Defining Joint Distortion for JPEG Steganography

Weixiang Li
CAS Key Laboratory of Electromagnetic Space Information
University of Science and Technology of China
Hefei, Anhui, China
wxli6049@mail.ustc.edu.cn

Weiming Zhang*
CAS Key Laboratory of Electromagnetic Space Information
University of Science and Technology of China
Hefei, Anhui, China
zhangwm@ustc.edu.cn

Kejiang Chen
CAS Key Laboratory of Electromagnetic Space Information
University of Science and Technology of China
Hefei, Anhui, China
chenkj@mail.ustc.edu.cn

Wenbo Zhou
CAS Key Laboratory of Electromagnetic Space Information
University of Science and Technology of China
Hefei, Anhui, China
welbeckz@mail.ustc.edu.cn

Nenghai Yu
CAS Key Laboratory of Electromagnetic Space Information
University of Science and Technology of China
Hefei, Anhui, China
ynh@ustc.edu.cn

ABSTRACT

Recent studies have shown that the non-additive distortion model of Decomposing Joint Distortion (*DeJoin*) can work well for spatial image steganography by defining joint distortion with the principle of Synchronizing Modification Directions (SMD). However, no principles have yet produced to instruct the definition of joint distortion for JPEG steganography. Experimental results indicate that SMD can not be directly used for JPEG images, which means that simply pursuing modification directions clustered does not help improve the steganographic security. In this paper, we inspect the embedding change from the spatial domain and propose a principle of Block Boundary Continuity (BBC) for defining JPEG joint distortion, which aims to restrain blocking artifacts caused by inter-block adjacent modifications and thus effectively preserve the spatial continuity at block boundaries. According to BBC, whether inter-block adjacent modifications should be synchronized or desynchronized is related to the DCT mode and the adjacent direction of inter-block coefficients (horizontal or vertical). When built into *DeJoin*, experiments demonstrate that BBC does help improve state-of-the-art additive distortion schemes in terms of relatively large embedding payloads against modern JPEG steganalyzers.

*Corresponding author

IH&MMSec'18, June 20–22, 2018, Innsbruck, Austria

ACM ISBN 978-1-4503-5625-1/18/06…$15.00
https://doi.org/10.1145/3206004.3206008

KEYWORDS

Steganography; JPEG image; non-additive model; joint distortion; block boundary continuity

ACM Reference Format:
Weixiang Li, Weiming Zhang, Kejiang Chen, Wenbo Zhou, and Nenghai Yu. 2018. Defining Joint Distortion for JPEG Steganography. In *Proceedings of 6th ACM Information Hiding and Multimedia Security Workshop (IH&MMSec'18).* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3206004.3206008

1 INTRODUCTION

Modern steganography is a science and art of covert communication that changes the original digital media slightly in order to hide secret messages without drawing suspicions from steganalysis [5, 13]. Currently, the most effective steganographic schemes are based on the framework of minimizing distortion, which defines the distortion as the sum of embedding cost at each individual cover element. And Syndrome-Trellis Codes (STCs) [4] provide a general and efficient coding method that can asymptotically approach the theoretical bound of average embedding distortion for arbitrary additive distortion function.

As a widely adopted format for image storage and transmission, JPEG steganography has become a research hotspot over the past few years. To date, there exist many content-adaptive algorithms designed for JPEG steganography, such as J-UNIWARD [10], UED [7], UERD [8], IUERD [15], HDS [18], RBV [19]. The embedding distortion of J-UNIWARD (UNIversal WAvelet Relative Distortion) [10] is computed as a sum of relative changes of coefficients in a directional filter bank decomposition of the decompressed cover image. Followed by the concept in spirit of "spread spectrum communication", UED (Uniform Embedding Distortion) [7] and UERD (Uniform Embedding Revisited Distortion) [8] with low complexity uniformly spread the embedding modifications to DCT coefficients of all possible magnitudes. IUERD (Improved UERD) [15] works quite well in the intersections

between smooth and texture regions by exploring the correlation among neighboring DCT blocks more efficiently. After decompressing the image, HDS (Hybrid DiStortion) [18] exploits block fluctuation via the prediction error of pixel and combines quantization step to form a hybrid distortion function, while RBV (Residual Block Value) [19] uses a wavelet filter bank to filter the decompressed image and obtains residual block values to measure block fluctuation, which can effectively identify complex discernible objects and their orientation from the spatial domain.

Above adaptive steganographic methods are based on additive distortion model, in which the modifications on cover elements are assumed to be independent and thus minimizing the overall distortions is equivalent to minimizing the sum of costs of all individual modified elements. Intuitively, non-additive distortion model is more suitable for natural images because the embedding changes on adjacent cover elements will interact mutually and the interplay among them would sometimes disturb the spatial continuity and correlation in natural images. Recent studies on spatial image steganography show that non-additive distortion models work best in resisting modern steganalysis equipped with high-dimensional features. Li et al. [14] and Denemark et al. [2] independently introduced a similar and effective strategy for exploiting the mutual impact of adjacent modifications. In [14], the cover image is decomposed into several sub-images, and additive distortion is individually minimized in each of the sub-images while the costs of cover elements within each sub-image are dynamically updated according to the modification directions of the embedded sub-images. The strategy used in [14] and [2] is generalized as "updating distortion" (abbreviated to $UpDist$) in this paper. Since that how to design efficient coding schemes for non-additive distortion function is commonly recognized as an important open problem for steganography by the academia [11], Zhang et al. [20] proposed a general framework called $DeJoin$ attempting to solve this problem, in which the joint distortion of cover element block is firstly defined and then decomposed into additive distortion on individual elements. It has been proved that $DeJoin$ can approach the lower bound of average joint distortion for a given payload. We mainly use $DeJoin$ for JPEG non-additive distortion steganography, and also combine it with $UpDist$ to enhance the steganographic security in this paper.

With the aforementioned non-additive models, finding some rules or principles for defining reasonable joint distortion or updating distortion availably in various kinds of covers has become a critical issue for non-additive distortion steganography. Regrettably, there emerges only one practical principle for defining non-additive distortion that works well for spatial images. The principle of Synchronizing Modification Directions (SMD) used in [2, 14, 20] aims to cluster modification directions of adjacent pixels by decreasing the costs on changes in the same direction and increasing that in the opposite direction. However, experimental results show that SMD could not be directly applied to JPEG images, meaning that simply pursuing the synchronization of modification directions among adjacent DCT coefficients (intra-block or inter-block) does not improve the security of steganography. Intuitively, the interplay among adjacent modifications is more complicated since changing one DCT coefficient will make diverse impacts on the whole 8×8 block. As for the most popular image format, it is still unclear how to define non-additive distortion for JPEG steganography.

