A Novel Upgraded Uniform Embedding Technique for JPEG Steganography

Savina Bansal, R. K. Bansal, and Rakesh Kumar

Abstract JPEG steganography finds extensive applications in the area of concealing the intelligent data in images. Minimal distortion embedding-based steganographic technique provides a precise distortion measure with feasible coding. This paper presents a new Upgraded Uniform Embedding Distortion (U^2ED) for JPEG steganography that tries to maintain the statistical imperceptibility in both spatial and DCT domains. A new distortion measure is proposed for both AC modes and DCT blocks in the proposed U^2ED scheme that aims to restrict the secret data into coarse and noisy regions, thus making it tough to be detected by the steganalyzer. Syndrome trellis codes are utilized to accomplish minimum distortion due to embedding of secret message. Effectiveness of the proposed scheme is analysed using BOSS base image database as used in literature using GFR steganalyzer combined with ensemble classifier. Experimental result based on 100 testing images indicates the goodness of the proposed U^2ED against Uniform Embedding Revisited Distortion (UERD) at different payloads in terms of probability of error. Also, the proposed scheme competes well with JPEG Universal Wavelet Relative Distortion (J-UNIWARD) for quality factor 75 and tends to outperform it for quality factor 95.

Keywords JPEG steganography · Uniform embedding · Syndrome trellis code

1 Introduction

Steganography aims towards achieving maximum statistical imperceptibility while hiding confidential information into a cover medium usually digital image. There is always a trade-off between statistical imperceptibility and sufficient payload capacity

S. Bansal (✉) · R. K. Bansal · R. Kumar
Giani Zail Singh Campus College of Engineering and Technology, Maharaja Ranjit Singh Punjab Technical University Bathinda, Punjab 151001, India
e-mail: savinabansal@mrsptu.ac.in

R. K. Bansal
e-mail: rkbansal@mrsptu.ac.in

A. K. Nagar et al. (eds.), *Intelligent Sustainable Systems*, Lecture Notes in Networks and Systems 333, https://doi.org/10.1007/978-981-16-6309-3_68

[14]. Therefore, the target of a steganographer is to better the statistical imperceptibility for a fixed payload. This can be attained by adaptively utilizing the location of the cover image pixels or coefficients for data embedding [16].

Currently, the minimal distortion embedding strategy is considered as the most secure approach for JPEG image steganography. This technique involves a well-designed cost function and a realistic coding scheme to lessen the distortion. The total cost for an additive cost function is computed as the summation of modification cost assigned to each cover elements. For content-based image steganography, large costs are allotted where the detection rate is expected to be higher (flat regions such as background) and smaller costs are allotted for textured or rough regions where the embedding changes will be tough to detect.

A practical coding approach to implement minimal distortion embedding is syndrome trellis coding [1]. It includes an easy to deal with optimizable cost function and a non-binary embedding method for coding the secret message while minimizing the additive cost. Syndrome trellis codes (STCs) have embedding rate near the rate-distortion limit, so the work left for steganographer is to design an effective distortion function.

2 Related Work

Uniform embedding as introduced by Guo et al. [2] for non-side informed JPEG steganography was later extended [3] for side informed JPEG steganography. Their scheme uses non-zero AC DCT coefficients for modification and uniformly distribute the embedding modification over quantized DCT coefficients by incorporating STCs. Holub et al. [4] presented a novel method called UNIWARD for content adaptive JPEG steganography. This method uses wavelet domain to derive distortion function for spatial (S-UNIWARD), side informed JPEG (SI-UNIWARD) and non-side informed JPEG (J-UNIWARD) domain and use all DCT coefficients as possible cover elements for modification. Their experimental results showed good security performance but constraint computational complexity. A major improvement over [3] was presented in Guo et al. [5] by considering the relative variations in coefficients of cover images and named this JPEG steganography technique as Uniform Embedding Revisited Distortion. DCT coefficients (including DC, zero AC, and non-zero AC) are utilized for embedding modification and uniformly spread the embedding changes over the cover elements. Their additive distortion function helps in achieving high embedding security and reduced computational time. Pan et al. [6] presented an improved distortion function for UERD [5] as IUERD. In their scheme, the distortion function efficiently utilized the mutual correlation between DCT blocks, which leads to more statistical undetectability and its effectiveness of securely embedding payload is checked using JRM steganalyzer.

