 628-631.

[21] Shao F, Jiang G, Lin W, Yu M, Dai Q (2013) Joint bit allocation and rate control for coding multi-view video plus depth based 3D video. IEEE Trans. Multimedia, 15: 1843–1454.

[22] Cheung G, Velisavljevic V, Ortega A (2011) On dependent bit allocation for multiview image coding with depth-image-based rendering. IEEE Transactions on Image Processing, 20: 3179–3194.

[23] Rajaei B, Maugey T, Pourreza H, Frossard P (2013) Rate-distortion analysis of multiview coding in a DIBR framework. Annals of Telecommunications, 68: 627–640.

[24] Gelman A, Dragotti PL, Velisavljevic V (2012) Multiview image coding using depth layers and an optimized bit allocation. IEEE Transactions on Image Processing, 21: 4092–4105.

[25] Hachicha W, Kaaniche M, Beghdadi A, Cheikh FA (2014) Rate distortion optimal bit allocation for stereo image coding. International Conference on Image Processing (ICIP 2014), pp. 5661-5665.

[26] Cheng H, Lerner C (2015) Bit allocation for lossy image set compression. IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (*PACRIM 2015*), pp. 052-057.

[27] Lei J, Li Z, Zhu T, He X, You L, Hou C (2016) Region-based bit allocation and rate control for depth video in HEVC. Multimedia tools and applications, pp. 001-017.

[28] Gao W, Kwong S, Yuan H, Wang X (2016) DCT coefficient distribution modeling and quality dependency analysis based frame-level bit allocation for HEVC. IEEE Transactions on Circuits and Systems for Video Technology, 26: 139-153.

[29] Pan D (1995) A tutorial on MPEG/audio compression. IEEE MultiMedia, 2(2): 060-074.

[30] Chen H, Yu TL (2005) Comparison of psychoacoustic principles and genetic algorithms in audio compression. Proceedings of the 18th International Conference on Systems Engineering, pp. 270-275.

[31] Beghdadi A, Larabi M-C, Bouzerdoum A, Iftekharuddin KM (2013) A survey of perceptual image processing methods. Signal Processing: Image Communication, 28(8): 811-831.

[32] Abu NA, Ernawan F (2015) A novel psychovisual threshold on large DCT for image compression. The Scientific World Journal, 2015: 001-011.

[33] Rodriguez-Sanchez R, Martinez-Baena J, Garrido A, Garcia JA, Fdez-Valdivia J, Aranda MC (2002) computer vision group, University of Granada. http://decsai.ugr.es/cvg/dbimagenes/c512.php

[34] Hunt O, Mukundan R (2004) A comparison of discrete orthogonal basis functions for image compression. Proceedings of the Conference on Image and Vision Computing New Zealand (IVCNZ-2004), pp. 053-058.

[35] Ahmed N, Natrajan T, Rao KR (1993) Discrete cosine transform. IEEE Transactions on Computers, C-23(1): 090-093.

[36] Britanak V (2010) New fast computational structures for an efficient implementation of the forward/backward MDCT in MP3 audio coding standard. Signal Processing, 90(2): 536-547.

[37] Patchoo W, and Fischer TR (2010) Rate-distortion modeling of spherical vector quantization performance in transform excitation audio coding. International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology, pp. 1075-1079.

[38] Wu G, Yang E-H (2010) A new efficient method of computing MDCT in MP3 audio coding. International Conference on Multimedia Information Networking and Security, pp. 063-067.

[39] Gockel H, Moore BCJ, Patterson RD (2002) Asymmetry of masking between complex tones and noise: the role of temporal structure and peripheral compression. Journal of Acoustical Society of America, 111(6): 2759-2770.

[40] Abu NA, Ernawan F (2014) Psychovisual threshold on large Tchebichef moment for image compression. Applied Mathematical Sciences, 8: 6951-6961.

[41] Painter T, Spanias A (2000) Perceptual coding of digital audio. Proceedings of the IEEE, pp. 451-515.

[42] Kotz S, Nadarajah S (2000) Extreme value distributions: theory and applications, London: Imperial College Press.

[43] Hore A, Ziou D (2010) Image quality metrics: PSNR vs. SSIM. Proceedings of the International Conference on Pattern Recognition, pp. 2366-2369.

[44] Chen Y-K, Cheng F-C, Tsai P (2011) A gray-level clustering reduction algorithm with the least PSNR. Expert Systems with Applications, 38(8): 10183-10187.

[45] Yim C, Bovik AC (2011) Quality assessment of deblocked images. IEEE Transactions on Image Processing, 20(1): 088-098.

[46] Wallace G (1992) The JPEG still picture compression standard. IEEE Transactions on Consumer Electronics, 38(1): xviii–xxxiv.

[47] Zhang X, Wang S, Gu K, Lin W, Ma S, Gao W (2017) Just-noticeable difference-based perceptual optimization for JPEG compression. IEEE Signal Processing Letters, 24(1): 96-100.

[48] Ernawan F, Nugraini SH (2014) The optimal quantization matrices for JPEG image compression from psychovisual threshold. Journal of Theoretical and Applied Information Technology, 70(3): 566-572.

[49] Rahmalan H, Abu NA, Wong SL (2010) Using tchebichef moment for fast and efficient image compression. Pattern Recognition and Image Analysis, 20(4): 505-512.

[50] Ernawan F, Abu NA, Suryana N (2014) An optimal tchebichef moment quantization using psychovisual threshold for image compression. Advanced Science Letters, 20(1): 070-074.

CHAPTER 4

Robust Image Watermarking Based on Psychovisual Threshold

Ferda Ernawan

Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang, Lebuhraya Tun Razak 26300 Gambang Kuantan, Pahang, Malaysia
e-mail: ferda@ump.edu.my

Abstract. Because of the facility of accessing and sharing digital images through the internet, digital images are often copied, edited and reused. Digital image watermarking is an approach to protect and manage digital images as intellectual property. The embedding of a natural watermark based on the properties of the human eye can be utilized to effectively hide a watermark image. This paper proposes a watermark embedding scheme based on the psychovisual threshold and edge entropy. The sensitivity of minor changes in DCT coefficients against JPEG quantization tables was investigated. A watermark embedding scheme was designed that offers good resistance against JPEG image compression. The proposed scheme was tested under different types of attacks. The experimental results indicated that the proposed scheme can achieve high imperceptibility and robustness against attacks. The watermark recovery process is also robust against attacks.

Keywords: *image watermarking; imperceptibility; modified entropy; psychovisual threshold; robustness; watermark embedding; watermarking scheme.*

1. Introduction

Nowadays, multimedia data such as images are easily converted into digital content. The protection of intellectual properties in the form of digital images faces very serious challenges such as piracy, illegal redistribution, forgery and theft [1]. These challenges make digital image watermarking an important issue, as it protects against unauthorized duplication of digital images. Image watermarking means to embed a watermark without degrading the perceptual image quality and at the same time making it difficult to remove [2]. Image watermarking should be able to comply with imperceptibility, robustness and security. Embedding and extracting the watermark image should be limited to authorized users only.

In modern digital image watermarking, the watermark insertion process exploits the characteristics of the human visual system (HVS). The watermark can be inserted in a redundant region of the HVS [3], especially in highly textured areas and significantly changing regions of an image or the image edges. The characteristics of the human visual system have previously been used in image watermarking applications [4,5]. Some image watermarking schemes were designed based on visual models of colour stimuli [6,7] or were derived from image compression to increase their robustness [8]. Embedding watermarks based on HVS properties is able to improve watermark robustness while still maintaining their imperceptibility. The watermark scheme in [9] utilizes the spatial masking principle of HVS to improve the watermark's strength. The HVS threshold has been utilized for improving the watermark's robustness for authentication [10] and protection [11].

In this paper, a specific location is proposed for embedding watermarks based on the psychovisual threshold. The contribution of the DCT coefficient to the reconstruction error was measured in natural and graphical images and analysed as an initial psychovisual threshold [12]. This threshold can be utilized to determine the location and strength of the watermark. In the proposed method, watermark embedding under the constraint of the psychovisual threshold was chosen in order for the watermark to be invisible to the human visual system and produces only imperceptible distortion. The entropy and edge entropy of each image block of the host image is considered to identify the region most suitable for watermark embedding. The watermark is inserted in a block with a minimum amount of edge entropy based on the psychovisual threshold. Watermark insertion based on the entropy of image pixels can improve the watermark's imperceptibility and robustness [13].

2. Psychovisual Threshold

A true-colour 24-bit image is converted to the YUV colour model. The advantage of the YUV colour model is that it separates chromatic and achromatic components. A YUV colour model consists of luminance Y, chrominance U and chrominance V, which have identical characteristics. The two-dimensional discrete cosine transform (DCT) is used to transform each component. The characteristics of DCT coefficients against reconstruction errors prescribe the psychovisual threshold for luminance and chrominance as shown in Figures 1 and 2 respectively. A green and a blue curve represent the average error reconstruction based on the minimum and maximum JPEG quantization values for each frequency order, respectively.

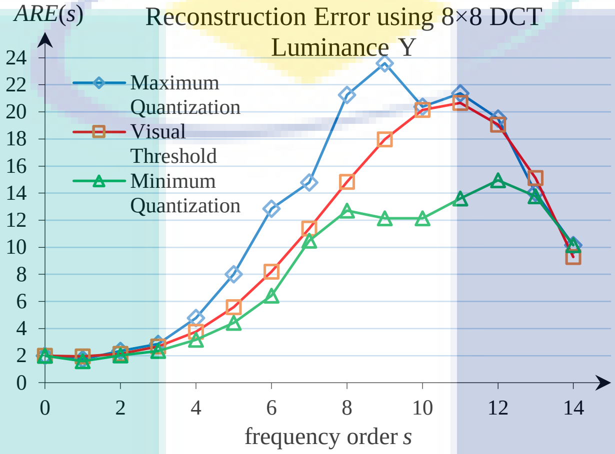


Figure 1. Average reconstruction for 40 real images error resulting from an increment of the DCT coefficient on the luminance.

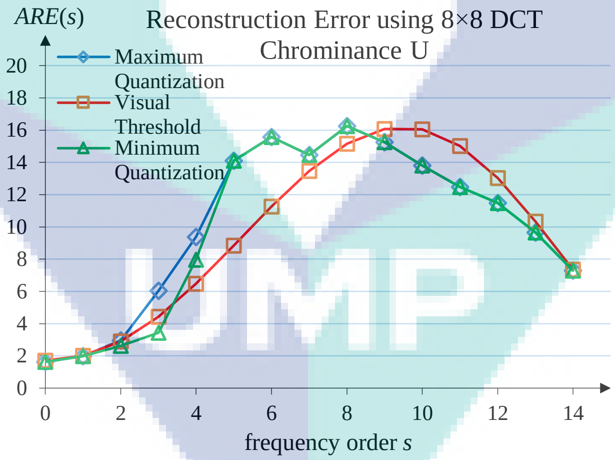


Figure 2. Average reconstruction error for 40 real images resulting from an increment of the DCT coefficient on the chrominance.

The average reconstruction error from an increment of the DCT coefficient on chrominance U is similar to the average reconstruction error on chrominance V. The sensitivity of the DCT coefficients on each frequency order against reconstruction errors produces an acceptable visual quality for the human visual system. The psychovisual threshold is set as a smooth transition curve of average error reconstruction as depicted by the red curve. According to Figures 1 and 2, the area under the psychovisual error threshold has potential resistance against JPEG quantization tables in image compression. In previous works, the psychovisual threshold has been applied to several image processing applications, such as image compression [14-19], adaptive image compression [20,21] and image watermarking [12].