In this paper, we present a principle called Block Boundary Continuity (BBC) for exploiting the interactive impact of changes between adjacent inter-block coefficients. Inspecting the embedding change from the spatial domain, the goal of the principle is to preserve spatial continuity at block boundaries via restraining blocking artifacts caused by adjacent modifications on inter-block coefficients, and thus effectively maintain the spatial continuity and neighboring relativity in natural images. According to BBC, the encouraged modification direction is not only related to the mode of DCT coefficient but also the adjacent direction of inter-block coefficients (horizontal or vertical), so the changed directions of some adjacent inter-block coefficients may be the same whereas others should be the opposite. When built into $DeJoin$, experimental results show that BBC can help improve the performances of recent additive distortion schemes in resisting the state-of-the-art JPEG steganalyzers.

The rest of this paper is organized as follows. In Section 2, the model of $DeJoin$ for non-additive distortion steganography are briefly reviewed. We elaborate the principle of BBC and show its effectiveness in resisting steganalysis via a simulation experiment in Section 3. The definition of JPEG joint distortion with BBC is provided in Section 4. Experimental results and comparisons are presented in Section 5. The paper is concluded in Section 6.

2 MODEL OF DECOMPOSING JOINT DISTORTION

In this paper, sets and matrices are written in boldface, and k-ary entropy function is denoted by $H_k(\pi_1, \cdots, \pi_k)$ for $\sum_{i=1}^{k} \pi_i = 1$. The embedding operation on cover element is ternary embedding with $\mathbf{I} = \{+1, -1, 0\}$, where 0 denotes no modification.

Previous adaptive steganography usually defines additive cost c_i on single cover element e_i for $i = 1, \cdots, n$. In the model of $DeJoin$ established in [20], joint distortion on element block need to be defined firstly according to the additive distortion and a specific principle. The joint distortion of each block is still additive, but it is unpractical to directly apply STCs because the number of modification patterns within a block is large that causes a high computational complexity. Therefore, $DeJoin$ carries out a two-round embedding strategy by decomposing the joint distortion into distortions on individual elements, and thus STCs can be used to embed message efficiently. Without loss of generality, Figure 1 illustrates the example of the decomposition coding process on the 1×2 block.

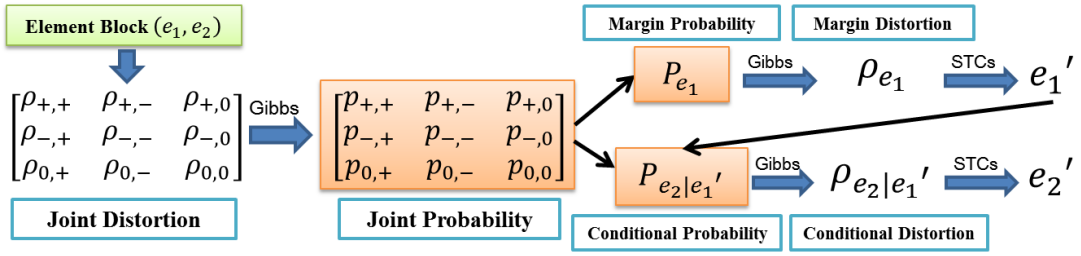


Figure 1: Illustration of *DeJoin* on 1×2 element block.

Assume that the joint distortion of the block $(e_{i,1}, e_{i,2})$ is $\rho^{(i)}(l, r)$, which denotes the distortion introduced by modifying $(e_{i,1}, e_{i,2})$ to $(e_{i,1} + l, e_{i,2} + r)$ for $(l, r) \in \mathbf{I}^2$. For a given message length L, following the maximum entropy principle, the optimal joint modification probability $\pi^{(i)}$ has a Gibbs distribution [3], and is given by

$$\pi^{(i)}(l,r) = \frac{\exp\big(-\lambda\rho^{(i)}(l,r)\big)}{\sum_{(p,q)\in\mathbf{I}^2}\exp\big(-\lambda\rho^{(i)}(p,q)\big)}, (l,r) \in \mathbf{I}^2, \quad (1)$$

which satisfies $L = \sum_{i=1}^{N} H_9(\pi^{(i)})$, where N is the number of element blocks.

In the first round, the margin probability $\pi_1^{(i)}$ on $e_{i,1}$ is calculated by

$$\pi_1^{(i)}(l) = \sum_{r\in\mathbf{I}} \pi^{(i)}(l,r), l \in \mathbf{I}. \quad (2)$$

As proved in [3], $\pi_1^{(i)}(l)$ can be transformed to the corresponding distortion $\rho_1^{(i)}(l)$ by

$$\rho_1^{(i)}(l) = \ln\frac{\pi_1^{(i)}(0)}{\pi_1^{(i)}(l)}, l \in \mathbf{I}, \quad (3)$$

and after that, ± 1 STCs can be applied to embed message into $e_{1,1}, \cdots, e_{i,1}, \cdots, e_{N,1}$ efficiently.

In the second round, the conditional probability $\pi_{2|l}^{(i)}(r)$ on $e_{i,2}$ is calculated by

$$\pi_{2|l}^{(i)} = \frac{\pi^{(i)}(l,r)}{\pi_1^{(i)}(l)}, r \in \mathbf{I}, l \in \mathbf{I}, \quad (4)$$

which denotes the probability such that $e_{i,2}$ is changed to $e_{i,2} + r$ under the condition of $e_{i,1}$ having been changed to $e_{i,1} + l$ in the first round. After transforming the conditional probability to the corresponding distortion as done in the first round, message is embedded into $e_{1,2}, \cdots, e_{i,2}, \cdots, e_{N,2}$ with ± 1 STCs.

Demonstrably, we will embed $L_1 = \sum_{i=1}^{N} H_3(\pi_1^{(i)})$ bits of message in the first round and $L_2 = \sum_{i=1}^{N}\sum_{l\in I}\pi_1^{(i)}(l)H_3(\pi_{2|l}^{(i)})$ in the second round. By chain rule, totally $L_1 + L_2 = \sum_{i=1}^{N} H_9(\pi^{(i)}) = L$ bits of message are embedded into the cover image. It has been proved in [20] that *DeJoin* can minimize the joint distortion defined on element blocks of any size, such as *DeJoin* on 2-element blocks (abbreviated to $DeJoin_2$) and 4-element blocks (abbreviated to $DeJoin_4$).

Obviously, *DeJoin* implicitly introduces non-additivity by distinguishing the joint distortions of different joint modification patterns in element block and executing a decomposition coding algorithm for embedding. Note that the definition of joint distortion is guided by some specific and instructive principles for various kinds of covers. We will discuss an effective principle for defining JPEG joint distortion in the next section.