Wang et al. [7] presented hybrid distortion function steganography (HDS) which exploit spatial and DCT domain to resist multidomain steganalysis. Their distortion function calculates block distortion as fluctuation in block and DCT coefficients changes as quantization step. In progression to exploit spatial and DCT domain for distortion calculation, Su et al. [8] decomposed a distortion measure known as GUED using an exponential model. Their results claim to have improved embedding security and reduced computation cost over UERD and UNIWARD. Further, Wei et al. [9] presented Residual Block Value method to measure the embedding cost due to modification in DCT block and quantization step for selection channel embedding risk. Hu et al. [10] presented a block entropy transformation scheme (BET) based on block entropy embedding (BEE). Their hybrid distortion function is designed using statistics of decompressed cover images and DCT coefficients.

3 Proposed Work

This work is an upgrade to an earlier work named Uniform Embedding Revisited Distortion (UERD) for non-side informed JPEG images [5], wherein the concept of generalized uniform embedding has been exploited. As this work is an upgradation of UERD so, its brief overview is given below for better understanding the work proposed here.

3.1 UERD Scheme

In UERD, the cost of modifying a coefficient depends on coefficient of variation parameter (CV is the fraction of standard deviation to the mean). The larger the CV, more the probability of changing coefficient and vice versa.

Let $x_{i,j}$ denotes a DCT coefficient at location (a, b) in the c DCT block. The cost function for $x_{i,j}$ is given as:

$$\rho_{i,j} = \rho(a, b), \text{(mode)} * \rho(m, n), \text{(block)} \quad (1)$$

where $\rho(a, b)$, (mode) and $\rho(m, n)$, (block) denotes the cost measures for the corresponding AC modes and DCT blocks, respectively. Specifically, $\rho(a, b)$, (mode) $=$ $q_{a,b}$ where $q_{a,b}$ indicates the quantization step for $x_{i,j}$, whereas $\rho(m, n)$, (block) is the function of energies of the $(m, n)^{\text{th}}$ block and its adjacent blocks. The block energy, say $D_{m,n}$, is defined as

$$D_{m,n} = \sum_{0}^{7} \sum_{0}^{7} \left| x_{a,b} \right| * q_{a,b} \quad (2)$$

here, $x_{a,b}$ specifies quantized DCT coefficient in $(m, n)^{\text{th}}$ DCT block. Thus, UERD only consider the DCT coefficients in build of distortion function $\rho_{i,j}$.

3.2 Proposed U^2ED Scheme

Here, new cost measures for AC modes and their corresponding DCT blocks are constructed depending on the statistics of both spatial domain and DCT domain. For DC mode, the distortion is calculated as the average of adjoining AC coefficients residing in the same DCT block. Inspired by J-UNIWARD [4], for $\rho(a, b)$, (mode), we compute the impact of unit modification on mode(a, b) calculated in wavelet domain as an alternative to quantization step $q_{a,b}$ in UERD. Let $W(a, b)$ denote the embedding modifications in spatial domain during modification of mode(a, b) by unity, and it can be obtained by utilizing two-dimensional inverse discrete cosine transformation (2D-IDCT) as:

$$W(a, b) = q_{a,b}.f(u, v; a, b) \tag{3}$$

where $q_{a,b}$ indicates quantization step for mode(a, b) and $f(u, v; a, b)$ is 2D-IDCT with index (u, v), i.e. $f(u, v; a, b) = \frac{Z_{a}.Z_{b}}{4}\cos\frac{\pi(2u+1)a}{16}\cos\frac{\pi(2v+1)b}{16}$, $Z_0 = 1/\sqrt{2}$, $Z_r = 1(r > 0), 0 \leq u, v \leq 7, 0 \leq a, b \leq 7$.

