Mauro Barni
Ingemar Cox
Ton Kalker
Hyoung Joong Kim (Eds.)

LNCS 3710

Digital Watermarking

4th International Workshop, IWDW 2005
Siena, Italy, September 2005
Proceedings

Springer

Mauro Barni Ingemar Cox
Ton Kalker Hyoung Joong Kim (Eds.)

Digital Watermarking

4th International Workshop, IWDW 2005
Siena, Italy, September 15-17, 2005
Proceedings

Springer

Volume Editors

Mauro Barni
University of Siena, Department of Information Engineering
Via Roma 56, 53100 Siena, Italy
E-mail: barni@dii.unisi.it

Ingemar Cox
University College London, Torrington Place, UK
E-mail: i.cox@ee.ucl.ac.uk

Ton Kalker
Hewlett-Packard Labs
1501 Page Mill Road, Palo Alto, CA 94305, USA
E-mail: Ton.Kalker@hp.com

Hyoung Joong Kim
Kangwon National University, Chunchon, 200-701, Korea
E-mail: khj@kangwon.ac.kr

Library of Congress Control Number: 2005931932

CR Subject Classification (1998): K.4.1, K.6.5, H.5.1, D.4.6, E.3, E.4, F.2.2, H.3, I.4

ISSN 0302-9743
ISBN-10 3-540-28768-X Springer Berlin Heidelberg New York
ISBN-13 978-3-540-28768-1 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springeronline.com

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11551492 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 3710

Preface

We are delighted to welcome the attendees of the Fourth International Workshop on Digital Watermarking (IWDW). Watermarking continues to generate strong academic interest. Commercialization of the technology is proceeding at a steady pace. We have seen watermarking adopted for DVD audio. Fingerprinting technology was successfully used to determine the source of pirated video material. Furthermore, a number of companies are using watermarking as an enabling technology for broadcast monitoring services. Watermarking of digital cinema content is anticipated. Future applications may also come from areas unrelated to digital rights management. For example, the use of watermarking to enhance legacy broadcast and communication systems is now being considered. IWDW 2005 offers an opportunity to reflect upon the state of the art in digital watermarking as well as discuss directions for future research and applications.

This year we accepted 31 papers from 74 submissions. This 42% acceptance rate indicates our commitment to ensuring a very high quality conference. We thank the members of the Technical Program Committee for making this possible by their timely and insightful reviews. Thanks to their hard work this is the first IWDW at which the final proceedings are available to the participants at the time of the workshop as a Springer LNCS publication.

This year's program reflects all the major interests of the watermarking community. The accepted papers cover a full range of topics, including robust and fragile watermarking, steganography and steganalysis, security and attacks, and fingerprinting and benchmarking. These papers address the theoretical and practical issues that we felt to be of broad interest to our community. Moreover, this year we will also have a very relevant special session on foundational and practical aspects of watermarking security.

Finally, this year's workshop is special since it is the first installment of IWDW to be held outside of Korea. It is our aim that future IWDW workshops will rotate between locations in Asia, Europe and the Americas. We hope you will find the workshop useful and enjoyable, and we look forward to meeting you again in the context of IWDW.

Welcome to IWDW 2005 in Siena!

July 2005

Ingemar Cox
Ton Kalker
Hyoung Joong Kim
Mauro Barni

Organization

General Chairs

Mauro Barni (University of Siena, Italy)
Daeho Kim (NSRI, Korea)

Technical Program Chairs

Ingemar J. Cox (UCL, UK)
Ton Kalker (HP, USA)
Hyoung Joong Kim (Kangwon National University, Korea)

Finance Chair

Roberto Caldelli (University of Florence, Italy)

Publicity Chair

Vito Cappellini (University of Florence, Italy)

Electronic Media Chair

Alessia De Rosa (University of Florence, Italy)

Publications Chair

Enrico Magli (Politecnico di Torino, Italy)