3. Embedding Location

The location of the embedding bits of a watermark in the low frequency order was chosen because it has more resistance against JPEG quantization tables in image compression. The loopholes of JPEG quantization tables are identified by differentiating between the average reconstruction error of the psychovisual threshold and the default 8×8 JPEG quantization tables. These gaps can be computed as follows in Eqs. (1) and (2):

$$Q_{GL} = Q_{VL} - Q_{CL} \tag{1}$$

$$Q_{GR} = Q_{VR} - Q_{CR} \tag{2}$$

The new quantization tables Q_{VL} and Q_{VR} for luminance and chrominance based on the psychovisual threshold are shown in Figure 3. The locations of loopholes $C_{5,1}$, $C_{4,2}$, $C_{3,3}$ luminance and $C_{3,2}$, $C_{2,2}$, $C_{2,3}$ chrominance in the JPEG quantization tables based on the psychovisual error threshold are indicated by the blackened cells:

16	14	13	15	19	28	37	55
14	13	15	19	28	37	55	64
13	15	19	28	37	55	64	83
15	19	28	37	55	64	83	103
19	28	37	55	64	83	103	117
28	37	55	64	83	103	117	117
37	55	64	83	103	117	117	111
55	64	83	103	117	117	111	90

18	18	23	34	45	61	71	92
18	23	34	45	61	71	92	92
23	34	45	61	71	92	92	104
34	45	61	71	92	92	104	115
45	61	71	92	92	104	115	119
61	71	92	92	104	115	119	112
71	92	92	104	115	119	112	106
92	92	104	115	119	112	106	100

Figure 3. Location of the embedded watermark within 8×8 DCT coefficients for luminance (left) and chrominance (right) of new quantization tables Q_{VL} and Q_{VR} based on the psychovisual threshold.

The watermark is expected to survive better in these locations against JPEG quantization tables Q_{CL} and Q_{CR} for luminance and chrominance, respectively. Watermark insertion in the blackened cell locations will not produce a significantly high quality degradation of the watermarked image. The entropy and edge entropy of the image pixels are employed to select the region block for watermark embedding. The entropy is used to measure the spatial correlation as defined by Eq. (3):

$$E = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{3}$$

where p_i denotes the occurrence probability of an event i with $0 \leq p_i \leq 1$ and $\sum_{i-1}^{n} p_i = 1$. Accordingly, the entropy together with the edge entropy of each block are considered to identify the block suitable for embedding. The edge entropy is defined as follows in Eq. (4):

$$E_{edge\_entropy} = \sum_{i=1}^{n} p_i \exp^{u_i} = \sum_{i=1}^{n} p_i \exp^{1-p_i} \tag{4}$$

where $u_i = 1 - p_i$ indicates the ignorance or uncertainty of the image pixels. The two measures of entropy of each block are then summed up and the values thus obtained are sorted in ascending order. The blocks with low entropy values are selected for watermark embedding until the number of selected blocks is equal to the watermark size.

Five true-colour images were selected as the host images to evaluate the watermarking scheme, i.e. “Baboon”, “Pepper”, “Boat”, “Airplane” and “Lena” [22]. The original high-fidelity images of size 512×512 pixels are shown in Figure 4.

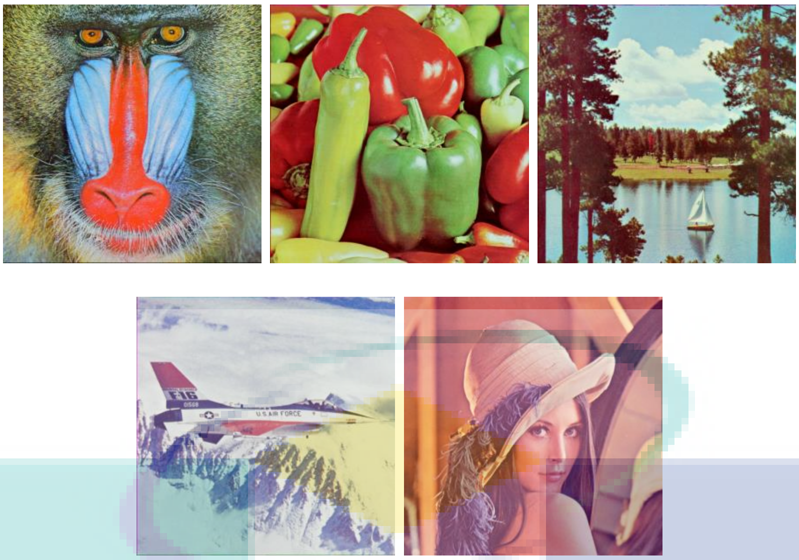


Figure 4. Original images “Baboon”, “Pepper”, “Boat”, “Airplane” and “Lena”.

4. Experimental Method

Embedding watermark schemes in a low frequency order will produce a higher quality degradation of the watermarked image. In 2002, Fridrich, *et al.* described how inserting bits of quantized DCT coefficients corresponding to medium frequencies provides a spare space to carry additional data against lossless image compression [23]. Embedding a watermark in a high frequency order makes the watermark less robust, with a higher probability of being lost if the watermarked image is compressed [24]. The watermark is inserted in the loopholes of the JPEG quantization tables because it has a complex texture area or edge in each block of the image. The human visual system is less sensitive to the edges of an image object [25]. The image watermarking scheme over edge entropy makes it possible to embed perceptually invisible watermarks and to make them more robust against attacks. A trade-off between robustness and imperceptibility is expected.

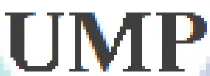


Figure 5. Original watermark image consisting of 25×75 pixels.

In this work, the Mersenne twister method was used to generate random numbers based on a secret key. The secret key was employed to encrypt and decrypt the watermark during insertion and extraction. The binary watermark W (“UMP”) with a size of 25×75 pixels is shown in Figure 5.

5.1 Watermark Insertion

A host image is first divided into an 8×8 block image. Then the entropy and edge entropy are used to select the region block for watermark embedding. The suitable block is then transformed by the two-dimensional DCT and is embedded through random numbers in specific locations based on the psychovisual threshold. A random number selects the loophole positions in each 8×8 DCT block for inserting bits of the watermark. The watermark that is embedded in the host image is subjected to JPEG quantization values. The quantization value that is used in the embedding process is given as follows in Eq. (5):

$$W_{QL} = \{18, 17, 16\} \text{ and } W_{QCR} = \{21, 26, 26\} \quad (5)$$

The embedded watermarks for luminance are randomized by a private key, as given by Eq. (6):

$$\begin{array}{ll} C_{5,1} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=0 \text{ and } RNG(i,3)=0 \\ C_{4,2} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=0 \text{ and } RNG(i,3)=1 \\ C_{3,3} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=1 \text{ and } RNG(i,3)=0 \\ C_{5,1} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=1 \text{ and } RNG(i,3)=1 \\ C_{4,2} & \text{if } RNG(i,1)=1 \text{ and } RNG(i,2)=0 \text{ and } RNG(i,3)=0 \\ C_{5,1} & \text{if } RNG(i,1)=1 \text{ and } RNG(i,2)=0 \text{ and } RNG(i,3)=1 \\ C_{4,2} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=1 \text{ and } RNG(i,3)=0 \\ C_{3,3} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=1 \text{ and } RNG(i,3)=1 \end{array} \tag{6}$$

The embedded watermarks for chrominance are also randomized by a similar private key, as given by Eq. (7):

$$\begin{array}{ll} C_{3,2} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=0 \text{ and } RNG(i,3)=0 \\ C_{2,2} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=0 \text{ and } RNG(i,3)=1 \\ C_{2,3} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=1 \text{ and } RNG(i,3)=0 \\ C_{3,2} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=1 \text{ and } RNG(i,3)=1 \\ C_{2,2} & \text{if } RNG(i,1)=1 \text{ and } RNG(i,2)=0 \text{ and } RNG(i,3)=0 \\ C_{3,2} & \text{if } RNG(i,1)=1 \text{ and } RNG(i,2)=0 \text{ and } RNG(i,3)=1 \\ C_{2,2} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=1 \text{ and } RNG(i,3)=0 \\ C_{2,3} & \text{if } RNG(i,1)=0 \text{ and } RNG(i,2)=1 \text{ and } RNG(i,3)=1 \end{array} \tag{7}$$

The calculation of the watermark quantity is given as follows in Eq. (8):

$$Q(i) = T \cdot W_Q \tag{8}$$

where the watermark weight to be embedded depends on threshold input T. Consider that when watermark $W = 1$, the watermark image is multiplied by "+1", whereas when watermark $W = 0$, it is multiplied by "−1" or subtracted from the host image.

The main steps of the embedding procedure can be described as follows:

Step 1: Take the host image block as input (block size is 8×8 pixels).
Step 2: Calculate the entropy and edge entropy of each image block to identify the block suitable for insertion. The two measures of entropy for each block are then summed up and the values thus obtained are sorted in ascending order. The block with the lowest value is selected for embedding until the number of selected blocks is equal to the watermark size.
Step 3: Transform the selected image block by the 2-D DCT.
Step 4: Generate a unique random number based on the secret key. The sequence value belongs to the set {0, 1}.
Step 5: Determine the selected location for watermark insertion based on a random number generator (RNG).
Step 6: Embed −1 or +1 into the selected location when the watermark value is 0 or 1 respectively.

The difference between the watermarked image and the original image was enhanced and shown in Figure 6.

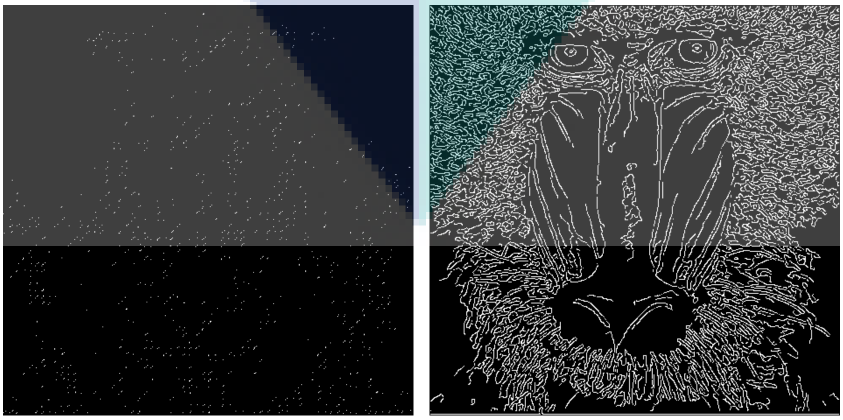


Figure 6. Enhanced embedding location based on the entropy of the "Baboon" image (left) and the edge entropy of the "Baboon" image (right).

5.2 Watermark Extraction

The watermark is extracted from the host image based on using the entropy and edge entropy to determine the selected block where the watermark is embedded. The watermark image is dispersed randomly on each selected block of the image based on the entropy and edge entropy. Extraction of the watermark involves a secret key to generate pseudo-random numbers. The watermark is detected by computing the correlation between the watermarked image and the watermark code.