The relative changes in wavelet domain are calculated using8-tap Daubechies wavelet directional filter bank, $K = \{k_1, k_2, k_3\}$ where $k_1 = l.h'$,$k_2 = h.l'$, $k_3 = h.h'$. Here, l is low-pass and h is high-pass one-dimensional decomposition filter. The relative changes in wavelet coefficients at mode(a, b), are represented as the convolution ('$\odot$') between spatial modifications $W(a, b)$ and directional filter K:

$$\rho(a, b), (\text{mode}) = \sum_{k=1}^{3}\sum \text{abs}(W(a, b) \odot K_k) \tag{4}$$

The distortion function for DCT block is dependent on image content (i.e. the embedding changes are restricted in complicated textured regions) because of the fact that the Warden cannot distinguish content and noise from highly coarse regions. Also, from literature it is evident that most of the steganalyzer (Discrete Cosine Transform Residuals, Gabor Filter Residuals) are constructed based on the noise residuals model, features set of steganalysis and machine learning. Therefore, statistical properties of image model are exploited to design a steganalysis resistive distortion function. From this, we can deduce that residuals of image in spatial domain are a beneficial measure to calculate the distortion of DCT blocks. For building the distortion function $\rho(m, n)$, (block), we explore the Wiener Filter residuals (Wiener filter is a low-pass adaptive denoising filter) in spatial domain because of its better performance at lower computation cost, followed by a variance estimator as per [11].

$$\hat{R} = \hat{\sigma}_n^2 = \frac{r_n - \hat{r}_n^2}{p^2 - q} \tag{5}$$

where r_n is residuals value, $\hat{r}_n$ is least square estimation of residuals and $p^2 \times q$ is the size of matrix for parametric models. For numerical stability, a lower limit of 0.01 is set for estimated variance. To epitomize the cost measure for $(m, n)^{\text{th}}$ DCT block, $\hat{R}$ is divided into 8 by 8 block in arrangement with the other blocks in the same domain.

$$E_{m,n} = \sum_{a=0}^{7} \sum_{b=0}^{7} \hat{R}_{m,n}(a, b) \tag{6}$$

The cost measure for $(m, n)^{\text{th}}$ DCT block $\rho(m, n)$, (block) considers the cost measure $E_{m,n}$ and its eight neighbouring blocks as well, for all quantized DCT coefficient $x_{i,j}$ in the $(m, n)^{\text{th}}$ block:

$$\rho(m, n), \text{(block)} = \left(E_{m,n}\right)^{-1} + 0.25 * \sum_{d \in \bar{d}} (E_d)^{-1} \tag{7}$$

where $\bar{d} = \{(m-1, n-1), (m-1, n), (m-1, n+1), (m, n-1),$

$(m, n+1), (m+1, n-1), (m+1, n), (m+1, n+1)\}$.

In succinct, we propose a new distortion function $\rho_{i,j}$ for U^{2}ED as shown in Eq. (8). The embedding costs for both DC and AC coefficients are calculated as follow.

$$\rho_{i,j} = \begin{cases} 0.5 * (\rho_{i+1,j} + \rho_{i,j+1}), & \text{if}(a, b) = (0, 0) \\ \rho(a, b), \text{(mode)} * \rho(m, n), \text{(block)}, & \text{otherwise} \end{cases} \tag{8}$$

4 Experimental Results and Performance Analysis

The effectiveness of the offered scheme is analysed from experimental results. The U^{2}ED scheme is compared with the most closely resembling UERD and J-UNIWARD schemes.

4.1 Experimental Setup

Image sets from BOSS base ver. 1.01 [12] database is used to implement all experiments. The database contains 10,000 images in Portable Gray Map (PGM) format with each image having dimensionality of 512 by 512 pixels. All images are then JPEG compressed for quality factor 75 and quality factor 95 using MATLAB command imwrite. In our experiments, 100 images (50 for training and 50 for testing) are randomly selected from database. The embedding payload ranges from 0.1 to

0.5 bpnzac having a step size of 0.1 bpnzac. GFR-17000D [15] JPEG steganalyzer is employed to evaluate the pragmatic security performance of the proposed U^{2}ED JPEG steganographic scheme with UERD and J-UNIWARD. The ensemble classifier [13] with Fisher linear discriminant (FLD) base learner is utilized for classification of cover and stego images in our experiments. The total set of selected images are divided into two parts. One half of the images are used for training of ensemble classifier and the other half for testing of trained classifier. To evaluate the pragmatic security performance the minimum total probability of error P_e parameter is used. The P_e is the calculated as average of false alarm and missed detection of the ensemble classifier, under equal probability of occurring a cover or stego image, as is used in [5]

$$P_e = \min \frac{P_{\text{fa}} + P_{\text{md}}}{2} \tag{9}$$

where P_{fa} represents probability of false positive (indicating the cover image as stego image), P_{md} is probability of false negative (indicating a missed detection). The performance is evaluated using mean P_e over ten random tests with respect to relative payload. The higher value of P_e reflects strong empirical security.