Technical Program Committee

Roberto Caldelli (U. of Florence, Italy)
Patrizio Campisi (U. of Roma III, Italy)
Alessia De Rosa (U. of Florence, Italy)
Jana Dittman (U. Magdeburg, Germany)
Jean-Luc Dugelay (Eurecom, France)
Touradj Ebrahimi (EPFL, Switzerland)

Jessica Fridrich (SUNY Binghamton, USA)
Teddy Furon (IRISA, France)
Miroslav Goljan (SUNY Binghamton, USA)
Yo-Sung Ho (K-JIST, Korea)
Jinwoo Hong (ETRI, Korea)
Jiwu Huang (Zhongshan U., China)
Ebroul Izquierdo (Q. Mary Univ. London, UK)
Mohan Kankanhalli (NUS, Singapore)
Stefan Katzenbeisser (Technical Univ. Munich, Germany)
Inald Lagendijk (Delft U. of Tech., The Netherlands)
Heung-Kyu Lee (KAIST, Korea)
Benoit Macq (UCL, Belgium)
Enrico Magli (Politecnico di Torino, Italy)
Nasir Memon (Polytechnic U., USA)
Kivanc Mihcak (Microsoft, USA)
Matt Miller (NEC, USA)
Pierre Moulin (U. of Illinois, USA)
Hideki Noda (Kyushu Inst. Tech., Japan)
Fernando Pérez-González (U. of Vigo, Spain)
Ioannis Pitas (U. of Thessaloniki, Greece)
Alessandro Piva (U. of Florence, Italy)
William Puech (LIRMM, France)
Hayder Radha (Michigan State U., USA)
Yong-Man Ro (ICU, Korea)
David Saad (Aston University, UK)
Kouichi Sakurai (Kyushu U., Japan)
Farook Sattar (Nanyang Tech. U., Singapore)
Yun-Qing Shi (New Jersey Inst. of Tech., USA)
Shan Suthaharan (Tennesee St. U., USA)
Kiyoshi Tanaka (Shinshu University, Japan)
Sviatoslav Voloshynovsky (U. of Geneva, Switzerland)
Chee Sun Won (Dongguk University, Korea)
Min Wu (U. of Maryland, USA)

Lecture Notes in Computer Science

For information about Vols. 1–3577

please contact your bookseller or Springer

Vol. 3710: M. Barni, I. Cox, T. Kalker, H.J. Kim (Eds.), Digital Watermarking. XII, 485 pages. 2005.

Vol. 3703: F. Fages, S. Soliman (Eds.), Principles and Practice of Semantic Web Reasoning. VIII, 163 pages. 2005.

Vol. 3698: U. Furbach (Ed.), KI 2005: Advances in Artificial Intelligence. XIII, 409 pages. 2005. (Subseries LNAI).

Vol. 3697: W. Duch, J. Kacprzyk, E. Oja, S. Zadrożny (Eds.), Artificial Neural Networks: Formal Models and Their Applications - ICANN 2005, Part II. XXXII, 1045 pages. 2005.

Vol. 3696: W. Duch, J. Kacprzyk, E. Oja, S. Zadrożny (Eds.), Artificial Neural Networks: Biological Inspirations - ICANN 2005, Part I. XXXI, 703 pages. 2005.

Vol. 3687: S. Singh, M. Singh, C. Apte, P. Perner (Eds.), Pattern Recognition and Image Analysis, Part II. XXV, 809 pages. 2005.

Vol. 3686: S. Singh, M. Singh, C. Apte, P. Perner (Eds.), Pattern Recognition and Data Mining, Part I. XXVI, 689 pages. 2005.

Vol. 3674: W. Jonker, M. Petković (Eds.), Secure Data Management. X, 241 pages. 2005.

Vol. 3672: C. Hankin, I. Siveroni (Eds.), Static Analysis. X, 369 pages. 2005.

Vol. 3671: S. Bressan, S. Ceri, E. Hunt, Z.G. Ives, Z. Bellahsène, M. Rys, R. Unland (Eds.), Database and XML Technologies. X, 239 pages. 2005.