3 THE PRINCIPLE OF BLOCK BOUNDARY CONTINUITY

Currently, there is no effective principle proposed to instruct the definition of joint distortion for JPEG steganography. In this section, we elaborate a principle called Block Boundary Continuity (BBC), which considers the spatial interactions of modifications on coefficients at the same DCT mode in adjacent blocks. The pair of coefficients at the same DCT mode in adjacent blocks is called the inter-block neighbors for short.

3.1 Embedding Change in Spatial Domain

In JPEG standard, the image is split into blocks of 8×8 pixels, and each block is converted to a frequency-domain representation by 2-D DCT transform

$$F(u,v) = \tfrac{1}{4}\xi(u)\xi(v)\left[\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y)\cdot\cos\tfrac{(2x+1)u\pi}{16}\cdot\cos\tfrac{(2y+1)v\pi}{16}\right], \quad (5)$$

where $f(\cdot)$ and $F(\cdot)$ are respectively the pixel value and the DCT coefficient. (x, y) represents the location of pixel in the spatial block where $x, y \in \{0, 1, \cdots, 7\}$ are respectively the row and column coordinate, and (u, v) is the DCT mode where $u, v \in \{0, 1, \cdots, 7\}$ are the horizontal and vertical spatial frequency respectively.

$$\xi(u), \xi(v) = \begin{cases} 1/\sqrt{2} & \text{if } u, v = 0 \\ 1 & \text{otherwise} \end{cases} \quad (6)$$

are the normalizing scale factors to make the transformation orthonormal. Then the quantized DCT coefficient $F_q(u, v)$ is computed with a selected quality factor and rounded to the nearest integer by

$$F_q(u,v) = round\Big(\frac{F(u,v)}{Q(u,v)}\Big), \quad (7)$$

where $Q(u,v)$ is the quantization step of the DCT mode(u,v), and JPEG steganography embeds message by modifying these quantized DCT coefficients. In view of the neighboring relativity in natural images, we attempt to inspect the embedding change from the spatial domain according to 2-D IDCT transform

$$f(x,y)=\frac{1}{4}\left[\sum_{u=0}^{7}\sum_{v=0}^{7}\xi(u)\xi(v)R(u,v)\cdot\cos\frac{(2x+1)u\pi}{16}\cdot\cos\frac{(2y+1)v\pi}{16}\right], \tag{8}$$

where $R(u,v)=F_q(u,v)\times Q(u,v)$ is the reconstructed approximate coefficient after inverse quantization. Suppose that a single quantized coefficient $F_q(u,v)$ is modified to $F_q^{'}(u,v)=F_q(u,v)+\Delta F$, and equivalently $R^{'}(u,v)=R(u,v)+\Delta R(u,v)=R(u,v)+Q(u,v)\times\Delta F$, the change on the spatial block can be computed by

$$\begin{aligned}&\Delta f_{u,v}(x,y,\Delta F)\\&\quad=f^{'}(x,y)-f(x,y)\\&\quad=\frac{1}{4}\xi(u)\xi(v)\cdot\Delta R(u,v)\cdot\cos\frac{(2x+1)u\pi}{16}\cdot\cos\frac{(2y+1)v\pi}{16}\\&\quad=\frac{1}{4}\xi(u)\xi(v)\cdot[Q(u,v)\times\Delta F]\cdot\cos\frac{(2x+1)u\pi}{16}\cdot\cos\frac{(2y+1)v\pi}{16}.\end{aligned} \tag{9}$$

For a given mode(u,v), we can draw the spatial change image $\Delta\mathbf{I_{u,v}}$, which consists of 64 gray values $\Delta f_{u,v}(x,y,\Delta F)$ representing positive or negative increment of the original pixel at position (x,y). We call $\Delta\mathbf{I_{u,v}}$ the DCT base image at mode(u,v). Without loss of generality, we select the quantization matrix $\mathbf{Q}$ of quality factor 75, and take $\Delta F=+1$ for illustration since $\Delta F=+1$ and $\Delta F=-1$ result in equal and merely opposite spatial change. Obviously, for quality factor 75 and $\Delta F=+1$, there are 64 DCT base images in total for all modes, and we draw them together in Figure 2(a). To clarify, the increments within each base image are normalized respectively, and like $\Delta\mathbf{I_{7,7}}$ in Figure 2(a), whiter means larger increment and darker means larger decrement within a base image.

In this paper, we just concentrate on the directions of changes on original pixel values, i.e. increase or decrease. When the magnitude is being neglected, the spatial change can be binarized to $\Delta f^B_{u,v}(x,y,\Delta F)=sgn\big(\Delta f_{u,v}(x,y,\Delta F)\big)$ by the sign function

$$sgn(\phi)=\begin{cases}1 & \text{if } \phi>0\\ 0 & \text{if } \phi=0\\ -1 & \text{if } \phi<0\end{cases}. \tag{10}$$

Since $sgn(\phi_1\phi_2)=sgn(\phi_1)\times sgn(\phi_2)$, the binarized spatial change

$$\begin{aligned}\Delta f^B_{u,v}(x,y,\Delta F)&=sgn\Big(\Delta f_{u,v}(x,y,\Delta F)\Big)\\&=sgn\big(\Delta F\cdot\cos\frac{(2x+1)u\pi}{16}\cdot\cos\frac{(2y+1)v\pi}{16}\big)\\&=\begin{cases}1 & \Longrightarrow increase\\ -1 & \Longrightarrow decrease\end{cases}.\end{aligned} \tag{11}$$

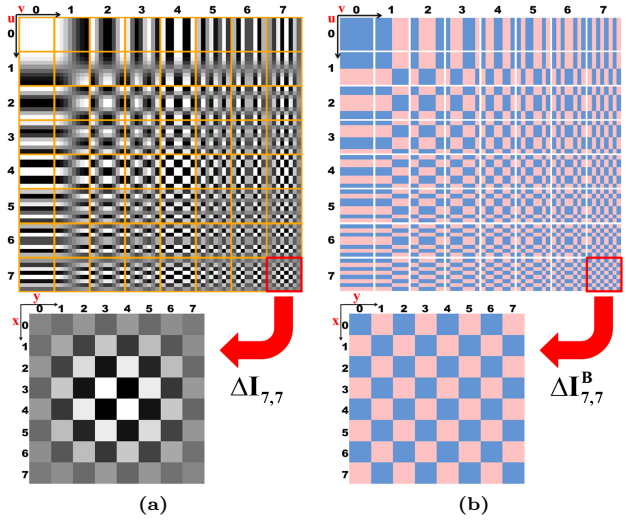


Figure 2: Two kinds of DCT base images associated with quality factor 75 and $\Delta F=+1$, and the corresponding examples of mode(7,7). (u,v) represents the DCT mode where $u,v\in\{0,1,\cdots,7\}$ are respectively the horizontal and vertical spatial frequency, and (x,y) represents the location of pixel in the spatial block where $x,y\in\{0,1,\cdots,7\}$ are respectively the row and column coordinate. (a) 64 base images $\Delta\mathbf{I_{u,v}}$, and a close-up of $\Delta\mathbf{I_{7,7}}$. (b) 64 binarized base images $\Delta\mathbf{I^B_{u,v}}$ with blue representing increase and red representing decrease, and a close-up of $\Delta\mathbf{I^B_{7,7}}$.