The main steps of watermark extraction can be described as follows:

Step 1: Select the image blocks with low entropy values. The blocks with low entropy values are selected for extracting the watermark until the number of selected blocks is equal to the watermark size.
Step 2: Transform the image block by the 2-D DCT in the image block.
Step 3: Generate pseudo-random numbers with the same private key. These random numbers are used to find the location of the embedded watermark.
Step 4: Extract the watermark using an inner product algorithm. In order to extract the extracted sequence of the $X^*\{x^*(i), (1 < i < N)\}$, where $x^*(i) \geq 1$ means that the watermark is 1 and $x^*(i) \leq 0$ means that the watermark is 0.

A correlation coefficient is used to determine the watermark image. The correlation coefficient can be computed as follows in Eq. (9):

$$\rho = X \cdot X^* \tag{9}$$

where $X{\cdot}X^*$ is the inner product of X and the extracted sequence of X^*. If the correlation coefficient between watermarked image X and extracted sequence X^* is larger than a certain threshold, it is determined that the watermark exists.

4.3 Watermark Image Evaluation

The concealment of the watermark image was evaluated by peak signal to noise ratio (PSNR) and normalized cross-correlation (NC). The PSNR is defined as followsin Eq. (10) [26]:

$$PSNR = 10\log_{10}\left(\frac{255^2}{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\sum_{k=0}^{2}\left\|g(i,j,k)-f(i,j,k)\right\|^2}\right) \tag{10}$$

where $g(i, j, k)$ represents the watermarked image, $f(i, j, k)$ represents the original host image, and k is the third index referring to the three RGB colors. PSNR is generally deployed for comparing imperceptibility performance [27]. The comparison between the recovered watermark and the original watermark was quantitatively analysed using the NC [28], which is defined as follows in Eq. (11):

$$NC = \frac{\sum_{i=1}^{K}\sum_{j=1}^{L} W(i,j).W'(i,j)}{\sqrt{\sum_{i=1}^{K}\sum_{j=1}^{L} W(i,j)^2 \sum_{i=1}^{K}\sum_{j=1}^{L} W'(i,j)^2}} \tag{11}$$

where $W(i, j)$ is the original watermark image and (i, j) is the recovered watermark image. $K{\times}L$ is the watermark image size and the value of NC is between 0 and 1. A higher value of NC means that the recovered watermark image is closer to the original watermark image.

6. Experimental Method

The watermarked images were tested with a number of attacks to evaluate the watermarking scheme's performance. The visual effects on the watermarked "Baboon" image and the corresponding extracted watermark under different types of attacks are shown in Figures 7 and 8, respectively. The extracted watermark image can be damaged but it can still be seen by the human eye.

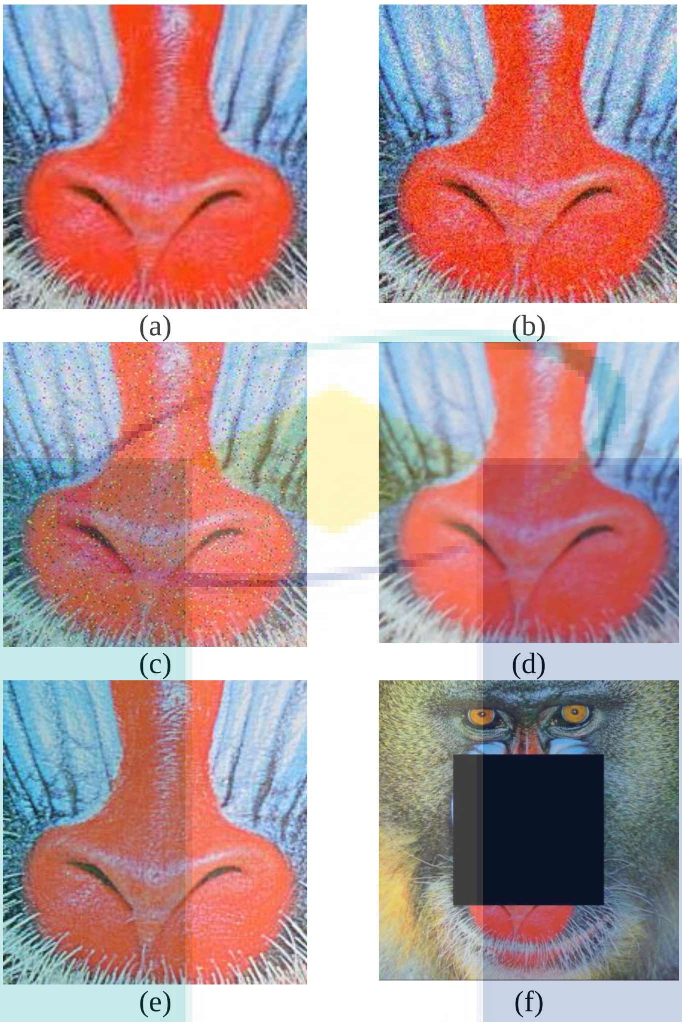
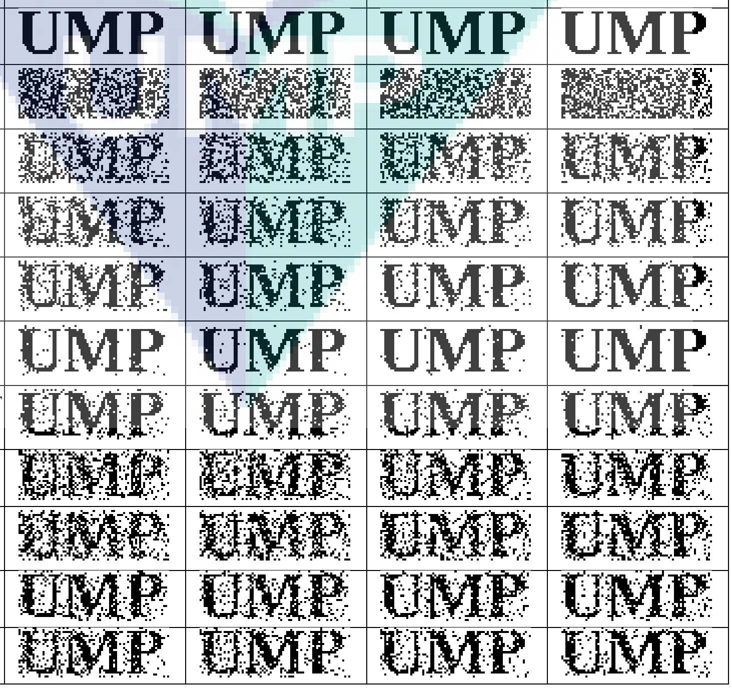


Figure 7. Attacked watermarked "Baboon" image: (a) JPEG compression, NC = 0.961; (b) Gaussian white noise 0.01, NC = 0.874; (c) salt and pepper noise 0.05, NC = 0.827; (d) median filter [3 3], NC = 0.944; (e) sharpening, NC = 0.991; and (f) cropping 25%, NC = 0.897.

Attacks	**Baboon**			
	***T* = 0.25**	***T* = 0.5**	***T* = 0.75**	***T* = 1**
No attack				
JPEG compression Q = 10				
JPEG compression Q = 30				
JPEG compression Q = 50				
JPEG compression Q = 70				
JPEG compression Q = 90				
Gaussian low pass filter [3 3]				
Gaussian noise 0.01				
Salt and pepper noise 0.05				
Median filter [3 3]				
Speckle noise 0.01				



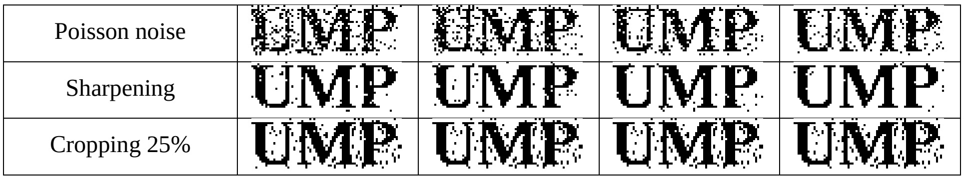


Figure 8. Visual quality comparison of the extracted watermarks for the "Baboon" image under different types of attacks.

Table 1. PSNR performance under different threshold values.

Threshold	Baboon	Pepper	Boat	Airplane	Lena
1	44.912	44.894	44.898	44.895	44.922
0.75	47.411	47.393	47.397	47.394	47.421
0.5	50.932	50.915	50.919	50.916	50.943
0.25	56.953	56.935	56.939	56.936	56.963

Table 2. Full Error performance under different threshold values.

Threshold	Baboon	Pepper	Boat	Airplane	Lena
1	0.673	0.679	0.687	0.676	0.670
0.75	0.505	0.509	0.515	0.507	0.503
0.5	0.337	0.340	0.343	0.338	0.335
0.25	0.169	0.171	0.172	0.169	0.168

Table 3. NC after different types of attacks on Watermarked Image under different threshold values.

Attacks	***T* = 0.25**		***T* = 0.5**		***T* = 0.75**		***T* = 1**	
	Baboon	**Pepper**	**Baboon**	**Pepper**	**Baboon**	**Pepper**	**Baboon**	**Pepper**
No attack	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
JPEG compression Q = 10	0.609	0.584	0.607	0.568	0.604	0.577	0.601	0.586
JPEG compression Q = 30	0.792	0.660	0.802	0.687	0.822	0.777	0.851	0.859
JPEG compression Q = 50	0.846	0.753	0.867	0.809	0.904	0.914	0.923	0.961
JPEG compression Q = 70	0.901	0.825	0.923	0.902	0.953	0.957	0.957	0.981
JPEG compression Q = 90	0.972	0.952	0.977	0.975	0.986	0.991	0.989	0.996
Gaussian low pass filter [3 3]	0.927	0.913	0.928	0.925	0.932	0.949	0.939	0.966
Gaussian noise 0.01	0.813	0.727	0.825	0.763	0.847	0.832	0.877	0.874
Salt and pepper noise 0.05	0.799	0.718	0.820	0.748	0.850	0.794	0.862	0.827
Median filter [3 3]	0.891	0.892	0.889	0.890	0.906	0.921	0.912	0.944
Speckle noise 0.01	0.875	0.821	0.890	0.860	0.913	0.924	0.924	0.949
Poisson noise	0.900	0.837	0.923	0.881	0.929	0.948	0.935	0.966
Sharpening	0.980	0.971	0.981	0.977	0.984	0.988	0.992	0.991
Cropping 25%	0.939	0.894	0.939	0.895	0.940	0.896	0.940	0.897

From Tables 1 and 2, it can be seen that the threshold used for embedding the watermark has a significant effect on the imperceptibility of the watermark. A larger threshold makes the embedded watermark more robust while it results in a lower quality of the watermarked image. Table 3 shows the NC comparison of the watermarked image under different threshold values. The watermarked image underwent different types of attacks. The experimental results indicate that the proposed scheme has a great resistance to format-compression attack: JPEG compression; denoising attacks: median filter, Gaussian low pass filter; noise attacks: Gaussian noise, salt and pepper noise, Poisson noise, speckle noise; image processing attacks: sharpening; geometrical attacks: cropping. This embedding watermark scheme provides perceptual invisibility to the human visual system and robustness against attacks. Embedding the watermark in deep-hole locations of the JPEG quantization tables makes it resistant against JPEG quantization tables in image compression and its location does not have a significant impact on the quality of the image reconstruction.

7. Conclusions

Digital image watermarking is useful for preventing unauthorized duplication of digital images. This research investigated watermark embedding in the lowest DCT psychovisual threshold based on entropy and edge entropy. The DCT psychovisual threshold indicates loopholes $C_{5,1}$, $C_{4,2}$, $C_{3,3}$ luminance and $C_{3,2}$, $C_{2,2}$, $C_{2,3}$ chrominance in the JPEG quantization tables, respectively. Embedding a watermark image in those loopholes of the JPEG quantization tables provides great resistance against JPEG compression. The proposed embedding watermark scheme was tested under different types of attacks. The experimental results showed that the watermarked images had high imperceptibility and the watermark recovery was robust against different types of attacks.

