SELECTION OF JPEG STEGANOGRAPHY ALGORITHMS USING A FEATURE BASED MODEL

Vijayanathan Senthooran

(128005D)

Degree of Master of Philosophy

Department of Information Technology

University of Moratuwa
Sri Lanka

November 2018

SELECTION OF JPEG STEGANOGRAPHY ALGORITHMS USING A FEATURE BASED MODEL

Vijayanathan Senthooran

(128005D)

Thesis is submitted in partial fulfillment of the requirements for the degree Master of Philosophy

Department of Information Technology

University of Moratuwa
Sri Lanka

November 2018

DECLARATION

I declare that this is my own work and this thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

The above candidate has carried out research for the MPhil thesis under my supervision.

Name of the Supervisor: Dr. Lochandaka Ranathunga

Signature of the Supervisor: Date:

ACKNOWLEDGEMENT

Firstly, I am indebted to my supervisor Dr. Lochandaka Ranathunga for his kind cooperation, guidance, and supervision extended throughout this research project. Without his insight, this work would not be possible. I would like to express my heartfelt thanks to the Vice Chancellor of the University of Moratuwa, the Dean of the Faculty of Information Technology and the Head of the Department of Information Technology, University of Moratuwa for granting me the opportunity to commence my research work at the University of Moratuwa and facilitating me to carry out same successfully. Further, I would be failing in my duty if I do not extend my sincere gratefulness to the Rector of the Vavuniya Campus of the University of Jaffna, the Dean of the Faculty of Applied Science and the Head of the Department of Physical Science, Vavuniya Campus of the University of Jaffna for granting me two years paid study leave and releasing me from all academic duties in the Campus, thus allowing me to concentrate fully on my research work during the said two years. Thanks are also due to the Director and staff of the National Science Foundation, Sri Lanka for granting a research scholarship under the grant number NSF/SCH/2013/04 to financially support this effort.

I would like to thank all the staff members of the University of Moratuwa for their cooperation and commitment extended in various ways in order to make this project a success. The support and motivation provided by my post graduate friends, Mr. A.M.R.R. Bandara, Mr. Saneera Hemantha Kulathilake and Mrs. N.M. Wagarachchi at the Faculty of Information Technology, University of Moratuwa too deserve mentioning with a debt of gratitude.

I would like to express my appreciation to Mr. T. Kartheeswaran for his immense support and contribution to make this thesis a success. Last but not least, I offer a debt of gratitude to my parents, my wife, my son and my brothers for all their encouragement and support extended right through this endeavor. Finally, I am grateful to all those who assisted me in numerous ways during the course of this research.

Abstract

JPEG image steganographic techniques use the DCT coefficients scaled by quantization table to make secure data hiding without degrading the image quality. The selection process of data embedding locations in lower frequency DCT coefficients should be carefully considered in each image blocks as these lower frequency coefficients are high sensitive to human eyes. Some of the existing related JPEG steganographic methods have been proposed with primary quantization table modification to hide message bits in the quantized DCT coefficients with minimal distortion by analyzing the properties of quantization table entry and relevant DCT coefficients. The performance of the JPEG steganographic methods is evaluated by the imperceptibility and embedding capacity. In the literature of quantization table modification based JPEG steganography, the middle frequency coefficients in each image block are utilized to embed maximum message size by modifying the middle part of the relevant quantization table values with minimizing the effect of visual perception. However, the data hiding techniques in lower frequency coefficients from the existing studies endure from imperceptibility while increasing the message size. This study suggests the lower frequency data hiding algorithms with utilizing middle frequency data hiding in terms of the modification of lower and middle part of the quantization table values by evaluating image quality parameters and it doesn't affect the perceptual detectability and improves embedding capacity. The proposed JPEG steganography investigates the modification of quantization table values with regarding to selected lower frequency DCT coefficients for data hiding and selects different data hiding patterns in lower frequency area in terms of modification of quantization table. Finally, it returns the pair of relevant modified quantization table and generated data hiding pattern for an image based on the empirical results of the PSNR values. The pair that contains modified quantization table and data hiding pattern shared by the sender is used as a secrete key to extract the message at the receiver side. From the preliminary studies, the selection of appropriate lower frequency coefficients in image block to hide the optimum size of secrete message with perceptual un-detectability is dependent on the combination of image features, message size and the hiding algorithm. Further, this study recommends a dynamic model to keep the consistency of the combination of image features, message size and the hiding algorithm in terms of quantization table modification and this model based steganography suggests a dynamic model to cover image statistics. Eventually, the model prevents visually perceptible changes for maximum embedding message bits. The proposed method achieves a good imperceptibility level and it is evaluated by the PSNR value range 30dB to 45dB and maximum message size more than 52 bits per block for the selected JPEG image dataset. The dynamic model fitted between the quantization tables and cover image statistics shows the statistical significance with the p-value 0.0007634 and the model generated between the data hiding pattern and statistical features of DCT coefficients shows the statistical significance with the p-value 4.598e-13. The dynamic model for the selected data hiding patterns in the lower frequency coefficients hides the message and it is stego invariant for message analyzers.