We also make $\Delta F=+1$ and plot 64 binarized base images $\Delta\mathbf{I^B_{u,v}}$ in Figure 2(b) with blue representing increase and red representing decrease. From $\Delta\mathbf{I^B_{u,v}}$, it is easy to make out the positive or negative change patterns on the spatial block when the single DCT coefficient $F_q(u,v)$ is modified by $\Delta F=+1$. For instance, modifying the DC coefficient $F_q(0,0)$ by $+1$ generates increases on all the 8×8 pixels; modifying $F_q(0,1)$ causes increases on the left half and decreases on the right half of the spatial block, while modifying $F_q(1,0)$ causes increases on the upper half and decreases on the lower half of the spatial block. Obviously, with the mode being higher, $\Delta\mathbf{I^B_{u,v}}$ becomes more complicated in horizontal and vertical direction.

3.2 Impact of Simultaneous Modifications on Inter-block Neighbors

In the process of steganography, the cases of simultaneously modifying inter-block neighbors are possible and common, especially among coefficients of low frequency in textured regions, for that we need to observe the combined influence on adjacent spatial blocks. Denote the inter-block neighbors at mode(u,v) by $\big(F_q^1(u,v),F_q^2(u,v)\big)$ and the simultaneous modifications on them by $(\Delta F_1,\Delta F_2)$. The simultaneous modifications contain four pairs of nonzero joint modification patterns, i.e., $(\Delta F_1,\Delta F_2)\in\big\{(+1,+1),(+1,-1),(-1,+1),(-1,-1)\big\}$,

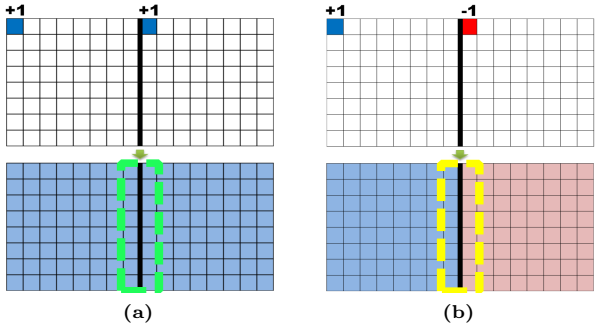


Figure 3: Combined impact on adjacent spatial blocks when modifying horizontal inter-block neighbors at mode DC with two kinds of joint modification patterns. (a) Pattern (+1,+1). (b) Pattern (+1,-1).

where $(+1,+1),(-1,-1)$ correspond to synchronizing modification directions and $(+1,-1),(-1,+1)$ denote desynchronizing modification directions. Without loss of generality, we take $(+1,+1)$ and $(+1,-1)$ as examples of synchronization and desynchronization, respectively.

As illustrated in Figure 3, when modifying the horizontal $\left(F_q^1(0,0),F_q^2(0,0)\right)$, the joint modification pattern of $(+1,+1)$ simultaneously increases pixel values of both adjacent blocks and thus preserves the consistency of boundary between the adjacent blocks, whereas $(+1,-1)$ causes increases on the left block but decreases on the right block, which leads to a discontinuity at block boundary, i.e. blocking artifact. In view of the neighboring relativity in natural images, $(+1,-1)$ is notably unreasonable because it breaks the local continuity on adjacent blocks and brings about blocking artifact, which would be captured and utilized by steganalysis as well. Hence for a higher steganographic security, $(+1,+1)$ is encouraged and $(+1,-1)$ should be avoided for DC inter-block neighbors.

Similarly, we need to maintain spatial continuity at block boundaries if simultaneously modifying horizontal or vertical inter-block neighbors at other modes. To clarify, the block boundary consists of two columns or rows bordering two adjacent spatial blocks, i.e., the union of the $7th$ column in the left block and the $0th$ column in the right for the horizontal, or the union of the $7th$ row in the upper block and the $0th$ row in the lower for the vertical.

As illustrated in Figure 2(b), since $\Delta\mathbf{I}^{\mathbf{B}}_{\mathbf{u,v}}$ are distinct with different DCT modes, $(+1,+1)$ may be not invariably suitable for each mode (differs from SMD in spatial images). Figure 4 displays the combined impacts on adjacent blocks when modifying horizontal or vertical $\left(F_q^1(0,1),F_q^2(0,1)\right)$ and $\left(F_q^1(1,0),F_q^2(1,0)\right)$, where $(+1,-1)$ seems to be more preferable for some modes. In Figure 4(a)-(b) of horizontal inter-block neighbors, $(+1,-1)$ preserves the spatial continuity at block boundary for mode(0,1) but breaks that for mode(1,0), so it should be encouraged for mode(0,1) but discouraged for mode(1,0). The results about the vertical can be observed from Figure 4(c)-(d), and Table 1 reports the