Acknowledgments

The authors would like to express very special thanks to Universiti Malaysia Pahang, Malaysia for providing financial support of this research project by UMP Research Grant Scheme (RDU160364).

References

[1] Halder, R., Pal, S. & Cortesi, A., *Watermarking Techniques for Relational Databases: Survey, Classification and Comparison*, Journal of Universal Computer Science, **16**(21), pp. 3164-3190, 2010.

[2] Cui, L. & Li, W., *Adaptive Multiwavelet-Based Watermarking through JPW Masking*, IEEE Transactions on Image Processing, **20**(4), pp. 1047-1060, 2011.

[3] Yang, Y., Sun, X., Yang, H., Lie, C.T. & Xiao, R., *A Contrast-Sensitive Reversible Visible Image Watermarking Technique*, IEEE Transactions on Circuit and Systems for Video Technology, **19**(5), pp. 656-667, 2009.

[4] Niu, Y., Kyan, M., Krishnan, S. & Zhang, Q., *A Combined Just Noticeable Distortion Model-Guided Image Watermarking*, Signal, Image and Video Processing, **5**(4), pp. 517-526, 2011.

[5] Parthasarathy A.K. & Kak, S., *An Improved Method of Content Based Image Watermarking*, IEEE Transactions on Broadcasting, **53**(2), pp. 468-479, 2007.

[6] Barni, M., Bartolini, F. & Piva, A., *Multichannel Watermarking of Color Images*, IEEE Transactions on Circuits and Systems for Video Technology, **12**(3), pp. 142-156, 2012.

[7] Tsui, T.K., Zhang, X.P. & Androutsos, D., *Color image watermarking using Multidimensional Fourier Transforms*. IEEE Transactions on Information Forensics and Security, **3**(1), pp. 016-028, 2008.

[8] Lu, C.S., Huang, S.K., Sze, C.J. & Liao, H.Y.M., *Cocktail Watermarking for Digital Image Protection*, IEEE Transactions on Multimedia, **2**(4), pp. 209-224, 2000.

[9] Kutter, M. & Wingkler, S., *A Vision-Based Masking Model for Spread Spectrum Image Watermarking*, IEEE Transactions on Image Processing, **11**(1), pp. 016-025, 2002.

[10] Wei, Z. & Ngan, K.N., *Spatio-Temporal Just Noticeable Distortion Profile for Grey Scale Image/Video in DCT Domain*, IEEE Transactions on Circuits and Systems for Video Technology, **19**(3), pp. 337-346, 2009.

[11] Lu, C.S. & Liao, H.Y.M., *Multipurpose Watermarking for Image Authentication and Protection*, IEEE Transactions on Image Processing, **10**(10), pp. 1579-1592, 2001.

[12] Abu, N.A., Ernawan, F., Suryana, N. & Sahib, S., *Image Watermarking Using Psychovisual Threshold over the Edge*, Information and Communication Technology, **7804**, pp. 519-527, 2013.

[13] Maity, S.P. & Kundu, M.K., *DHT Domain Digital Watermarking with Low Loss in Image Informations*. International Journal of Electronics and Communications, **64**(3), pp. 243-257, 2010.

[14] Abu, N.A. & Ernawan, F., *A Novel Psychovisual Threshold on Large DCT for Image Compression*, The Scientific World Journal, pp. 001-011, 2015.

[15] Ernawan, F. & Nugraini, S.H., *The Optimal Quantization Matrices for JPEG Image Compression From Psychovisual Threshold*. Journal of Theoretical and Applied Information Technology, **70**(3), pp. 566-572, 2014.

[16] Abu, N.A. & Ernawan, F., *Psychovisual Threshold on Large Tchebichef Moment for Image Compression*, Applied Mathematical Sciences, **8**(140), pp. 6951-6961, 2014.

[17] Ernawan, F., Abu, N.A., & Suryana, N., *An Optimal Tchebichef Moment Quantization Using Psychovisual Threshold for Image Compression*, Advanced Science Letters, **20**(1), pp. 70-74, 2014.

[18] Abu, N.A., Ernawan, F. & Sahib, S., *Psychovisual Model on Discrete Orthonormal Transform*, International Conference on Mathematical Sciences and Statistics, Kuala Lumpur, Malaysia, pp. 309-314, 2013.

[19] Abu, N.A., Ernawan, F. & Suryana, N., *A Generic Psychovisual Error Threshold for the Quantization Table Generation on JPEG Image Compression*, 9th International Colloquium on Signal Processing and its Applications, Kuala Lumpur, Malaysia, pp. 39-43, 2013.

[20] Ernawan, F., Abu, N.A. & Suryana, N., *Adaptive Tchebichef Moment Transform Image Compression Using Psychovisual Model*, Journal of Computer Science, **9**(6), pp. 716-725, 2013.

[21] Ernawan, F., Abu, N.A. & Suryana, N., *An Adaptive JPEG Image Compression Using Psychovisual Model*, Advanced Science Letters, **20**(1), pp. 26-31, 2014.

[22] Rodriguez-Sanchez, R., Martinez-Baena, J., Garrido, A., Garcia, J.A., Fdez-Valdivia, J. & Aranda, M.C., *Computer Vision Group*, University of Granada, [2014-03-20], http://decsai.ugr.es/cvg/dbimagenes/c512.php.

[23] Fridrich, J., Goljan, M., Du, R., Lossless Data Embedding for All Image Formats, in Security and Watermarking of Multimedia Contents IV, Proceedings of SPIE, **4675**, pp. 572-583, 2002.

[24] Amer, I., Hishmat, P., Badawy W., Jullien, G., *Comparisons and Analysis of DCT-based Image Watermarking Algorithm*, Advanced Techniques in Computing Sciences and Software Engineering, pp. 055-058, 2010.

[25] Khalighi, S. & Rabiee, H., *A Contourlet-based Image Watermarking Scheme with High Resistance to Removal and Geometrical Attacks*. EURASIP Journal on Advances in Signal Processing, **2010**, pp. 001-013, 2010.

[26] Yim, C. & Bovik, A.C., *Quality Assessment of Deblocked Images*, IEEE Transactions on Image Processing, **20**(1), pp. 088-098, 2011.

[27] Wang, Z., Bovik, A.C., Sheikh, H.R. & Simoncelli, E.P., *Image Quality Assessment: From Error Visibility to Structural Similarity*, IEEE Transactions on Image Processing, **13**(4), pp. 600-612, 2004.

[28] You, X., Du, L., Cheung, Y. & Chen, Q., *A Blind Watermarking Scheme Using New Nontensor Product Wavelet Filter Banks*, IEEE Transactions on Image Processing, **19**(12), pp. 3271-3284, 2010.

CHAPTER 5

An Improved Imperceptibility and Robustness of 4x4 DCT-SVD Image Watermarking Using Modified Entropy

Ferda Ernawan, Mritha Ramalingam, Ali Safa Sadiq, Zuriani Mustaffa
Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang
Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang, Malaysia
e-mail: ferda@ump.edu.my

***Abstract*— A digital protection against unauthorized distribution of digital multimedia is highly on demand. Digital watermarking is a defence in multimedia protection for authorized ownership. This paper proposes an improved watermarking based on 4×4 DCT-SVD blocks using modified entropy in image watermarking. A modified entropy is used to select unnoticeable blocks. The proposed watermarking scheme utilizes the lowest entropy values to determine unnoticeable regions of watermarked image. This paper investigates on the relationship between $U_{(2,1)}$ and $U_{(3,1)}$ coefficients of the U matrix 4×4 DCT-SVD in image watermarking. The proposed watermarking scheme produces a great level of robustness and imperceptibility of watermarked image against different attacks. The proposed scheme shows the improvement in terms of structural similarity index and normalized correlation of watermarked image.**

***Index Terms*—modified entropy; discrete cosine transform; imperceptibility; watermark embedding; robustness.**

I. Introduction

The advancement of multimedia technology has contributed to the unauthorized distribution of digital images. This leads to provide ownership authentication and multimedia protection against unapproved redistributing multimedia data. Digital image watermarking is an alternative solution to preserve the ownership from distribution and duplication of digital images. Digital image watermarking is a method which aids in preventing the copying of digital data and protects it by imperceptibly hiding a mark that has authorized information into the original data.

Image watermarking has been designed in spatial or frequency methods. Image watermarking with directly altering pixels in watermarking scheme leads to easy and low computational cost [1–2]. Image watermarking schemes based on frequency domain produce more robust than image watermarking with spatial domain [3]. The embedded watermarks in frequency domain are distributed irregular over the image when it is inversely transformed in spatial domain. This scheme is able to improve watermark robustness while still maintaining their imperceptibility [4].

Frequency transforms have been applied in image watermarking schemes such as DCT [5–8], SVD [9][10], tchebichef moment transform (TMT) [11-12] and discrete wavelet transforms (DWT) [13]. DCT has been used in image compressions [14-17], image watermarking, steganography image and other image processing applications. DCT is used as the basis of digital watermarking due to its advantages such as high-energy compaction, less computational algorithms, high robustness and easy implementation in watermarking applications.

In other hands, DWT is more superior to DCT such as in [18]. However, in real application DWT requires high computational cost and uses wavelet transform. Thus, the DWT also produces shift invariant and it skips the down sampling process of each level in the DWT filtering. Otherwise, SVD is most commonly used as a transformation technique in watermarking because of its strong properties [19]. SVD is a technique for getting geometric features from an image. The combination of DCT and SVD can elevate the performance of watermarking scheme [3].

This paper proposes 4×4 DCT-SVD watermarking scheme using modified entropy. The modified entropy is used to select blocks of image which aids to achieve a maximum robustness and it does not produce un-noticeable distortion for embedded watermarks. A watermarking scheme based on 4×4 DCT-SVD is proposed to improve the capacity of embedded watermarks. The smaller selected blocks have potential to achieve high imperceptibility. The number of selected blocks using modified entropy has to match the number of watermark bits. This paper investigates the impact of embedded watermarks in the U matrix of 4×4 SVD in the first column in terms of quantization step. In addition, we apply image encryption to protect the rightful ownership.

II. Background

The characteristics of an image watermarking are its imperceptibility, robustness, watermark capacity, and watermark security [18]. The hybrid DCT-SVD has been developed in many digital image watermarking. The watermark insertion in the values (S) of the matrix SVD produces false positive issue. Researchers investigate the embedded watermarks in the values (U) or (V) of the matrix SVD [19-21].

Lai's watermarking scheme [21] has investigated the relation between $U_{3,1}$ and $U_{4,1}$ of U matrix on 8×8 DCT-SVD. This scheme allows attackers to easily extract the watermark image. In the watermarking recovery, there is no rightful ownership protection. Additionally, the actual owner is able to extract the watermark from arbitrary images. These issues should be considered before we apply Lai's scheme [21] for the rightful owners in watermark applications. Moreover, Lai's scheme has not been investigated the tradeoff between imperceptibility and robustness. The threshold in Lai's scheme plays a key role to determine the imperceptibility of watermarked images. The Lai's watermarking scheme can achieve a good quality level of PSNR values. The watermark's robustness produces a great level of bit correction rate (BCR). However, the watermark recovery is not resistance towards Gaussian noise and median filter.