4.2 *Performance of U^{2}ED*

We compare the proposed U^{2}ED scheme for non-side informed images with the UERD and J-UNIWARD schemes for JPEG steganography. Figure 1 and Fig. 2 show the security performance of the involved schemes against the GFR-17000D steganalyzer, for quality factor 75 and 95 respectively. As gathered from Fig. 1 for quality factor 75, U^{2}ED outperforms UERD by a sizeable margin of 8, 12 and 11% for

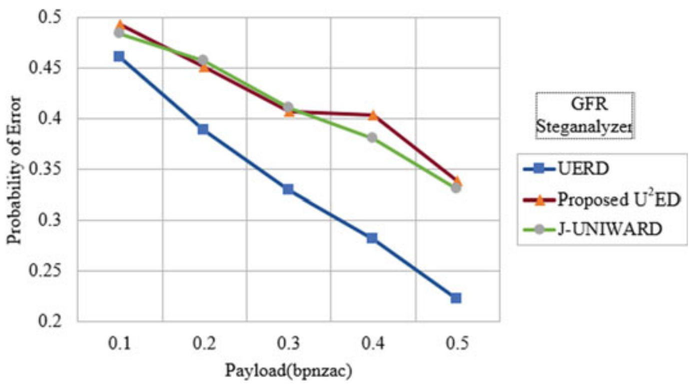

Fig. 1 Payload versus probability of error for JPEG quality factor 75

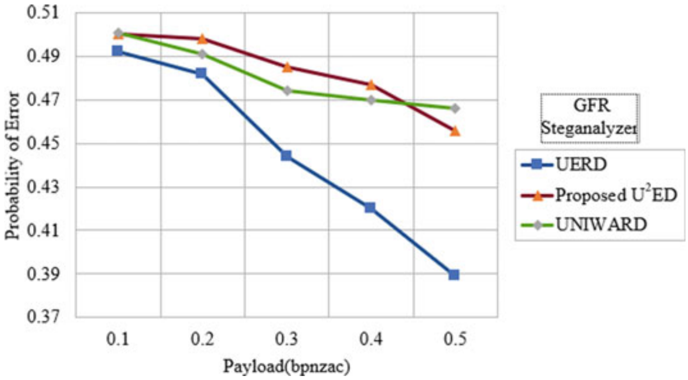

Fig. 2 Payload versus probability of error for JPEG quality factor 95

payload 0.3, 0.4, 0.5 bit per non-zero AC coefficients (bpnzac) respectively. Also, U^2ED has comparable security performance with J-UNIWARD for payloads less than 0.3 bpnzac and outperforms it for payloads greater than 0.3 bpnzac in all the tested JPEG images.

For quality factor 95, as gathered from Fig. 2 U^2ED outperforms UERD with a margin of 4%, 5%, and 7% for payloads 0.3, 0.4, and 0.5 bpnzac respectively. And, it outperforms J-UNIWARD for payload less than 0.5 bpnzac.

5 Conclusion

This paper presents a new distortion function named U^2ED for non-side informed JPEG steganography. This function utilizes the statistics of both spatial and DCT domains. The proposed distortion measure aims to maintain the statistical imperceptibility in both spatial and DCT domains by restricting the embedding changes in hard-to-detect regions. Experimental results indicate the superiority of U^2ED (in terms of surviving steganalysis attack) over UERD and J-UNIWARD for given set of training and testing data. More exhaustive results shall be presented in our future work.