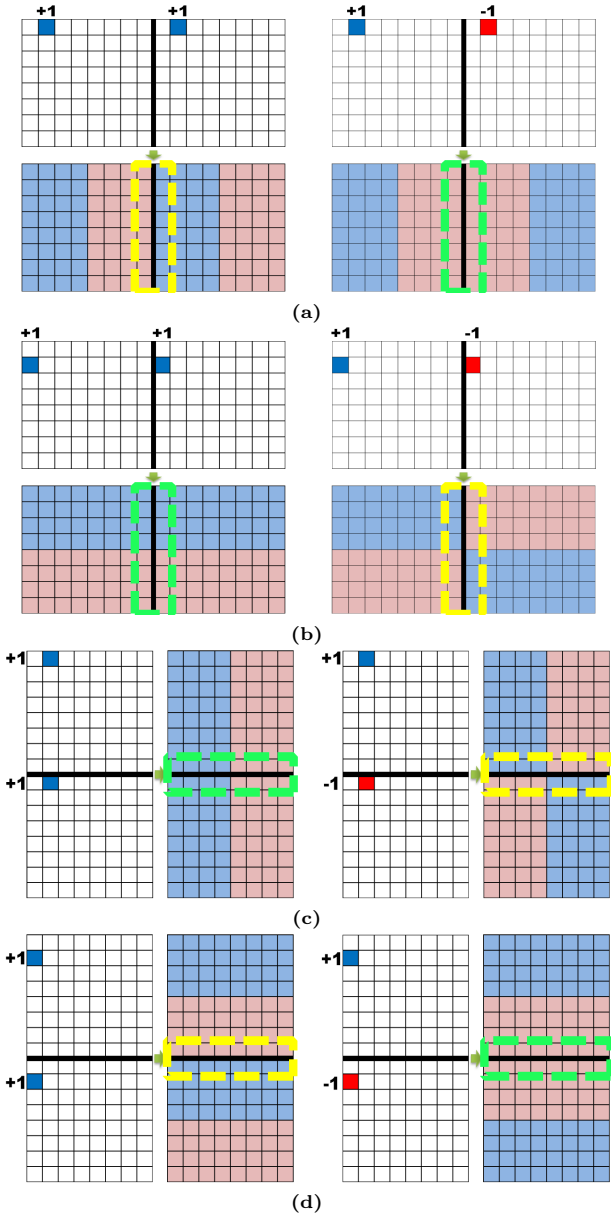


Figure 4: Combined impact on adjacent spatial blocks when modifying horizontal or vertical inter-block neighbors at mode(0,1) or mode(1,0) with (+1,+1) and (+1,-1). (a) Horizontal inter-block neighbors at mode(0,1). (b) Horizontal inter-block neighbors at mode(1,0). (c) Vertical inter-block neighbors at mode(0,1). (d) Vertical inter-block neighbors at mode(1,0).

corresponding encouraged joint modification patterns, which indicates that encouraged joint modification patterns are related to the mode of coefficient and the adjacent direction of inter-block neighbors.

Table 1: The encouraged joint modification pattern on horizontal or vertical inter-block neighbors at mode(0,1) or mode(1,0).

Mode	Horizontal	Vertical
(0,1)	$(+1,-1)$	$(+1,+1)$
(1,0)	$(+1,+1)$	$(+1,-1)$

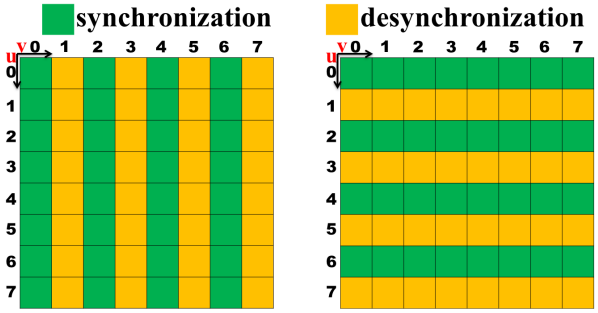


Figure 5: Encouraged joint modification patterns of 64 DCT modes with two types of adjacent direction, where green represents synchronization and orange represents desynchronization. (a) Horizontal inter-block neighbors. (b) Vertical inter-block neighbors.

3.3 Principle of BBC

Through observation from $\Delta\mathbf{I}^{\mathbf{B}}_{\mathbf{u,v}}$, we summarize the strategy associated with encouraged joint modification patterns of all the 64 DCT modes and two adjacent directions of inter-block neighbors, as shown in Figure 5. The strategy is named the principle of Block Boundary Continuity (BBC) as follows.

Principle of BBC.

(i) *For horizontal* $\left(F_q^1(u,v), F_q^2(u,v)\right)$ *at mode(u,v), the encouraged joint modification pattern*

$$(\Delta F_1, \Delta F_2) = \begin{cases} (+1,+1) \text{ and } (-1,-1) & \text{for } v = 0,2,4,6 \\ (+1,-1) \text{ and } (-1,+1) & \text{for } v = 1,3,5,7 \end{cases}.$$

(ii) *For vertical* $\left(F_q^1(u,v), F_q^2(u,v)\right)$ *at mode(u,v), the encouraged joint modification pattern*

$$(\Delta F_1, \Delta F_2) = \begin{cases} (+1,+1) \text{ and } (-1,-1) & \text{for } u = 0,2,4,6 \\ (+1,-1) \text{ and } (-1,+1) & \text{for } u = 1,3,5,7 \end{cases}.$$

According to the principle of BBC, it is encouraged to synchronize modification directions at modes of $v = 0,2,4,6$ and desynchronize that at modes of $v = 1,3,5,7$ for horizontal inter-block neighbors, and meanwhile for vertical inter-block neighbors, synchronization should be encouraged at modes of $u = 0,2,4,6$ and desynchronization at modes of $u = 1,3,5,7$. With BBC, the neighboring relativity in natural images is maintained, thus the modifications in DCT domain would be more secure.

Here we take modifications on horizontal inter-block neighbors as an example to prove the correctness of the BBC principle. We further derive the formulas for judging whether the joint modification pattern should be encouraged or not, which are needful for the definition of joint distortion.

Proof. The block boundary of horizontally adjacent blocks is the union of the $7th$ column in the left block and the $0th$ column in the right block. Let $\Phi_{u,v}(x,\Delta F_1,\Delta F_2) = \Delta f^B_{u,v}(x,7,\Delta F_1)\cdot\Delta f^B_{u,v}(x,0,\Delta F_2)$ determine the consistency of the binarized spatial changes at two columns, and $\Phi_{u,v}(x,\Delta F_1,\Delta F_2) = 1$ means continuity while conversely $\Phi_{u,v}(x,\Delta F_1,\Delta F_2) = -1$ means discontinuity. From (11),

$$\begin{aligned}
&\Phi_{u,v}(x,\Delta F_1,\Delta F_2) \\
&= sgn\big(\Delta F_1\cdot\Delta F_2\cdot\cos^2\frac{(2x+1)u\pi}{16}\cdot\cos\frac{15v\pi}{16}\cdot\cos\frac{v\pi}{16}\big) \\
&= sgn\big(\Delta F_1\cdot\Delta F_2\cdot\cos\frac{15v\pi}{16}\cdot\cos\frac{v\pi}{16}\big) \\
&= sgn\big(\Delta F_1\cdot\Delta F_2\cdot\cos(v\pi-\frac{v\pi}{16})\cdot\cos\frac{v\pi}{16}\big) \\
&= sgn\big(\Delta F_1\cdot\Delta F_2\cdot[\cos v\pi\cdot\cos\tfrac{v\pi}{16}+\sin v\pi\cdot\sin\tfrac{v\pi}{16}]\cdot\cos\tfrac{v\pi}{16}\big) \\
&= sgn\big(\Delta F_1\cdot\Delta F_2\cdot\cos v\pi\cdot\cos^2\frac{v\pi}{16}\big) \\
&= sgn\big(\Delta F_1\cdot\Delta F_2\cdot\cos v\pi\big) \\
&= \Delta F_1\cdot\Delta F_2\cdot\cos v\pi \\
&= \begin{cases} \Delta F_1\cdot\Delta F_2 & \text{if } v = 0,2,4,6 \\ -\Delta F_1\cdot\Delta F_2 & \text{if } v = 1,3,5,7 \end{cases}.
\end{aligned}$$