This paper proposes 4×4 DCT-SVD watermarking scheme by investigating $U_{(2,1)}$ and $U_{(3,1)}$ coefficients of the U matrix 4×4 DCT-SVD. DCT-SVD is used instead of DWT, because DCT achieves less computational cost and has easy implementation towards the rightful owner's applications. The two-dimensional DCT of an image A is defined:

$$B_{pq} = \alpha_p \beta_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad (1)$$

for $p = 0, 1, 2, \ldots, M-1$ and $q = 0, 1, 2, \ldots, N-1$ where

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & p > 0 \end{cases} \qquad \beta_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & q > 0 \end{cases} \quad (2)$$

The inverse of discrete cosine transform can be computed by:

$$A_{pq} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \alpha_p \beta_q B_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad (3)$$

for $p,q = 0, 1, 2, 3$. The singular value decomposition of D is given by $D = USV^T$, where singular vectors U and V are orthogonal matrices and singular vector S is diagonal matrix (λ_i) of singular values λ_i=1, 2, 3, 4 arranged in decreasing order.

III. Proposed Watermarking Scheme

This experiments use four grayscale images from CVG-UGR image database [22]. Four grayscale images which have 512×512 pixels are shown in Figure 1. The proposed watermarking scheme is tested with two binary images of different sizes. The watermark images are depicted in Figure 2.

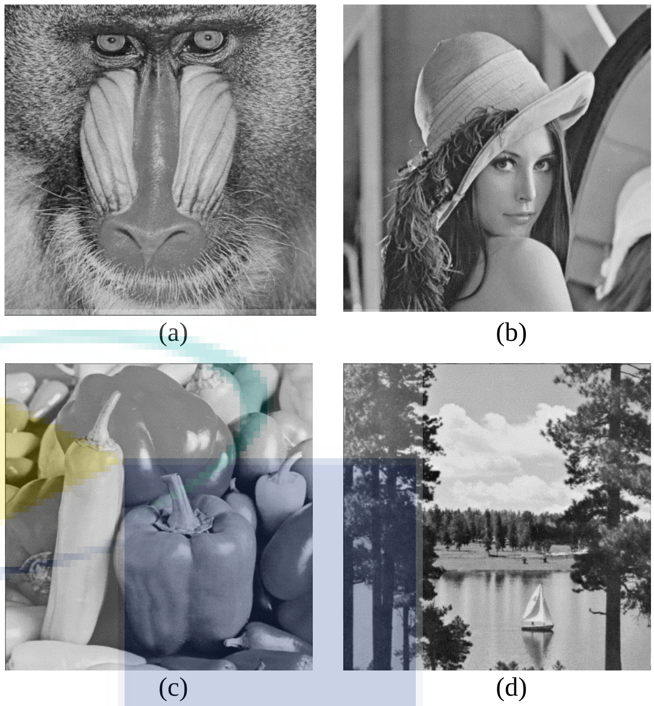


Figure 1: Four images (a) Baboon, (b) Lena, (c) Peppers, (d) Sailboat, 512×512, 8 bits/pixel.

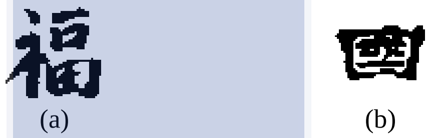


Figure 2: Original watermark image: (a) watermark image A with 48×48 pixels and (b) watermark image B with 32×48 pixels.

The binary watermark images of different sizes are used to be inserted into four host images. Before the watermark is being inserted, the watermark binary images are encrypted by applying logical XOR operation between a private key image generation and watermark binary image. A key image generation has unique random pixel values, 8 bits/pixel as shown in Figure 3. Only the actual owner that has a private key as presented by an image 8 bits/pixel can extract the watermark image. The insertion and extraction of watermarking process are depicted in Figures 4 and 5.

Figure 3: (a) encrypted watermark A, 48×48 pixels, 1 bit/pixel (b) key of watermark A, 48×48 pixels, 8 bits/pixel (c) encrypted watermark B, 32×48 pixels, 1 bit/pixel (d) key of watermark B, 48×32 pixels, 8 bits/pixel.

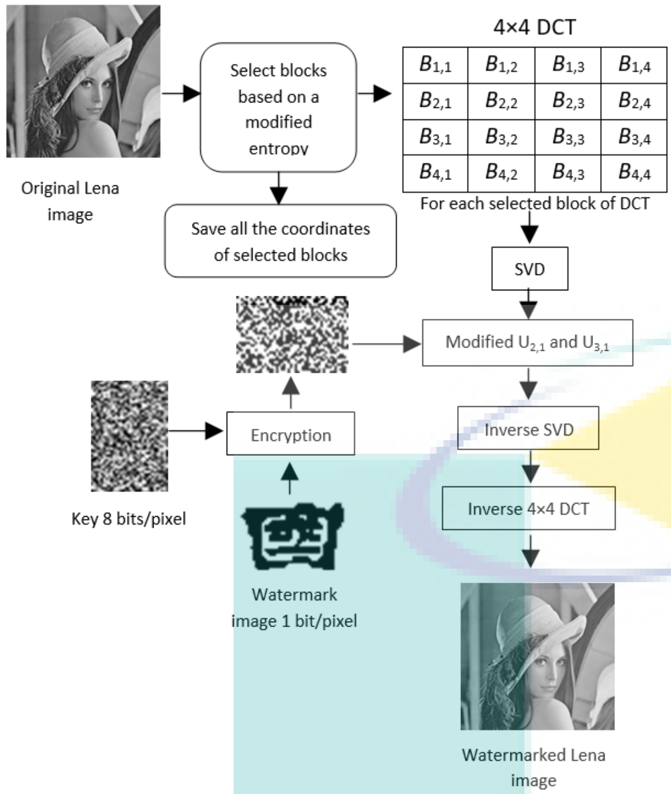


Figure 4: Insertion steps

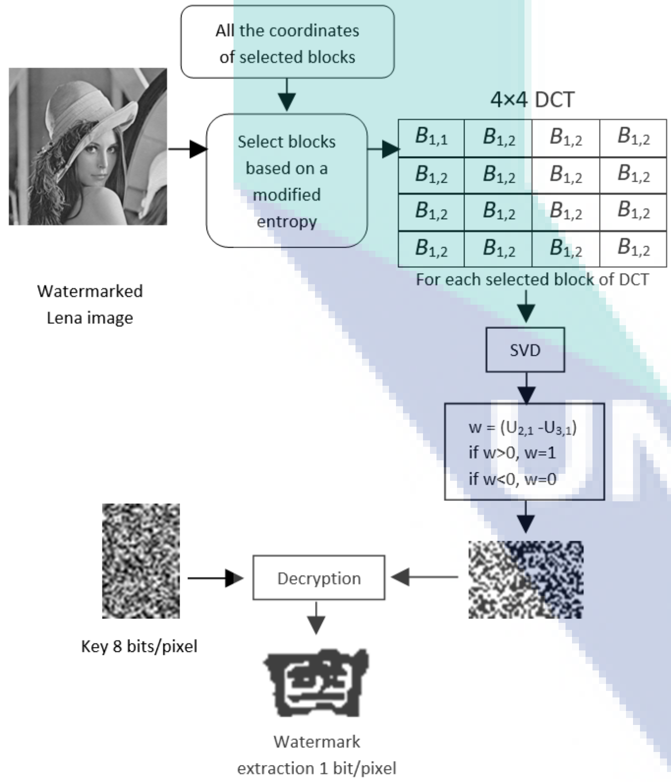


Figure 5: Extraction steps

A. *Watermark Insertion*

Embedding watermark sequence is given as follows:

Step1: Original image is first divided into 4×4 block image pixels.

Step2: Calculate the modified entropy for each image block. The modified entropy is defined by:

$$E = -\sum_{i=1}^{n} p_i \log_2(p_i) + p_i \exp^{1-p_i} \tag{4}$$

p_i denotes the occurrence probability of an event i with $0 \le p_i \le 1$. The values obtained from modified entropy are sorted, then the lowest values are utilized to select image blocks for embedded watermark.

Step3: The selected blocks are transformed by 4×4 DCT.

Step4: Apply SVD to 4×4 DCT coefficients. The results of SVD implementation by $U_{(4\times4)}$ is given as follows:

$$U = \begin{bmatrix} U_{(1,1)} & U_{(1,2)} & U_{(1,3)} & U_{(1,4)} \\ U_{(2,1)} & U_{(2,2)} & U_{(2,3)} & U_{(2,4)} \\ U_{(3,1)} & U_{(3,2)} & U_{(3,3)} & U_{(3,4)} \\ U_{(4,1)} & U_{(4,2)} & U_{(4,3)} & U_{(4,4)} \end{bmatrix} \tag{5}$$

For each selected block, the first column coefficients $U_{(2,1)}$ and $U_{(3,1)}$ are modified based on watermark binary values.

Step 5: For each selected block, $U_{(2,1)}$ and $U_{(3,1)}$ coefficients are changed and then compared to the threshold. If the watermark binary image is equal to 1, the coefficients of ($U_{(2,1)}$-$U_{(3,1)}$) must be a positive value and greater than threshold (T). Otherwise, if the watermark is 0, the relation ($U_{(2,1)}$-$U_{(3,1)}$) must be a negative value then it should greater than threshold (T). This conditions are violated, the coefficients of $U_{(2,1)}$ and $U_{(3,1)}$ must be modified based on the rules as given:

$$y = \frac{\left(\left|U_{(2,1)}\right| + \left|U_{(3,1)}\right|\right)}{2}, \tag{6}$$

$$\text{if } w_i = 1, \begin{cases} \tilde{U}_{(2,1)} = y + T/2 \\ \tilde{U}_{(3,1)} = y - T/2 \end{cases},$$

$$\text{if } w_i = 0, \begin{cases} \tilde{U}_{(2,1)} = y - T/2 \\ \tilde{U}_{(3,1)} = y + T/2 \end{cases},$$

where w_i denotes the watermark of i pixel with w equal to 0 or w equal to 1. $\tilde{U}$ represents a modified coefficient.

Step6: Inverse the SVD, then inverse the DCT for all selected blocks to generate watermarked image.

B. *Watermark Recovery*

Extracting watermark sequence is described as follows:

Step1: Watermarked image is first divided into 4×4 block image pixels.

Step2: Ascending values of modified entropy are used to identify the selected block for watermark embedding.

Step3: Apply 4×4 DCT to obtain DCT coefficients for each selected blocks.

Step4: Implement SVD to DCT coefficients of selected block.

Step5: The relation ($U_{(2,1)}$-$U_{(3,1)}$) of U matrix is calculated. If the result is positive, then the watermark recovery is 1, otherwise the watermark is 0.

C. *Watermarked Evaluation*

The watermarked imperceptibility is measured by structural similarity (SSIM) index and reconstruction errors. Reconstruction errors are calculated by measuring the difference between watermarked pixels and original pixels. SSIM is a method which measures the quality by capturing the similarity and it can be computed by:

$$SSIM(x,y) = [l(x,y)]^{\alpha} \cdot [c(x,y)]^{\beta} \cdot [s(x,y)]^{\gamma} \qquad (7)$$

where α>0, β>0, γ>0. A detail description can be found in [23]. Robustness of watermark recovery is estimated by normalized correlation (NC) that measures the correlation between the watermark pixels extraction and original watermark pixels. The NC is given as follows:

$$NC = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} W(i,j).W^{*}(i,j)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} W(i,j)^{2} \sum_{i=1}^{M}\sum_{j=1}^{N} W^{*}(i,j)^{2}}} \qquad (8)$$

where $W(i,j)$ represents original watermark pixels, $W^{*}(i,j)$ denotes watermark recovery and $M \times N$ represents the watermark size.