References

1. Filler, T., Judas, J., Fridrich, J.: Minimizing additive distortion in steganography using syndrome-trellis codes. IEEE Trans. Inf. Forensics Secur. **6**(3), 920–935 (Sept. 2011). https://doi.org/10.1109/TIFS.2011.2134094
2. Guo, L., Ni, J., Shi, Y.Q.: An efficient JPEG steganographic scheme using uniform embedding. IEEE Int. Workshop Inf. Forensics Security (WIFS) **2012**, 169–174 (2012). https://doi.org/10.1109/WIFS.2012.6412644

3. Guo, L., Ni, J., Shi, Y.Q.: Uniform embedding for efficient JPEG steganography. IEEE Trans. Inf. Forensics Secur. **9**(5), 814–825 (May 2014). https://doi.org/10.1109/TIFS.2014.2312817
4. Holub, V., Fridrich, J., Denemark, T.: Universal distortion function for steganography in an arbitrary domain. EURASIP J. on Info. Security **2014**, 1 (2014). https://doi.org/10.1186/1687-417X-2014-1
5. Guo, L., Ni, J., Su, W., Tang, C., Shi, Y.: Using statistical image model for JPEG steganography: uniform embedding revisited. IEEE Trans. Inf. Forensics Secur. **10**(12), 2669–2680 (Dec. 2015). https://doi.org/10.1109/TIFS.2015.2473815
6. Pan, Y., Ni, J., Su, W.: "Improved uniform embedding for efficient JPEG steganography," in Proc. Int. Conf. Cloud Computing and Security Cham, Switzerland: Springer, pp. 125–133 (2016)
7. Wang, Z., Zhang, X., Yin, Z.: "Hybrid distortion function for JPEG steganography." J. Electron. Imaging **25**(5), 050501 (6 Sept 2016). https://doi.org/10.1117/1.JEI.25.5.050501
8. Su, W., Ni, J., Li, X., Shi, Y.: A new distortion function design for JPEG steganography using the generalized uniform embedding strategy. IEEE Trans. Circuits Syst. Video Technol. **28**(12), 3545–3549 (2018). https://doi.org/10.1109/TCSVT.2018.2865537
9. Wei, Q., Yin, Z., Wang, Z., et al.: Distortion function based on residual blocks for JPEG steganography. Multimed Tools Appl **77**, 17875–17888 (2018). https://doi.org/10.1007/s11042-017-5053-7
10. Hu, X., Ni, J., Shi, Y.: Efficient JPEG steganography using domain transformation of embedding entropy. IEEE Signal Process. Lett. **25**(6), 773–777 (June 2018). https://doi.org/10.1109/LSP.2018.2818674
11. Fridrich, J., Kodovský, J.: "Multivariate gaussian model for designing additive distortion for steganography." 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 2949–2953 (2013). https://doi.org/10.1109/ICASSP.2013.6638198
12. Bas, P., Filler, T., Pevný, T.: "Break our steganographic system": the ins and outs of organizing BOSS. In: Filler, T., Pevný, T., Craver, S., Ker, A. (eds.) Information Hiding. IH 2011. Lecture Notes in Computer Science, vol. 6958. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-24178-9_5
13. Kodovsky, J., Fridrich, J., Holub, V.: Ensemble classifiers for steganalysis of digital media. IEEE Trans. Inf. Forensics Secur. **7**(2), 432–444 (2012). https://doi.org/10.1109/TIFS.2011.2175919
14. Kaur, S., Bansal, S., Bansal, R.K.: Steganography and classification of image steganography techniques. Int. Conf. Comput. Sustain. Global Develop. (INDIACom) **2014**, 870–875 (2014). https://doi.org/10.1109/IndiaCom.2014.6828087
15. Song, X., Liu, F., Yang, C., Luo, X., Zhang, Y.: "Steganalysis of adaptive JPEG steganography using 2D Gabor filters." In: Proc. 3rd ACM Workshop Inf. Hiding Multimedia Security, pp. 15–23 (2015)
16. Kaur, S., Bansal, S., Bansal, R.K.: An efficient adaptive data hiding scheme for image steganography. In: Satapathy, S., Bhatt, Y., Joshi, A., Mishra, D. (eds.) Proceedings of the International Congress on Information and Communication Technology. Advances in Intelligent Systems and Computing, vol. 438. Springer, Singapore (2016). https://doi.org/10.1007/978-981-10-0767-5_40