Vol. 3670: M. Bravetti, L. Kloul, G. Zavattaro (Eds.), Formal Techniques for Computer Systems and Business Processes. XIII, 349 pages. 2005.

Vol. 3665: K. S. Candan, A. Celentano (Eds.), Advances in Multimedia Information Systems. X, 221 pages. 2005.

Vol. 3664: C. Türker, M. Agosti, H.-J. Schek (Eds.), Peer-to-Peer, Grid, and Service-Orientation in Digital Library Architectures. X, 261 pages. 2005.

Vol. 3663: W.G. Kropatsch, R. Sablatnig, A. Hanbury (Eds.), Pattern Recognition. XIV, 512 pages. 2005.

Vol. 3662: C. Baral, G. Greco, N. Leone, G. Terracina (Eds.), Logic Programming and Nonmonotonic Reasoning. XIII, 454 pages. 2005. (Subseries LNAI).

Vol. 3660: M. Beigl, S. Intille, J. Rekimoto, H. Tokuda (Eds.), UbiComp 2005: Ubiquitous Computing. XVII, 394 pages. 2005.

Vol. 3659: J.R. Rao, B. Sunar (Eds.), Cryptographic Hardware and Embedded Systems – CHES 2005. XIV, 458 pages. 2005.

Vol. 3658: V. Matoušek, P. Mautner, T. Pavelka (Eds.), Text, Speech and Dialogue. XV, 460 pages. 2005. (Subseries LNAI).

Vol. 3655: A. Aldini, R. Gorrieri, F. Martinelli (Eds.), Foundations of Security Analysis and Design III. VII, 273 pages. 2005.

Vol. 3654: S. Jajodia, D. Wijesekera (Eds.), Data and Applications Security XIX. X, 353 pages. 2005.

Vol. 3653: M. Abadi, L.d. Alfaro (Eds.), CONCUR 2005 – Concurrency Theory. XIV, 578 pages. 2005.

Vol. 3649: W.M.P. van der Aalst, B. Benatallah, F. Casati, F. Curbera (Eds.), Business Process Management. XII, 472 pages. 2005.

Vol. 3648: J.C. Cunha, P.D. Medeiros (Eds.), Euro-Par 2005 Parallel Processing. XXXVI, 1299 pages. 2005.

Vol. 3646: A. F. Famili, J.N. Kok, J.M. Peña, A. Siebes, A. Feelders (Eds.), Advances in Intelligent Data Analysis VI. XIV, 522 pages. 2005.

Vol. 3645: D.-S. Huang, X.-P. Zhang, G.-B. Huang (Eds.), Advances in Intelligent Computing, Part II. XIII, 1010 pages. 2005.

Vol. 3644: D.-S. Huang, X.-P. Zhang, G.-B. Huang (Eds.), Advances in Intelligent Computing, Part I. XXVII, 1101 pages. 2005.

Vol. 3642: D. Ślezak, J. Yao, J.F. Peters, W. Ziarko, X. Hu (Eds.), Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Part II. XXIII, 738 pages. 2005. (Subseries LNAI).

Vol. 3641: D. Ślezak, G. Wang, M. Szczuka, I. Düntsch, Y. Yao (Eds.), Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Part I. XXIV, 742 pages. 2005. (Subseries LNAI).

Vol. 3639: P. Godefroid (Ed.), Model Checking Software. XI, 289 pages. 2005.

Vol. 3638: A. Butz, B. Fisher, A. Krüger, P. Olivier (Eds.), Smart Graphics. XI, 269 pages. 2005.

Vol. 3637: J. M. Moreno, J. Madrenas, J. Cosp (Eds.), Evolvable Systems: From Biology to Hardware. XI, 227 pages. 2005.

Vol. 3636: M.J. Blesa, C. Blum, A. Roli, M. Sampels (Eds.), Hybrid Metaheuristics. XII, 155 pages. 2005.