IV. Experimental Results

The different watermark sizes are embedded into four host images. The imperceptibility and robustness are determined by the quantization step, it measures the tradeoff between them. The robustness of watermarked image is measured by NC, whereas the imperceptibility is measured by SSIM. We investigate the optimal threshold of Lai's watermarking scheme and the proposed 4×4 DCT-SVD using modified entropy. The optimal threshold of Lai's scheme can be obtained by T = 0.019 as shown in Figure 6. The proposed watermarking scheme can achieve the tradeoff between robustness and imperceptibility with T = 0.42. Figure 7 shows the relationship between $U_{(2,1)}$ and $U_{(3,1)}$ on 4×4 DCT-SVD by quantization steps.

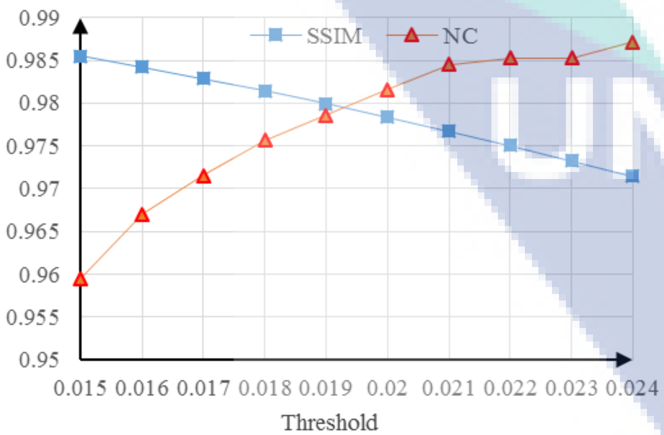


Figure 6: Relationship between imperceptibility and robustness of modified $U_{(3,1)}$ and $U_{(4,1)}$ on 8×8 DCT-SVD.

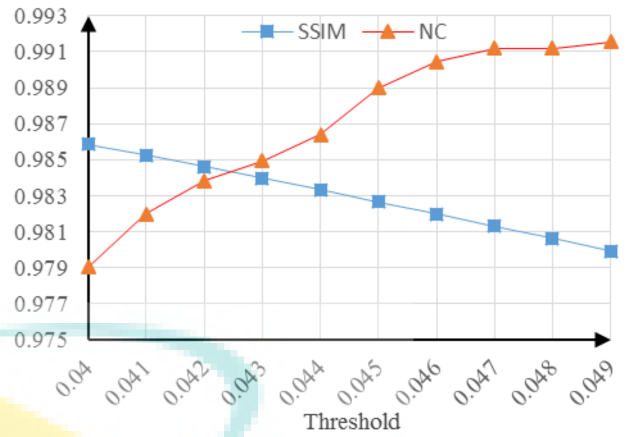


Figure 7: Relationship between imperceptibility and robustness of modified $U_{(2,1)}$ and $U_{(3,1)}$ on 4×4 DCT-SVD.

Table 1
Comparison of Lai's Scheme and Proposed Scheme in terms of ARE, PSNR, SSIM Using Watermark A

Images	Lai's Scheme [21] T = 0.019			Propose T = 0.042		
	ARE	PSNR	SSIM	ARE	PSNR	SSIM
Baboon	2.794	30.984	0.959	0.821	35.196	0.988
Lena	0.988	43.126	0.980	0.559	42.763	0.985
Pepper	1.084	39.918	0.977	0.506	43.097	0.987
Sailboat	1.524	38.066	0.965	0.712	40.784	0.976
Average	1.598	38.024	0.970	0.650	40.460	0.984

Table 2
Comparison of Lai's Scheme and Proposed Scheme in terms of ARE, PSNR, SSIM Using Watermark B

Images	Lai's Scheme [21] T = 0.019			Propose T = 0.042		
	ARE	PSNR	SSIM	ARE	PSNR	SSIM
Baboon	1.544	33.460	0.978	0.531	37.118	0.992
Lena	0.649	45.455	0.985	0.381	44.334	0.989
Pepper	0.654	43.866	0.985	0.343	44.682	0.991
Sailboat	0.860	42.686	0.976	0.486	42.349	0.983
Average	0.927	41.367	0.981	0.435	42.121	0.989

Tables 1 and 2 show ARE, PSNR, SSIM performances of Lai's scheme and our scheme using different watermark sizes. The quantitative measurement results show that our scheme produces less absolute reconstruction errors compared to Lai's scheme. SSIM values of the proposed scheme are higher than Lai's scheme. The embedded watermarks in smaller block produce high watermarked image quality. The proposed smaller blocks taken can increase the capacity of watermark bits.

Referring to Table 2, when the watermark size is small, it improves the reconstruction errors of watermarked image. The proposed scheme shows that the average of four watermarked images produce better image quality than Lai's scheme. The watermarked images are tested against seven types of attacks, e.g., Gaussian noise, JPEG compression, Gaussian low pass filter, salt & pepper, median filter, sharpening, and cropping. The comparison normalized correlation of watermark recovery is listed in Tables 3 and 4 with different sizes of watermark images.

Table 3
Comparison of Lai's Scheme and Proposed Scheme in terms of NC Using Watermark A

Attacks	Lai's Scheme [21] T = 0.019		Propose T = 0.042	
	Baboon	Lena	Baboon	Lena
Gaussian noise 0.001	0.890	0.708	0.973	0.981
JPEG compression	0.994	0.979	0.989	0.984
Gaussian low pass filter [3 3]	0.865	0.866	0.959	0.990
Salt and pepper noise 0.005	0.920	0.900	0.979	0.982
Median filter [3 3]	0.812	0.834	0.918	0.986
Sharpening	0.938	0.913	0.990	0.993
Cropping centre 12.5%	0.952	0.980	0.924	0.991
Average	0.910	0.883	0.962	0.987

Based on Tables 3 and 4, the proposed image watermarking scheme can perform in a good level of watermark recovery with different sizes of watermark images. It shows that the proposed watermarking scheme offer more robust than Lai's scheme. The visual watermark recovery after its attacks are shown in Figure 8. The proposed watermarking scheme can achieve great robustness from different types of attacks.

Figure 8: Attacked watermarked Lena images and watermark recovery: (a) Gaussian Noise 0.001, (b) JPEG compression, (c) Salt and Pepper 0.005, (d) Cropping centred 25%.

Table 4
Comparison of Lai's Scheme and Proposed Scheme in terms of NC Using Watermark B

Attacks	Lai's Scheme [21] T = 0.019		Propose T = 0.042	
	Baboon	Lena	Baboon	Lena
Gaussian noise 0.001	0.858	0.677	0.978	0.983
JPEG compression	0.988	0.991	0.990	0.983
Gaussian low pass filter [3 3]	0.894	0.840	0.962	0.991
Salt and pepper noise 0.005	0.912	0.866	0.970	0.978
Median filter [3 3]	0.840	0.801	0.914	0.989
Sharpening	0.943	0.881	0.994	0.988
Cropping centre 12.5%	0.920	0.985	0.906	0.995
Average	0.908	0.863	0.959	0.987

The comparison of the visual watermark recovery under different types of attacks between Lai's scheme and proposed watermarking scheme is shown in Table 5 and Table 6. The visual perceptions of watermark recovery show the proposed scheme produces more robust than Lai's watermarking scheme. Our watermark scheme has shown successful resistance against some attacks. In addition, the watermark extraction can be recovered with high visual quality.

Table 5
Visual Comparison of the Watermark Recovery Using Watermark A

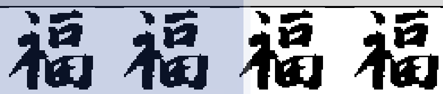
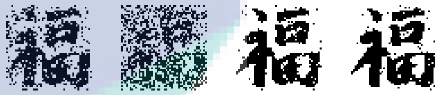
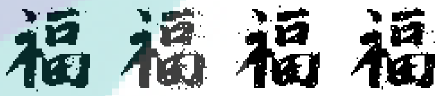
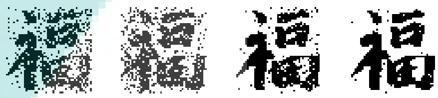
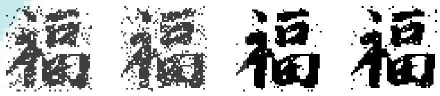
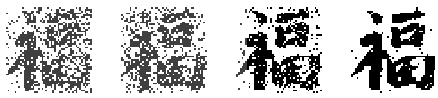
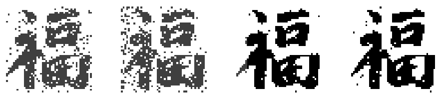
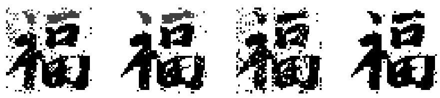
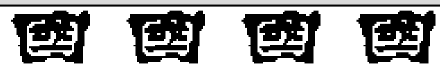
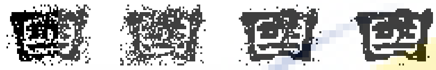
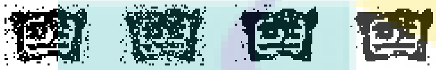
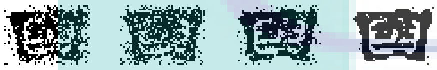
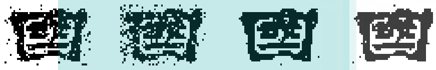
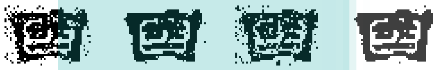
Attacks	Lai's Scheme [21] T = 0.019		Propose T = 0.042	
	Baboon	Lena	Baboon	Lena
Non-attack				
Gaussian noise 0.001				
JPEG compression				
Gaussian low pass filter [3 3]				
Salt and pepper noise 0.005				
Median filter [3 3]				
Sharpening				
Cropping centre 12.5%				

Table 6
Visual Comparison of the Watermark Recovery Using Watermark B

Attacks	Lai's Scheme [21] $T = 0.019$ Baboon	Lai's Scheme [21] $T = 0.019$ Lena	Propose $T = 0.042$ Baboon	Propose $T = 0.042$ Lena
Non-attack				
Gaussian noise 0.001				
JPEG compression				
Gaussian low pass filter [3 3]				
Salt and pepper noise 0.005				
Median filter [3 3]				
Sharpening				
Cropping centre 12.5%				

V. CONCLUSION

An image watermarking scheme using 4×4 DCT-SVD using modified entropy is presented. It uses a modified entropy that can help to identity un-noticeable region blocks for watermark embedding. The proposed watermarking scheme has modified the relation between $U_{2,1}$ and $U_{3,1}$ of U matrix on 4×4 DCT-SVD. It is demonstrated by experimental results that our scheme provides higher imperceptibility in the watermarked images with different watermark sizes. The watermarking scheme produces high potential resistance towards different types of attacks. This watermark scheme provides high capacity of embedded watermarks. The watermarking scheme produces less errors reconstruction of watermarked image with high structural similarity index. A watermarking scheme using 4×4 DCT-SVD using modified entropy has high-energy compaction, low-computational cost and easy to be implemented in the real watermarking applications.

ACKNOWLEDGMENT

The authors would like to express special thanks to Universiti Malaysia Pahang, Malaysia for providing financial support of this research project by UMP Research Grant Scheme (RDU160364).