Key words: JPEG Steganography, imperceptibility, embedding capacity, quantization table, DCT transformation.

TABLE OF CONTENTS

Declaration i

Acknowledgement ii

Abstract iii

Table of Contents iv

List of Figures viii

List of Tables xi

List of Appendices xii

Chapter 1 Introduction 1

1.1 Information Security and Steganography 2

1.2 Stenography and the ability of secrete communication 5

1.3 Types of steganography based on its mechanism 6

1.3.1 Key based 6

1.3.2 Countermeasures against steganogrphy 7

1.3.3. Embedding methodology based categorization 8

1.3.4 Carrier based categorization 9

1.4. Steganography in digital images 11

1.4.1 Data hiding in digital images 11

1.4.2 Image based steganography 12

1.4.3 Image Steganographic notions 13

1.4.4 Fundamental properties of image steganographic systems 14

1.4.5 Image steganography methods 17

1.4.6 Evaluation of image steganography 20

1.5. JPEG steganography 23

1.5.1 DCT based steganography 23

1.5.2 DCT and JPEG compression 24

1.5.3 Steganography in JPEG domain 27

1.5.4 Influence of quantization process in JPEG steganography 32

1.5.5 Optimization of quantization tables in JPEG steganography 33

1.6 Model based steganography ... 36
1.7 Solution provided by the proposed approach ... 37
1.8. Motivation and research problem ... 38
1.9 Aim of the study and Objectives ... 39
1.10. Thesis outline ... 40
1.11 Chapter summary ... 40
Chapter 2 Literature Review ... 42
2.1 Review of JPEG steganography ... 42
2.2 Quantization table based JPEG steganographic techniques ... 44
2.3 Summary ... 53
Chapter 3 Jpeg Steganography Based on Quantization Table Modification ... 54
3.1. Role of quantization table in JPEG steganography ... 54
3.2. Mathematical description of JPEG steganography ... 55
3.3. Steganography in JPEG compressed images ... 59
3.4. Overview of the proposed method ... 60
3.4.1. A practical JPEG steganography method based on quantization table modification ... 61
3.4.2. Generation of fifteen patterns using PSNR ... 64
3.4.3. Proposed practical method to select the quantization table and relevant hiding pattern ... 65
3.4.4. Extracting procedure ... 68
3.4.5 Find the relationship between the cover image features, selected quantization table and hiding patterns ... 68
3.4.6. Correlation between image features, quantization table and hiding pattern ... 71
3.4.7. Identification of quantization table and data hiding patterns using Model ... 71
3.4.8. Dynamic model based JPEG steganography ... 73
3.5. Chapter Summary ... 74
Chapter 4 Experimental Setup ... 76
4.1. Image set ... 76
4.2. Quantization table modification ... 77
4.3. Data hiding pattern generation ... 79

4.4. Cross experiments combined quantization table and data hiding patterns 87

4.4.1. Experimental design of proposed approach 88

4.5. Relationship between quantization table, data hiding pattern and image features 88

4.5.1. Feature Extraction 88

4.5.2. Find the correlation between the quantization table, data hiding pattern and image features using R library. 89

4.5.3 Identification of data hiding patterns and quantization tables with respect to image features by generating dynamic model 89

4.6. Performance Evaluation Parameters 90

4.6.1. The imperceptibility test 90

4.6.2. The Image quality test 90

4.6.3. The security of image steganography 91

4.7 Chapter Summary 91

Chapter 5 Results and Discussion 93

5.1. The PSNR of standard pattern 93

5.2. Results and Analysis of data hiding patterns generation 95

5.2.1. PSNR analysis for Pattern 1 95

5.3. The results and justification for proposed practical JPEG steganography method 99

5.3.1 The performance of proposed method 100

5.4. Image quality Test for proposed method 102

5.4.1. Analysis of Normalized Cross Correlation 103

5.4.2. Analysis of Structural Content 104

5.4.3. Analysis of Relative entropy (RE) 105

5.5. Embedding capacity of proposed randomly generated hiding patterns in lower frequencies 106

5.6. The results of file size 107

5.7. Relationship between the quantization table, data hiding patterns and image features – Identification Model 108

5.7.1. Identification Model 108

5.8. Model validation 121

5.9 Comparison of the proposed method with recent literature……………………121

5.10 Chapter Summary ………………………………………………………122

Chapter 6 Conclusion and Future work 1233

References 125

Appendix A: Comparison of psnr values for standard hiding pattern and fifteen generated hiding patterns 1366

Appendix B: Results of selected quantization tables and relevant data patterns for image set... 1433

Appendix C: Results of the comparison of visual perception of sample cover and stego images with respect to quantization table and selected data hiding pattern…144

Appendix D: R analysis for different combination of feature vectors 1499

Appendix E: Sample unknown images for model validation…………………….. 1544

Appendix_F: Unstisfactory results for unknown images for the model …………...155

Appendix G: Publications based of this research study……………………………156

LIST OF FIGURES

Figure 1.1: The principal difference between the function of cryptography and steganography. 2

Figure 1.2: Illustration of steganography system. 6

Figure 1.3: General principle of image steganography system. 14

Figure 1.4: Block diagram of JPEG compression showing the lossy and lossless. 27 stages during the encoding process.