Vol. 3634: L. Ong (Ed.), Computer Science Logic. XI, 567 pages. 2005.

Vol. 3633: C. Bauzer Medeiros, M. Egenhofer, E. Bertino (Eds.), Advances in Spatial and Temporal Databases. XIII, 433 pages. 2005.

Vol. 3632: R. Nieuwenhuis (Ed.), Automated Deduction – CADE-20. XIII, 459 pages. 2005. (Subseries LNAI).

Vol. 3630: M.S. Capcarrere, A.A. Freitas, P.J. Bentley, C.G. Johnson, J. Timmis (Eds.), Advances in Artificial Life. XIX, 949 pages. 2005. (Subseries LNAI).

Vol. 3629: J.L. Fiadeiro, N. Harman, M. Roggenbach, J. Rutten (Eds.), Algebra and Coalgebra in Computer Science. XI, 457 pages. 2005.

Vol. 3628: T. Gschwind, U. Aßmann, O. Nierstrasz (Eds.), Software Composition. X, 199 pages. 2005.

Vol. 3627: C. Jacob, M.L. Pilat, P.J. Bentley, J. Timmis (Eds.), Artificial Immune Systems. XII, 500 pages. 2005.

Vol. 3626: B. Ganter, G. Stumme, R. Wille (Eds.), Formal Concept Analysis. X, 349 pages. 2005. (Subseries LNAI).

Vol. 3625: S. Kramer, B. Pfahringer (Eds.), Inductive Logic Programming. XIII, 427 pages. 2005. (Subseries LNAI).

Vol. 3624: C. Chekuri, K. Jansen, J.D.P. Rolim, L. Trevisan (Eds.), Approximation, Randomization and Combinatorial Optimization. XI, 495 pages. 2005.

Vol. 3623: M. Liśkiewicz, R. Reischuk (Eds.), Fundamentals of Computation Theory. XV, 576 pages. 2005.

Vol. 3621: V. Shoup (Ed.), Advances in Cryptology – CRYPTO 2005. XI, 568 pages. 2005.

Vol. 3620: H. Muñoz-Avila, F. Ricci (Eds.), Case-Based Reasoning Research and Development. XV, 654 pages. 2005. (Subseries LNAI).

Vol. 3619: X. Lu, W. Zhao (Eds.), Networking and Mobile Computing. XXIV, 1299 pages. 2005.

Vol. 3618: J. Jedrzejowicz, A. Szepietowski (Eds.), Mathematical Foundations of Computer Science 2005. XVI, 814 pages. 2005.

Vol. 3617: F. Roli, S. Vitulano (Eds.), Image Analysis and Processing – ICIAP 2005. XXIV, 1219 pages. 2005.

Vol. 3615: B. Ludäscher, L. Raschid (Eds.), Data Integration in the Life Sciences. XII, 344 pages. 2005. (Subseries LNBI).

Vol. 3614: L. Wang, Y. Jin (Eds.), Fuzzy Systems and Knowledge Discovery, Part II. XLI, 1314 pages. 2005. (Subseries LNAI).

Vol. 3613: L. Wang, Y. Jin (Eds.), Fuzzy Systems and Knowledge Discovery, Part I. XLI, 1334 pages. 2005. (Subseries LNAI).

Vol. 3612: L. Wang, K. Chen, Y. S. Ong (Eds.), Advances in Natural Computation, Part III. LXI, 1326 pages. 2005.

Vol. 3611: L. Wang, K. Chen, Y. S. Ong (Eds.), Advances in Natural Computation, Part II. LXI, 1292 pages. 2005.

Vol. 3610: L. Wang, K. Chen, Y. S. Ong (Eds.), Advances in Natural Computation, Part I. LXI, 1302 pages. 2005.

Vol. 3608: F. Dehne, A. López-Ortiz, J.-R. Sack (Eds.), Algorithms and Data Structures. XIV, 446 pages. 2005.