REFERENCES

[1] Y. AL-Nabhani, H. A. Jalab, A. Wahid and R. M. Noor, "Robust watermarking algorithm for digital images using discrete wavelet and probabilistic neural network," *Journal of King Saud University-Computer and Information Sciences*, vol. 27, no. 4, pp. 393-401, Oct. 2015.

[2] Z. Wenyin and F. Y. Shih, "Semi-fragile spatial watermarking based on local binary pattern operators," *Optics Communications*, vol. 284, no. 16-17, pp. 3904-3912, Aug. 2011.

[3] A. Mishra, C. Agarwal, A. Sharma and P. Bedi, "Optimized gray-scale image watermarking using DWT–SVD and firefly algorithm," *Expert Systems with Applications*, vol. 41, no. 17, pp. 7858–7867, Dec. 2014.

[4] H. Tao, L. Chongmin, J. M. Zain and A. N. Abdalla, "Robust Image Watermarking Theories and Techniques: A Review," *Journal of Applied Research and Technology*, vol. 12, no. 1, pp. 122–138, Feb. 2014.

[5] G. Bhatnagar, "A new facet in robust digital watermarking framework," AEU-*International Journal of Electronics and Communications*, vol. 66, no. 4, pp. 275-285, Apr. 2012.

[6] G. Bhatnagar and Q. M. J. Wu, "Biometrics inspired watermarking based on a fractional dual tree complex wavelet transform," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 182-195, Jan. 2013.

[7] C. Song, S. Sudirman and M. Merabti, "A robust region-adaptive dual image watermarking technique," *Journal of Visual Communication and Image Representation*, vol. 23, no. 3, pp. 549-568, Apr. 2012.

[8] N. A. Abu, F. Ernawan, N. Suryana and S. Sahib, "Image watermarking using psychovisual threshold over the edge," 2013 Information and Communication Technology, vol. 7804, pp. 519-527.

[9] G. Bhatnagar, Q. M. J. Wu and B. Raman, "A new robust adjustable logo watermarking scheme," *Computer & Security*, vol. 31, no. 1, pp. 040-058, Feb. 2012.

[10] G. Bhatnagar, Q. M. J. Wu and B. Raman, "A new logo watermarking based on redundant fractional wavelet transform," *Mathematical and Computer Modelling*, vol. 58, no. 1-2, pp. 204-218, July 2013.

[11] S. M. Elshoura and D. B. Megherbi, "Analysis of noise sensitivity of Tchebichef and Zernike moments with application to image watermarking," *Journal of Visual Communication and Image Representation*, vol. 24, no. 5, pp. 567-578, July 2013.

[12] S. M. Elshoura and D. B. Megherbi, "A secure high capacity full-gray-scale-level multi-image information hiding and secret image authentication scheme via Tchebichef moments," *Signal Processing: Image Communication*, vol. 28, no. 5, pp. 531-552, May 2014.

[13] O. S. Faragallah, "Efficient video watermarking based on singular value decomposition in the discrete wavelet transform domain," *AEU-International Journal of Electronics and Communications*, vol. 67, no. 3, pp. 189-196, Mar. 2013.

[14] N. A. Abu and F. Ernawan, "A novel psychovisual threshold on large DCT for image compression," *The Scientific World Journal*, vol. 2015, pp. 001-011, Mar. 2015.

[15] F. Ernawan, N. A. Abu and N. Suryana, "An adaptive JPEG image compression using psychovisual model," *Advanced Science Letters*, vol. 20, no. 1, pp. 026-031, Jan. 2014.

[16] N. A. Abu, F. Ernawan and N. Suryana, "A Generic Psychovisual Error Threshold for the quantization table generation on JPEG Image Compression," 2013 9th International Colloquium on Signal Processing and its Applications, pp. 039-043.

[17] F. Ernawan and S. H. Nugraini, "The optimal quantization matrices for JPEG image compression from psychovisual threshold," *Journal of Theoretical and Applied Information Technology*, vol. 70, no. 3, pp. 566-572, Dec. 2014.

[18] N. M. Makbol, B. E. Khoo and T. H. Rassem, "Block-based discrete wavelet transform singular value decomposition image watermarking image watermarking scheme using human visual system characteristics," IET Image Processing, vol. 10, no. 1, pp. 034-052, Jan. 2016.

[19] Q. Su, Y. Niu, H. Zou and X. Liu, "A blind dual color images watermarking based on singular value decomposition," *Applied Mathematics and Computation*, vol. 219, no. 16, pp. 8455-8466, Apr. 2013.

[20] Q. Su, Y. Niu, Y. Zhao, S. Pang and X. Liu, "A dual color images watermarking scheme based on the optimized compensation of singular value decomposition," *AEU-International Journal of Electronics and Communications*, vol. 67, no. 8, pp. 652-664, Aug. 2013.

[21] C. C. Lai, "An improved SVD-based watermarking scheme using human visual characteristics," *Optics Communications*, vol. 284, no. 4, pp. 938-944, Feb. 2011.

[22] University of Granada. Computer Vision Group. CVG-UGR Image Database. [2012-10-22]. http://decsai.ugr.es/cvg/dbimagenes/c512.php

[23] C. Yim and A. C. Bovik, "Quality assessment of deblocked images," *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 088-098, Jan. 2011.

CHAPTER 6
CONCLUSION

6.1 INTRODUCTION

This chapter discusses the conclusion of this research project. Research constraint will be explained in Chapter 6.2 and the future work will be discussed in Chapter 6.3. Chapter 6.4 will conclude this research project.

6.2 RESEARCH CONSTRAINT

This section discusses about the research constraint that occur in this research. One of the research constraint is the limited dataset of the tested images. They only consist of 40 real and 40 graphical images that have been chose to test of the proposed digital watermarking scheme. The experimental result is not sufficient to prove the best performance of the proposed watermarking scheme. In future, the tested images will be increased 200 images to show more accuracy of the experimental results.

Besides, the experimental result only tested on the same size of the images. All the tested images have the same size of 512×512 pixels with the bit depth of 24-bit RGB. Our scheme uses the watermark size of 32×32 pixels in the experiment. We have limited size of the watermark in order to ensure the high quality of the watermarked image.

6.3 FUTURE WORK

The proposed psychovisual threshold has proven that it can improve the imperceptibility and robustness of the watermark image compare to other existing schmes. In the future, it can be applied in the video watermarking. The hybrid psychovisual threshold and motion analysis in the video watermarking has potential significant contribution in terms of the imperceptibility and robustness.

Besides that, the psychovisual threshold can be used to assign bit allocation in image compression. The proposed scheme has verified that it can improve the quality of image reconstruction at minimum bit rate compare to the existing current image compression. The proposed scheme has been tested in 40 real and 40 graphical images. In addition, our scheme does not provide blocking effect to the compressed image. Furthermore, the bit allocation strategy based on the psychovisual threshold can be implemented into the real device, e.g. digital camera.

6.4 CONCLUSION

This research has achieved the first objective which is to review the existing digital watermarking techniques for copyright protection. The current digital watermarking techniques have been used for comparison results. The second objective also has been achieved in this research. The second objective is to design and experiment on the psychovisual threshold in the block-based digital watermarking and transformed domain. Chapters 2, 4, and 5 have discussed the experimental results. Chapter 3 has shown the effectiveness of bit allocation strategy based on the psychovisual threshold in image compression. Last objective is to evaluate the performance of the psychovisual threshold in the digital watermarking and image compression. The proposed scheme has been compared to the current existing benchmark techniques. In digital watermarking, our scheme produces higher robustness and imperceptibility than other techniques in terms of NC and SSIM values. Our scheme also shows that it can improve the quality of the compressed image at minimum bit rate.

REFERENCES

Barni, M., Bartolini, F., Piva, A., Multichannel Watermarking of Color Images. IEEE Transactions on Circuits and Systems for Video Technology, 12(3), pp. 142-156, 2012. (Journal)

Breitmeyer, B.G. Visual Masking: Past Accomplishments, Present Status, Future Developments, Journal of Advances in Cognitive Psychology, Versita, Warsaw, 3(1), pp. 009-020, 2007. (Journal)

Chou, C.-H. and Li, Y.-C. A Perceptually Tuned Subband Image Coder Based on The Measure of Just-Noticeable-Distortion Profile, IEEE Transactions on Circuits and Systems for Video Technology, 5(6), pp. 467 – 476, 1995. (Journal)

Cui, L. and Li, W., Adaptive Multiwavelet-Based Watermarking Through JPW Masking, IEEE Transactions on Image Processing, 20(4), pp. 1047-1060, 2011. (Journal)

Halder, R., Pal, S., Cortesi, A., Watermarking Techniques for Relational Databases: Survey, Classification and Comparison, Journal of Universal Computer Science, 16(21), pp. 3164-3190, 2010. (Journal)

Karunasekera, S.A. and Kingsbury, N.G. A Distortion Measure for Blocking Artifacts in Images based on Human Visual Sensitivity, IEEE Transactions on Image Processing, 4(6), pp. 713 – 724, 1995. (Journal)

Kutter, M. and Wingkler, S., A Vision-Based Masking Model for Spread Spectrum Image Watermarking. IEEE Transactions on Image Processing, 11(1), pp. 016-025, 2002. (Journal)

Lu, C.S., Huang, S.K., Sze, C.J., Liao, H.Y.M., Cocktail Watermarking for Digital Image Protection. IEEE Transactions on Multimedia, 2(4), pp. 209-224, 2000. (Journal)

Lu, C.S. and Liao, H.Y.M., Multipurpose Watermarking for Image Authentication and Protection. IEEE Transactions on Image Processing, 10(10), pp. 1579-1592, 2001. (Journal)

Nadenau, M.J., Reichel, J. and Kunt, M. Performance Comparison of Masking Models Based on a New Psychovisual Test Method with Natural Scenery Stimuli, Signal Processing: Image Communication, 17(10), November 2002, pp. 807–823, 2002. (Journal)

Niu, Y., Kyan, M., Krishnan, S., Zhang, Q., A Combined Just Noticeable Distortion Model-Guided Image Watermarking. Signal, Image and Video Processing, 5(4), pp. 517-526, 2011. (Journal)

Painter T. and Spanias, A. Perceptual Coding of Digital Audio, Proceedings of the IEEE, 88(4), April 2000, pp. 451 – 515, 2000. (Conference Proceedings)

Pappas, T.N., Zujovic, J. and Neuhoff, D.L.,. Image Analysis and Compression: Renewed Focus on Texture. *Proc. SPIE 7543, Visual Information Processing and Communication*, 75430N, pp. 001-012, 2010. (Conference Proceedings)

Parthasarathy A.K. and Kak, S. An Improved Method of Content Based Image Watermarking. IEEE Transactions on Broadcasting, 53(2), pp. 468-479, 2007. (Journal)

Tsui, T.K., Zhang, X.P., Androutsos, D., Color image watermarking using Multidimensional Fourier Transforms. IEEE Transactions on Information Forensics and Security, 3(1), pp. 016-028, 2008. (Journal)

Wei, Z. and Ngan, K.N., Spatio-Temporal Just Noticeable Distortion Profile for Grey Scale Image/Video in DCT Domain. IEEE Transactions on Circuits and Systems for Video Technology, 19(3), pp. 337-346, 2009. (Journal)

Yang, Y., Sun, X., Yang, H., Lie, C.T., Xiao, R., A Contrast-Sensitive Reversible Visible Image Watermarking Technique. IEEE Transactions on Circuit and Systems for Video Technology, 19(5), pp. 656-667, 2009. (Journal)