So, taking $\Delta F_1\cdot\Delta F_2 = 1$ (synchronization) for $v = 0,2,4,6$ and $\Delta F_1\cdot\Delta F_2 = -1$ (desynchronization) for $v = 1,3,5,7$, can maintain the continuity $\big(\Phi_{u,v}(x,\Delta F_1,\Delta F_2) = 1\big)$, which derives the principle about horizontal inter-block neighbors. □

Since $\Phi_{u,v}(x,\Delta F_1,\Delta F_2)$ is independent of x, we denote it by

$$\Phi^{hor}_{u,v}(\Delta F_1,\Delta F_2) = \Delta F_1\cdot\Delta F_2\cdot\cos v\pi. \tag{12}$$

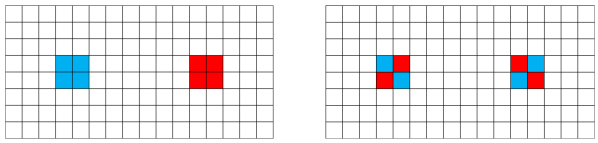
Similarly, for vertical inter-block neighbors, $\Phi_{u,v}(y,\Delta F_1,\Delta F_2) = \Delta f^B_{u,v}(7,y,\Delta F_1)\cdot\Delta f^B_{u,v}(0,y,\Delta F_2)$ is denoted by

$$\Phi^{ver}_{u,v}(\Delta F_1,\Delta F_2) = \Delta F_1\cdot\Delta F_2\cdot\cos u\pi. \tag{13}$$

Obviously, $\Phi^{hor}_{u,v}(\Delta F_1,\Delta F_2), \Phi^{ver}_{u,v}(\Delta F_1,\Delta F_2) = 1$ correspond to encouraged joint modification patterns, and conversely $\Phi^{hor}_{u,v}(\Delta F_1,\Delta F_2), \Phi^{ver}_{u,v}(\Delta F_1,\Delta F_2) = -1$ correspond to discouraged joint modification patterns, which will be used in the definition of joint distortion.

3.4 Verifying the Practicability of BBC

To verify whether the BBC principle is reasonable for JPEG steganography, we perform a simulation as follows. Firstly, 10,000 gray-scale images of size 512×512 pixels from BOSSBase1.01 [1] are JPEG compressed with quality factor 75, and coefficients at mode(u,v) are extracted from DCT blocks to form a sub-image of size 64×64, which is then divided into non-overlapping blocks of size 8×16. Secondly, a noise pattern is added to each image block to simulate the effect of data embedding. Two kinds of noise patterns in Figure 6 are



(a) (b)

Figure 6: Two kinds of noise patterns where blue means +1 and red means -1. (a) Noise pattern A. (b) Noise pattern B.

Table 2: The MMD and the steganalytic performance of individually adding noise pattern A or B on coefficients at mode DC or mode(1,1).

Mode	Pattern	MMD	Testing Error
DC	A	2.494×10^{-3}	0.2540
	B	5.716×10^{-3}	0.0777
(1,1)	A	7.446×10^{-3}	0.0478
	B	5.443×10^{-3}	0.0879

used, where modification directions are the same in Pattern A but the opposite in Pattern B both in horizontal and vertical directions. Thirdly, the DCTR-8,000D [9] features of the first 1,000 images are obtained and the MMD (maximum mean discrepancy) [16], which quantifies the distance between the feature set of cover images and that of stego images, is computed for each noise pattern. Finally, we employ modern steganalyzer with DCTR-8,000D to evaluate the performance of each noise pattern on resisting steganalysis. Generally, a lower MMD or a higher classification error corresponds to a higher level of security.

Without loss of generality, we only demonstrate the example of noise addition on coefficients at mode DC or mode(1,1) individually. The results in Table 2 show that noise addition of Pattern A is more secure for mode DC while Pattern B is less harmful for mode(1,1), which ideally conforms to the strategy in Figure 5. We also perform simulations with BBC on coefficients at other modes and come to the same conclusion. Consequently, it is reasonable to employ the BBC principle on directing the definition of joint distortion for JPEG steganography.

4 DEFINING JPEG JOINT DISTORTION WITH BBC

Under the guidance of the BBC principle, we define joint distortion for DCT coefficients at the same mode. Firstly, the initial distortion on single coefficient is defined by state-of-the-art additive methods. Secondly, coefficients at each mode are extracted from DCT blocks to form a sub-image $\mathbf{D_{u,v}}$ of size $\frac{m}{8}\times\frac{n}{8}$ (assume the size of the image is $m\times n$), which is then divided into non-overlapping joint blocks of the needed size. The joint distortion on joint block in each sub-image $\mathbf{D_{u,v}}$ is computed on the basis of the initial distortion and the BBC principle. Finally, joint distortions from all $\mathbf{D_{u,v}}$s are composed into a sequence of joint distortion that can be sent into *DeJoin*.

The division of a JPEG image into 1×2 joint blocks (abbreviated to 2-Coeffs) and 2×2 joint blocks (abbreviated to 4-Coeffs) is depicted in Figure 7. For 2-Coeffs, $\mathbf{D_{u,v}}$ is divided into the joint block sequence $B^{(i)}_{u,v}=(d^{i,1}_{u,v},d^{i,2}_{u,v})$ for $i=1,\cdots,N$ where $N=(\frac{m}{8}\times\frac{n}{8})/2$. For 4-Coeffs, $\mathbf{D_{u,v}}$ is divided into the joint block sequence $B^{(i)}_{u,v}=\begin{pmatrix} d^{i,1}_{u,v}, d^{i,2}_{u,v} \\ d^{i,3}_{u,v}, d^{i,4}_{u,v}\end{pmatrix}$ for $i=1,\cdots,N$ where $N=(\frac{m}{8}\times\frac{n}{8})/4$.