Vol. 3607: J.-D. Zucker, L. Saitta (Eds.), Abstraction, Reformulation and Approximation. XII, 376 pages. 2005. (Subseries LNAI).

Vol. 3606: V. Malyshkin (Ed.), Parallel Computing Technologies. XII, 470 pages. 2005.

Vol. 3604: R. Martin, H. Bez, M. Sabin (Eds.), Mathematics of Surfaces XI. IX, 473 pages. 2005.

Vol. 3603: J. Hurd, T. Melham (Eds.), Theorem Proving in Higher Order Logics. IX, 409 pages. 2005.

Vol. 3602: R. Eigenmann, Z. Li, S.P. Midkiff (Eds.), Languages and Compilers for High Performance Computing. IX, 486 pages. 2005.

Vol. 3599: U. Aßmann, M. Aksit, A. Rensink (Eds.), Model Driven Architecture. X, 235 pages. 2005.

Vol. 3598: H. Murakami, H. Nakashima, H. Tokuda, M. Yasumura, Ubiquitous Computing Systems. XIII, 275 pages. 2005.

Vol. 3597: S. Shimojo, S. Ichii, T.W. Ling, K.-H. Song (Eds.), Web and Communication Technologies and Internet-Related Social Issues - HSI 2005. XIX, 368 pages. 2005.

Vol. 3596: F. Dau, M.-L. Mugnier, G. Stumme (Eds.), Conceptual Structures: Common Semantics for Sharing Knowledge. XI, 467 pages. 2005. (Subseries LNAI).

Vol. 3595: L. Wang (Ed.), Computing and Combinatorics. XVI, 995 pages. 2005.

Vol. 3594: J.C. Setubal, S. Verjovski-Almeida (Eds.), Advances in Bioinformatics and Computational Biology. XIV, 258 pages. 2005. (Subseries LNBI).

Vol. 3593: V. Mařík, R. W. Brennan, M. Pěchouček (Eds.), Holonic and Multi-Agent Systems for Manufacturing. XI, 269 pages. 2005. (Subseries LNAI).

Vol. 3592: S. Katsikas, J. Lopez, G. Pernul (Eds.), Trust, Privacy and Security in Digital Business. XII, 332 pages. 2005.

Vol. 3591: M.A. Wimmer, R. Traunmüller, Å. Grönlund, K.V. Andersen (Eds.), Electronic Government. XIII, 317 pages. 2005.

Vol. 3590: K. Bauknecht, B. Pröll, H. Werthner (Eds.), E-Commerce and Web Technologies. XIV, 380 pages. 2005.

Vol. 3589: A M. Tjoa, J. Trujillo (Eds.), Data Warehousing and Knowledge Discovery. XVI, 538 pages. 2005.

Vol. 3588: K.V. Andersen, J. Debenham, R. Wagner (Eds.), Database and Expert Systems Applications. XX, 955 pages. 2005.

Vol. 3587: P. Perner, A. Imiya (Eds.), Machine Learning and Data Mining in Pattern Recognition. XVII, 695 pages. 2005. (Subseries LNAI).

Vol. 3586: A.P. Black (Ed.), ECOOP 2005 - Object-Oriented Programming. XVII, 631 pages. 2005.

Vol. 3584: X. Li, S. Wang, Z.Y. Dong (Eds.), Advanced Data Mining and Applications. XIX, 835 pages. 2005. (Subseries LNAI).

Vol. 3583: R.W. H. Lau, Q. Li, R. Cheung, W. Liu (Eds.), Advances in Web-Based Learning – ICWL 2005. XIV, 420 pages. 2005.

Vol. 3582: J. Fitzgerald, I.J. Hayes, A. Tarlecki (Eds.), FM 2005: Formal Methods. XIV, 558 pages. 2005.

Vol. 3581: S. Miksch, J. Hunter, E. Keravnou (Eds.), Artificial Intelligence in Medicine. XVII, 547 pages. 2005. (Subseries LNAI).