Figure 1.5: The JPEG Steganography process. 28

Figure 1.6: Frequency locations in a DCT block where embedding takes place in proposed method 30

Figure 1.7: Block diagram of the process JPEG steganography in compressed domain. 31

Figure 3.1: Standard JPEG quantization table. 56

Figure 3.2: Modification of middle elements in primary quantization table. 62

Figure 3.3: Indented modified elements in top left part of the primary quantization.62

Figure 3.4: Different coefficients specified in the proposed approach. 64

Figure 3.5: Block diagram of the proposed practical method. 67

Figure 3.6: DCT feature extraction in selected coefficients. 69

Figure 3.7: The proposed dynamic model based JPEG Steganography 73

Figure 4.1: Sample Cover Images. 76

Figure 4.2: Extracted quantization table from cover image set. 77

Figure 4.3: Modified quantization table according to Literature. 78

Figure 4.4: (f1 = 1/4). 78

Figure 4.5: (f2 = 1/2). 78

Figure 4.6: (f3 = 3/4). 78

Figure 4.7: Number of bits to be hidden in middle frequency area. 79

Figure 4.8: Pattern 1. 80

Figure 4.9: Pattern 2. 80

Figure 4.10: Pattern 3. 81

Figure 4.11: Pattern 4. 81

Figure 4.12: Pattern 5 82

Figure 4.13: Pattern 6 82

Figure 4.14: Pattern 7 83

Figure 4.15: Pattern 8 83

Figure 4.16: Pattern 9 84

Figure 4.17: Pattern 10 84

Figure 4.18: Pattern 11 85

Figure 4.19: Pattern 12. 85

Figure 4.20: Pattern 13. 86

Figure 4.21: Pattern 14. 86

Figure 4.22: Pattern 15. 87

Figure 4.23: Block diagram of cross experiments with q quantization tables and hiding patterns. 88

Figure 4.24: Sample stego images with respected to selected quantization table and hiding pattern. 88

Figure 5.1: The PSNR values for standard hiding patterns with modified quantization tables. 93

Figure 5.2: The PSNR values for standard hiding patterns with four modified quantization tables. 94

Figure 5.3: Comparison of PSNR values of Standard hiding pattern and Patter 1 for images set. 96

Figure 5.4: Residual plot of PSNR values for image set 97

Figure 5.5: The comparison of PSNR values of proposed hiding method with standard hiding pattern for image set. 101

Figure 5.6: Residual plot of PSNR values for image set 102

Figure 5.7: PSNR range for quantization tables 102

Figure 5.8: PSNR range for generated data hiding patterns 103

Figure 5.9: The Comparison of NCC for standard hiding method and proposed hiding method. 104

Figure 5.10: The comparison of SC values for standard hiding method and proposed hiding method. 105

Figure 5.11: The results of relative entropy of cover and stego image set. 105

Figure 5.12: Comparison of PSNR range for standard hiding and proposed method. 107

Figure 5.13: The analysis of cover and stego image file size 108

LIST OF TABLES

Table 5.1: Number of selected secondary quantization tables for the image set. 95

Table 5.2: The results of other patterns and their PSNR range. 96

Table 5.3: Data hiding patterns with maximum message size. 97

Table 5.4: Statistical difference between the PSNR values. 101

Table 5.5: Data hiding patterns with maximum message size 106

Table 5.6: Results of the model validation for MIT Forest Images 111

Table 5.7: F-Test for QT 119

Table 5.8: F-Test for Patt 119

Table 5.9: Results of model validation for unknown image set 120

Table 5.10: Comparison of proposed method with recent literature 121

Table 5.11: Comparison of proposed method with recent literature 122

LIST OF APPENDICES

Appendix	Description	Page
Appendix-A	Comparison of PSNR values for standard Hiding pattern and fifteen generated hiding Patterns.	136
Appendix-B	Results of selected quantization tables and Relevant data patterns for image set.	143
Appendix-C	Results of the comparison of visual perception Of sample cover and stego images with respected to selected quantization table and data hiding pattern	144
Appendix-D	R analysis for different combination of feature Vectors.	149
Appendix-E	Sample unknown images for model validation	154
Appendix-F	Unsatisfactory results of unknown images for the model	155
Appendix-G	Publications based on the research study	156