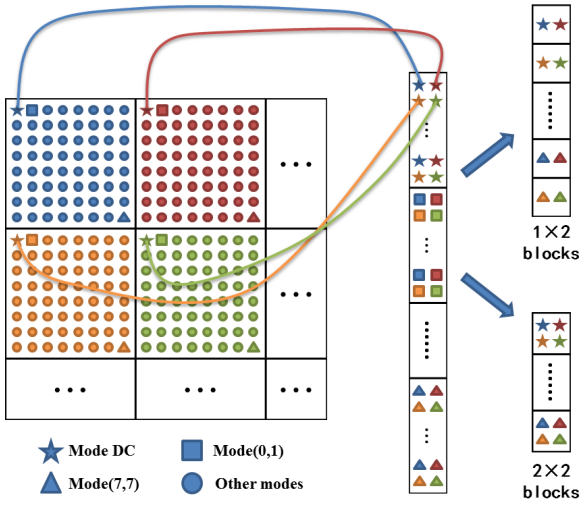


Figure 7: The division of a JPEG image into 1×2 joint blocks and 2×2 joint blocks.

4.1 Defining Horizontal 2-Coeffs Joint Distortion

For the joint block $B^{(i)}_{u,v}=(d^{i,1}_{u,v},d^{i,2}_{u,v})$ in $\mathbf{D_{u,v}}$, denote the initial cost on $d^{i,1}_{u,v}$ by $c^{i,1}_{u,v}(\Delta F_1)$ for $\Delta F_1\in\mathbf{I}$ and the initial cost on $d^{i,2}_{u,v}$ by $c^{i,2}_{u,v}(\Delta F_2)$ for $\Delta F_2\in\mathbf{I}$. The joint distortion on $B^{(i)}_{u,v}$ is defined by

$$\rho^{(i)}_{u,v}(\Delta F_1,\Delta F_2)=\omega_{u,v}(\Delta F_1,\Delta F_2)\times\left(c^{i,1}_{u,v}(\Delta F_1)+c^{i,2}_{u,v}(\Delta F_2)\right), \tag{14}$$

and the scaling function $\omega_{u,v}(\Delta F_1,\Delta F_2)$ is computed by

$$\omega_{u,v}(\Delta F_1,\Delta F_2)=\begin{cases}1/\alpha & \text{if } \Phi^{hor}_{u,v}(\Delta F_1,\Delta F_2)=1\\ \alpha & \text{if } \Phi^{hor}_{u,v}(\Delta F_1,\Delta F_2)=-1\\ 1 & \text{otherwise}\end{cases}, \tag{15}$$

where $\alpha>1$ is to differentiate the costs of encouraged and discouraged joint modification patterns. According to (12), $\Phi^{hor}_{u,v}(\Delta F_1,\Delta F_2)=1$ corresponds to the encouraged joint modification patterns, of which the costs will be reduced by dividing α, and $\Phi^{hor}_{u,v}(\Delta F_1,\Delta F_2)=-1$ corresponds to the discouraged joint modification patterns, of which the costs will be enlarged by multiplying α. A larger α means a better

differentiation but inevitably causes more modifications. Since too many modifications would make a greatly bad influence on steganographic security, α could not be too large. We set $\alpha = 1.5$ experimentally, and embed message by using $DeJoin_2$ on 2-Coeffs from all $\mathbf{D_{u,v}}$s to minimize the joint distortion defined in (14).

4.2 Combining $DeJoin_2$ with $UpDist$

Since the joint distortion on horizontal 2-Coeffs only reflects the mutual impact of modifications in horizontal direction, we can incorporate the mutual impact in vertical direction by applying $UpDist$ upon 2-Coeffs if taking 2-Coeffs as a super-coefficient. To do that, we firstly embed half of payload into the 2-Coeffs in *odd* rows by using $DeJoin_2$. As illustrated in Figure 8, the joint distortion on the 2-Coeffs in *even* rows will be updated according to the changed results of the *odd* 2-Coeffs (in the same $\mathbf{D_{u,v}}$) above and under it. We finally embed the rest payload into the *even* 2-Coeffs by using $DeJoin_2$ on the updated joint distortion that considers the impact of modifications both in horizontal and vertical directions. The combination of $UpDist$ and $DeJoin_2$ is abbreviated to $UpDist$-$DeJoin_2$.

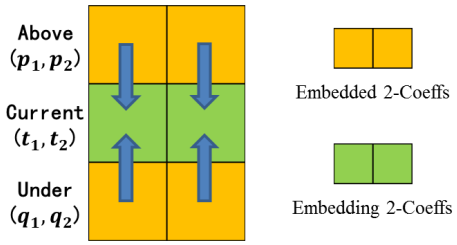


Figure 8: Illustration of $UpDist$ upon 2-Coeffs.

Assume that in the first round, the above *odd* block (p_1, p_2) has been changed to $(p_1+\Delta F_1', p_2+\Delta F_2')$ and the under *odd* block (q_1, q_2) has been changed to $(q_1 + \Delta F_1'', q_2 + \Delta F_2'')$, and then the joint distortion $\rho_{u,v}(\Delta F_1, \Delta F_2)$ on the current *even* block (t_1, t_2) will be updated to $\rho_{u,v}^{update}(\Delta F_1, \Delta F_2)$ by

$$\rho_{u,v}^{update}(\Delta F_1, \Delta F_2) = \prod_{(\Delta_1,\Delta_2)} \varepsilon_{u,v}(\Delta_1, \Delta_2) \times \rho_{u,v}(\Delta F_1, \Delta F_2), \tag{16}$$

where

$$\varepsilon_{u,v}(\Delta_1, \Delta_2) = \begin{cases} \beta & \text{if } \Phi_{u,v}^{ver}(\Delta_1, \Delta_2) = 1 \\ 1 & \text{otherwise} \end{cases}, \tag{17}$$

and

$$(\Delta_1, \Delta_2) \in \left\{(\Delta F_1, \Delta F_1'), (\Delta F_1, \Delta F_1''), (\Delta F_2, \Delta F_2'), (\Delta F_2, \Delta F_2'')\right\}$$

corresponds to four pairs of vertically adjacent modifications. With (13), $\beta < 1$ has the same effect as α in (15), and we set $\beta = 0.6$ experimentally in this paper.