Vol. 3580: L. Caires, G.F. Italiano, L. Monteiro, C. Palamidessi, M. Yung (Eds.), Automata, Languages and Programming. XXV, 1477 pages. 2005.

Vol. 3579: D. Lowe, M. Gaedke (Eds.), Web Engineering. XXII, 633 pages. 2005.

Vol. 3578: M. Gallagher, J. Hogan, F. Maire (Eds.), Intelligent Data Engineering and Automated Learning - IDEAL 2005. XVI, 599 pages. 2005.

Table of Contents

Session I: Steganography and Steganalysis

Session II: Fingerprinting

Session III: Watermarking I

Session IV: Attacks

Session V: Special Session on Watermarking Security

Session VI: Watermarking of Unconventional Media

Session VII: Channel Coding and Watermarking

Session VIII: Theory

Session IX: Watermarking II

Session X: Applications

A New Approach to Estimating Hidden Message Length in Stochastic Modulation Steganography

Junhui He[1], Jiwu Huang[1,*], and Guoping Qiu[2]

[1] School of Information Science and Technology,
Sun Yat-sen University, Guangzhou, China, 510275
isshjw@zsu.edu.cn
[2] School of Computer Science, University of Nottingham, NG8 1BB, UK

Abstract. Stochastic modulation steganography hides secret message within the cover image by adding a weak noise signal with a specified probabilistic distribution. The advantages of stochastic modulation steganography include high capacity and better security. Current steganalysis methods that are applicable to the detection of hidden message in traditional least significant bit (LSB) or additive noise model based steganography cannot reliably detect the existence of hidden message in stochastic modulation steganography. In this paper, we present a new steganalysis approach which can reliably detect the existence and accurately estimate the length of hidden message in stochastic modulation steganography. By analyzing the distributions of the horizontal pixel difference of the images before and after stochastic modulation embedding, it is shown that for non-adaptive steganography, the distribution of the stego-image's pixel difference can be modeled as the convolution of the distribution of the cover image's pixel difference and that of the quantized stego-noise difference, and that the estimation of the hidden message length in stochastic modulation can be achieved by estimating the variance of the stego-noise. To estimate the variance of the stego-noise, hence determining the existence and the length of hidden message, we first model the distribution of the cover image's pixel difference as a generalized Gaussian and estimate the parameters of the distribution using grid search and Chi-square goodness of fit test, and then exploit the relationship between the distribution variance of the cover image's pixel difference and that of the stego-noise difference. We present experimental results to demonstrate that our new approach is effective for steganalyzing stochastic modulation steganography. Our method provides a general theoretical framework and is applicable to other non-adaptive embedding algorithms where the distribution models of the stego-noise are known or can be estimated.

1 Introduction

Steganography [1] conceals the occurrence of communication by embedding message into the cover medium such as an image, an audio recording, or a video film

[*] Correspondence author.

M. Barni et al. (Eds.): IWDW 2005, LNCS 3710, pp. 1–14, 2005.

and has received much attention in secret communication. Image is one of the most important cover media for steganography.

To be useful, a steganographic system should be able to provide a relatively high capacity of information hiding. At the same time, the embedded secret message should be undetectable. If the existence of secret message can be detected by an attacker with a probability higher than random guessing, the corresponding steganography technique is considered to be invalid. Similar to cryptography, steganography techniques may suffer from many active or passive attacks (referred as steganalysis [2]) such as detecting the existence of hidden message, searching the steganography key or estimating the secret message length.

The LSB-based steganography is one of the conventional techniques capable of hiding a long secret message in the cover image without introducing perceptible distortions. It works by replacing the LSBs of sequentially or randomly selected pixels in the cover image with the secret message bits. The ways in which pixels are selected are usually determined by a secret key. Without the knowledge of this key, it is difficult for an attacker to extract the embedded message.