Zhai, G., Wu, X., Yang, X., Lin W. and Zhang W. A Psychovisual Quality Metric in Free-Energy Principle, IEEE Transactions on Image Processing, January 2012, 21(1), pp. 041–052, 2012. (Journal)

APPENDICES

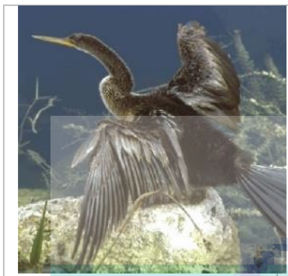
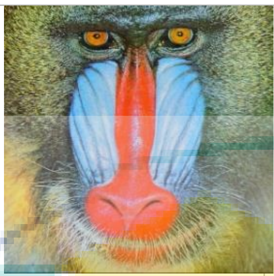
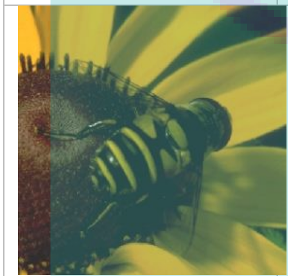
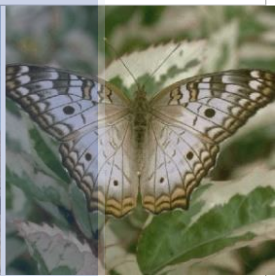
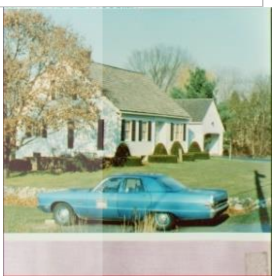
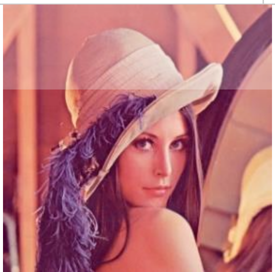
A. DIGITAL IMAGES

This research conducted the quantitative experiments on 80 images. Hence, there are 80 images chose to be tested, these images can be categories into 40 real image and 40 graphical image. These images save in bitmap format with the pixels of 512×512 and 24-bit RGB.

Real images are images that captured by camera around the environment. These images come from human being, animals, plants, and buildings. The pixels' information for these kinds of images are more complex because they include the brightness and saturation of the images. In other hand, graphical images are images that generated by the computer graphics software like Maya, Photoshop, Illustrator, 3ds Max and others. The pixels' information is much simple compared to the real images.

Table below displays the images that has been chose. Table A.1 is the sample of real images and Table A.2 is the sample of graphical images. Those images are showed in thumbnail format.

Table A.1 List of 40 real images

			
R1.bmp	R2.bmp	R3.bmp	R4.bmp
			
R5.bmp	R6.bmp	R7.bmp	R8.bmp
			
R9.bmp	R10.bmp	R11.bmp	R12.bmp
			
R13.bmp	R14.bmp	R15.bmp	R16.bmp
			
R17.bmp	R18.bmp	R19.bmp	R20.bmp

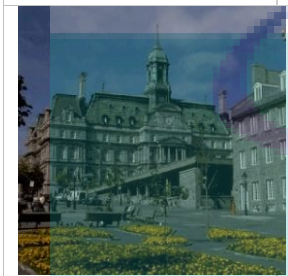
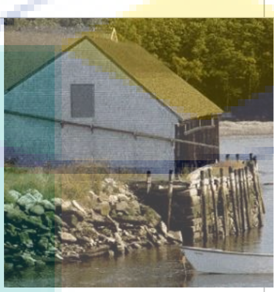
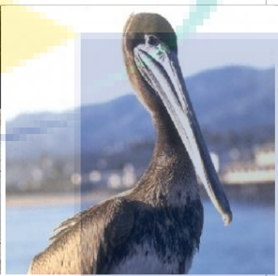
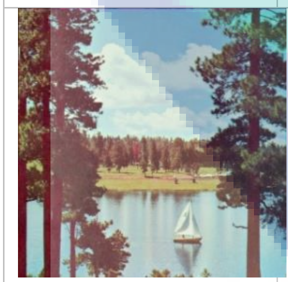
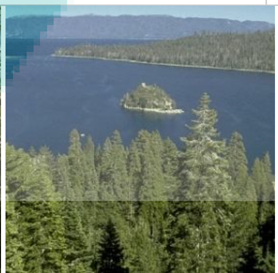
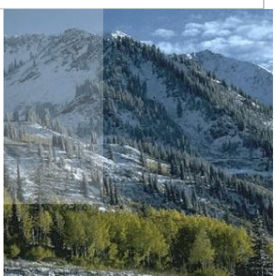
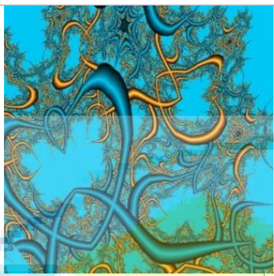
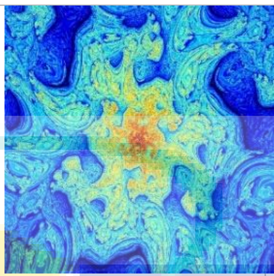
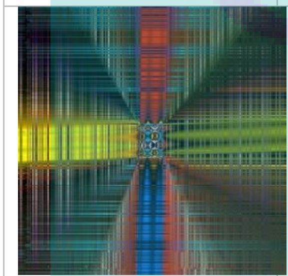
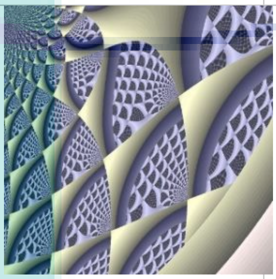
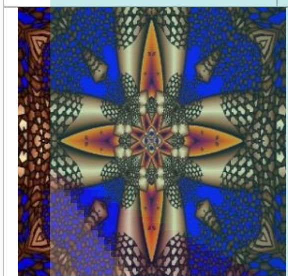
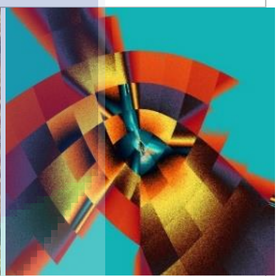
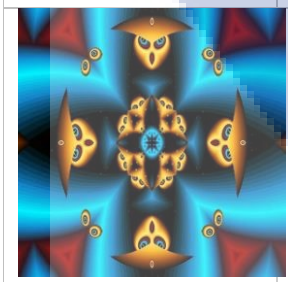
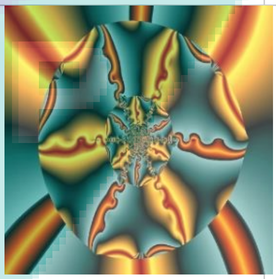
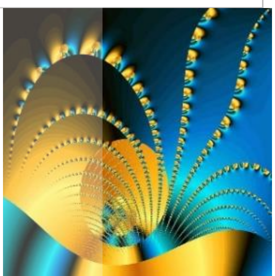
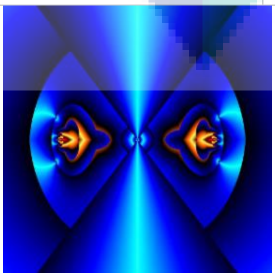
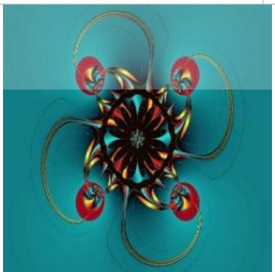
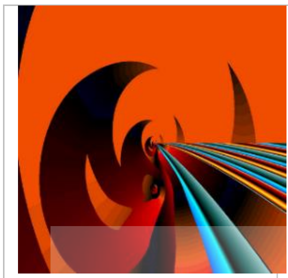
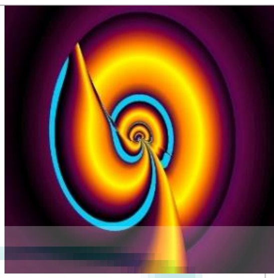
			
R21.bmp	R22.bmp	R23.bmp	R24.bmp
			
R25.bmp	R26.bmp	R27.bmp	R28.bmp
			
R29.bmp	R30.bmp	R31.bmp	R32.bmp
			
R33.bmp	R34.bmp	R35.bmp	R36.bmp
			
R37.bmp	R38.bmp	R39.bmp	R40.bmp

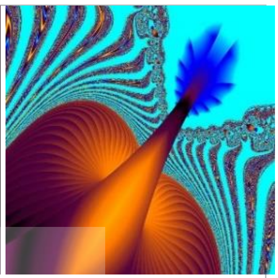
Table A.2 List of 40 graphical images

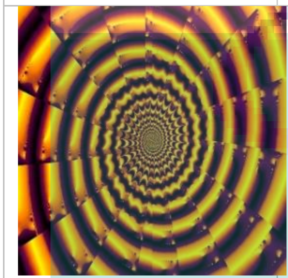
			
G1.bmp	G2.bmp	G3.bmp	G4.bmp
			
G5.bmp	G6.bmp	G7.bmp	G8.bmp
			
G9.bmp	G10.bmp	G11.bmp	G12.bmp
			
G13.bmp	G14.bmp	G15.bmp	G16.bmp
			
G17.bmp	G18.bmp	G19.bmp	G20.bmp


G21.bmp

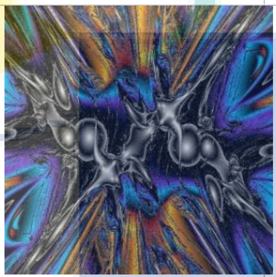

G22.bmp


G23.bmp


G24.bmp


G25.bmp

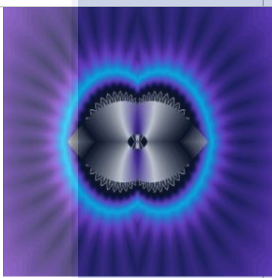

G26.bmp

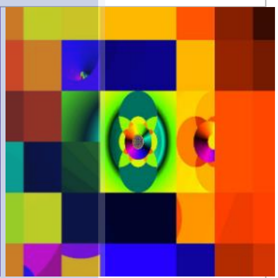

G27.bmp

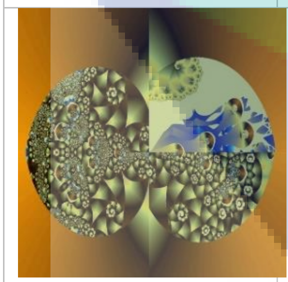

G28.bmp

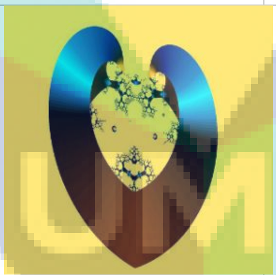

G29.bmp


G30.bmp

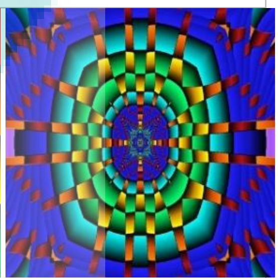

G31.bmp


G32.bmp

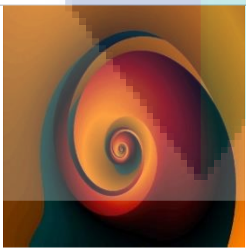

G33.bmp


G34.bmp


G35.bmp


G36.bmp


G37.bmp


G38.bmp


G39.bmp


G40.bmp