Many steganography tools using LSB-based steganography techniques, including Steghide, S-Tools, Steganos, SteganoDos, Winstorms, etc., are available on the Internet[1]. In recent years, LSB-based steganography has been widely investigated and many steganalytic approaches, such as Chi-square statistical attack [3], generalized Chi-square statistical attack [4], Regular-Singular method [5], detection based on difference histogram [6] and Sample Pairs analysis [7], have been proposed. These steganalysis methods can detect hidden message with high reliability or accurately estimate the length of secret message embedded with LSB-based steganography.

However, there are some more advanced steganograhpy algorithms, examples including, Hide [8], the spread spectrum image steganography (SSIS) [9], and the stochastic modulation steganography [10], are robust against most of the steganalysis methods mentioned above. These techniques, referred to as additive noise steganography in this paper, hide secret message in the cover image by adding stego-noise with a specific probabilistic distribution and have better security.

With the advance of research in steganalysis, Hide and SSIS steganography have been successfully steganalyzed by neighbor colors histogram (NCH) analysis [11] and histogram characteristic function center of mass (HCF-COM) analysis [12], respectively. The NCH method counts the number of neighbors of each unique color in the image to reliably detect the existence of hidden message in Hide steganography. However, it is only applicable to the images that do not have a large number of unique colors. If the cover image is grayscale or high quality color image, this attack works less reliably and may have high false positives. The HCF-COM method shows that some additive noise embedding methods are equivalent to low pass filtering the cover image's histogram and builds a classifier which performs very well on SSIS. However, the method

[1] `http://www.stegoarchive.com`

needs proper choice of training images and it may be hard or impractical to find a universal threshold for a sufficiently wide class of images. According to the principles of these two steganalysis methods, it will be extremely difficult for them to accurately estimate the secret message length in stochastic modulation steganography. Although a steganalytic technique based on the analysis of translation coefficients between the pixel difference histograms and capable of estimating the secret message length in LSB steganography has been proposed in [6], it may not be directly applicable to the steganalysis of additive noise steganography.

In this paper, we propose a new steganalysis method for reliably detecting the existence and for accurately estimating the length of secret message embedded with stochastic modulation steganography. We model the distribution of the cover image's pixel difference as generalized Gaussian, and model the distribution of the stego-image's pixel difference as the convolution of the distribution of the cover image's pixel difference and that of the quantized stego-noise difference. We estimate the generalized Gaussian's parameters using grid search and Chi-square goodness of fit test, and estimate the variance of the stego-noise which in turn determines the length of hidden message by exploiting the relationship between the distribution variance of the cover image's pixel difference and that of the quantized stego-noise difference. We present experimental results which show that the proposed method is effective.

The rest of this paper is organized as follows. In Sect. 2, we first briefly review the stochastic modulation steganography, we then discuss the statistical models of the image's pixel difference before and after message embedding, and finally, we describe the estimation of the length of the hidden message in detail. Experimental results and analysis are given in Sect. 3. We conclude our work in Sect. 4.

2 Steganalysis of Stochastic Modulation

It is known that the pixel difference histogram of natural image can be modeled as a generalized Gaussian distribution (GGD) [13]. However, this may be not true for the distribution of the stego-image's pixel difference due to the stego-noise added by steganography. For the stochastic modulation steganography, we may assume that the stego-noise is independent from the cover image. Therefore, the distribution of stego-image's pixel difference is a convolution of the probabilistic distribution of the stego-noise difference and that of the cover image's pixel difference. Based on the independence assumption, we will derive an estimator to estimate the hidden message length through the following subsections.

Let $\{c_{i,j}\}$ and $\{s_{i,j}\}$ denote the cover image and the stego-image, respectively, where $c_{i,j} \in \{0, \cdots, 255\}$, $s_{i,j} \in \{0, \cdots, 255\}$, $i \in \{1, \cdots, \mathrm{M}\}$ and $j \in \{1, \cdots, \mathrm{N}\}$. The message m_k ($k = 1, \cdots, \mathrm{K}$, where K denotes the absolute length of secret message in bits) consists of a binary random sequence and $m_k \in \{+1, -1\}$. Let $n_{i,j}$ denote the stego-noise, which will be rounded off to the quantized stego-noise $z_{i,j}$ during embedding. The random variables ξ_{c}, ξ_{s}, ξ_{n}

and ξ_{z} model the cover image's pixel $c_{i,j}$, the stego-image's pixel $s_{i,j}$, the stego-noise $n_{i,j}$ and the quantized stego-noise $z_{i,j}$. Similarly, the cover image's pixel difference $dc_{i,j}$, the stego-image's pixel difference $ds_{i,j}$, the stego-noise difference $dn_{i,j}$, and the quantized stego-noise difference $dz_{i,j}$ are modeled as samples of the random variables $d\xi_{\mathrm{c}}$, $d\xi_{\mathrm{s}}$, $d\xi_{\mathrm{n}}$ and $d\xi_{\mathrm{z}}$.

2.1 Stochastic Modulation Steganography

Stochastic modulation steganography [10] adds stego-noise with a specific probability distribution in the cover image to embed the secret message. A steganography capacity as high as 0.8 bpp (bits per pixel) may be achieved with the use of a special parametric parity function. The parametric parity function $p\left(c_{i,j}, z_{i,j}\right)$ used in stochastic modulation steganography is required to satisfy the anti-symmetric property for $c_{i,j}$, i.e. $p\left(c_{i,j}+z_{i,j}, z_{i,j}\right) = -p\left(c_{i,j}-z_{i,j}, z_{i,j}\right)$ $(z_{i,j} \neq 0)$. The definition of the parity function proposed in [10] is given as follows.

(a). If $c_{i,j} \in [1, 2z_{i,j}]$, $p\left(c_{i,j}, z_{i,j}\right) = \begin{cases} (-1)^{c_{i,j}+z_{i,j}} & \text{if } z_{i,j} > 0, \\ 0 & \text{if } z_{i,j} = 0. \end{cases}$

(b). If $c_{i,j} \notin [1, 2z_{i,j}]$, $p\left(c_{i,j}, z_{i,j}\right)$ can be computed according to the anti-symmetric property and the above item (a).

The embedding procedure of stochastic modulation is described as below.

(1). Sequential or random visiting path and the stego-noise $n_{i,j}$ are generated using a secret key.

(2). For each pixel $c_{i,j}$ along the visiting path, one sample $n_{i,j}$ of the stego-noise ξ_{n} is rounded off to an integer $z_{i,j}$. If $z_{i.j} = 0$, the pixel $c_{i,j}$ is skipped and move to the next pixel in the visiting path, at the same time, the next stego-noise sample is input and rounded; If $z_{i.j} \neq 0$, the pixel $c_{i,j}$ will be modified according to the value of the parity function, i.e.

$$\begin{array}{rlll} \text{if} & p\left(c_{i,j}+z_{i,j}, z_{i,j}\right) = m_k & \text{then} & s_{i,j} = c_{i,j} + z_{i,j}, \\ \text{elseif} & p\left(c_{i,j}+z_{i,j}, z_{i,j}\right) = -m_k & \text{then} & s_{i,j} = c_{i,j} - z_{i,j}. \end{array}$$

During the embedding process, those pixels which may fall out of the range $[0, 255]$ will be truncated to the nearest value in this range with the desired parity.

2.2 Statistical Model of Difference

In this article, the distributions of the horizontal difference of images and stego noise are studied. The definitions of horizontal difference are given by (1).

$$\begin{aligned} dc_{i,j} &= c_{i,j} - c_{i,j+1} \quad, \\ ds_{i,j} &= s_{i,j} - s_{i,j+1} \quad, \\ dn_{i,j} &= n_{i,j} - n_{i,j+1} \quad, \\ dz_{i,j} &= z_{i,j} - z_{i,j+1} \quad, \end{aligned} \tag{1}$$