4.3 Defining 4-Coeffs Joint Distortion

For the joint block $B_{u,v}^{(i)} = \begin{pmatrix} d_{u,v}^{i,1}, d_{u,v}^{i,2} \\ d_{u,v}^{i,3}, d_{u,v}^{i,4} \end{pmatrix}$ in $\mathbf{D_{u,v}}$, denote the initial cost on $d_{u,v}^{i,j}$ by $c_{u,v}^{i,j}(\Delta_j)$ for $j \in \{1, 2, 3, 4\}$. The joint distortion on $B_{u,v}^{(i)}$ is defined by

$$\rho_{u,v}^{(i)}(\Delta_1, \Delta_2, \Delta_3, \Delta_4) = \omega_{u,v}(\Delta_1, \Delta_2, \Delta_3, \Delta_4) \times \sum_{j=1}^{4} c_{u,v}^{i,j}(\Delta_j). \tag{18}$$

The scaling function $\omega_{u,v}(\Delta_1, \Delta_2, \Delta_3, \Delta_4)$ is a ratio between the number of discouraged modification pairs (determined by $o_{u,v}(\Delta_p, \Delta_q)$) and the number of encouraged modification pairs (determined by $s_{u,v}(\Delta_p, \Delta_q)$), which is computed by

$$\omega_{u,v}(\Delta_1, \Delta_2, \Delta_3, \Delta_4) = \frac{\theta + \sum_{(p,q)} o_{u,v}(\Delta_p, \Delta_q)}{\theta + \sum_{(p,q)} s_{u,v}(\Delta_p, \Delta_q)}, \tag{19}$$

where

$$s_{u,v}(\Delta_p,\Delta_q)=\begin{cases} 1 \text{ if} \big(mod(p+q,2)=1 \;\&\&\; \Phi_{u,v}^{hor}(\Delta_p,\Delta_q)=1\big) \\ 1 \text{ if} \big(mod(p+q,2)=0 \;\&\&\; \Phi_{u,v}^{ver}(\Delta_p,\Delta_q)=1\big) \\ 0 \text{ otherwise} \end{cases}, \tag{20}$$

$$o_{u,v}(\Delta_p,\Delta_q)=\begin{cases} 1 \text{ if} \big(mod(p+q,2)=1 \;\&\&\; \Phi_{u,v}^{hor}(\Delta_p,\Delta_q)=-1\big) \\ 1 \text{ if} \big(mod(p+q,2)=0 \;\&\&\; \Phi_{u,v}^{ver}(\Delta_p,\Delta_q)=-1\big) \\ 0 \text{ otherwise} \end{cases}, \tag{21}$$

and $(p, q) \in \big\{(1, 2), (1, 3), (2, 4), (3, 4)\big\}$ corresponds to four pairs of horizontally $\big(\text{mod}\,(p + q, 2) = 1\big)$ or vertically $\big(\text{mod}\,(p + q, 2) = 0\big)$ adjacent modifications within $B_{u,v}^{(i)}$. With (12) and (13), $\theta > 1$ has the same effect as α and β. We set $\theta = 3$ experimentally and embed message by using $DeJoin_4$ on 4-Coeffs from all $\mathbf{D_{u,v}}$s to minimize the joint distortion defined in (18).

5 EXPERIMENTAL RESULTS AND ANALYSIS

In this section, experimental results are presented to demonstrate the feasibility and effectiveness of the BBC principle. We compare the performances of BBC-based schemes with several state-of-the-art additive schemes, including UERD [8], J-UNIWARD [10], IUERD [15], HDS [18] and RBV [19], on resisting the detection of DCTR-8,000D [9] and GFR-17,000D [17] under several quality factors.

5.1 Experiment Setup

All experiments are conducted on BOSSBase 1.01 [1], which contains 10,000 gray-scale images of size 512×512 pixels. All of the images are compressed into JPEG domain with quality factor $QF = 50, 75, 90$ respectively, which are then adopted as datasets for experimental comparisons. We replace STCs with an optimal embedding simulator [6] to reduce experimental complexity. The payloads range from 0.1 to 0.5 bpnzac (bit per nonzero AC coefficient) with a step of 0.1 bpnzac. The detector is trained by using state-of-the-art DCTR-8,000D [9] and GFR-17,000D [17] with the FLD ensemble [12] by default, which minimizes the total classification error probability

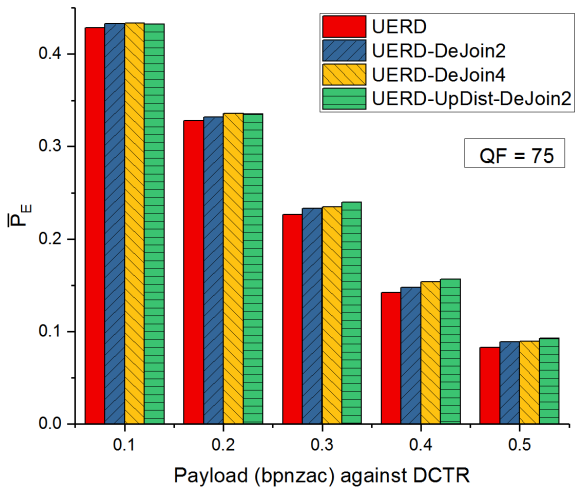


Figure 9: Detection errors for UERD and the corresponding BBC-based schemes against DCTR under QF=75.

under equal priors $P_E = \min_{P_{FA}} \frac{1}{2}(P_{FA} + P_{MD})$ where P_{FA} and P_{MD} are the false-alarm probability and the missed-detection probability respectively. The ultimate security is qualified by average error rate $\overline{P_E}$ averaged over 10 random 5000/5000 splits of the dataset, and larger $\overline{P_E}$ means stronger security.

5.2 Comparison and Visualization of $DeJoin_2$, $DeJoin_4$ and $UpDist\text{-}DeJoin_2$

Taking UERD as the initial distortion, we compare the steganographic securities of $DeJoin_2$, $DeJoin_4$ and $UpDist\text{-}DeJoin_2$ on resisting DCTR-8,000D with payloads of 0.1-0.5 bpnzac under $QF = 75$. As reported in Figure 9, three BBC-based schemes can outperform UERD, and because of incorporating the mutual impact of modifications in vertical direction, $DeJoin_4$ and $UpDist\text{-}DeJoin_2$ are more secure than $DeJoin_2$. For intuitively understanding the effect of BBC in JPEG steganography, an example is provided in Figure 10 to visualize the embedding changes in spatial domain. The sample cover image of size 128×128 pixels, containing smooth, edges and textured regions, is cropped from the full-size image "1013.jpg". It is clear that BBC-based schemes do maintain spatial continuity at block boundaries, and $UpDist\text{-}DeJoin_2$ preserves a wider range of continuity so that it can slightly outperform $DeJoin_4$. Hence, we select $UpDist\text{-}DeJoin_2$ for the following experiments.

5.3 Comparison with State-of-the-art Additive Distortion Functions

Since UERD and J-UNIWARD have become the mainstream methods in JPEG steganography, we compare the steganographic securities of UERD-$UpDist\text{-}DeJoin_2$ and J-UNI-$UpDist\text{-}DeJoin_2$ with that of the initial distortion UERD and J-UNIWARD against DCTR-8,000D and GFR-17,000D

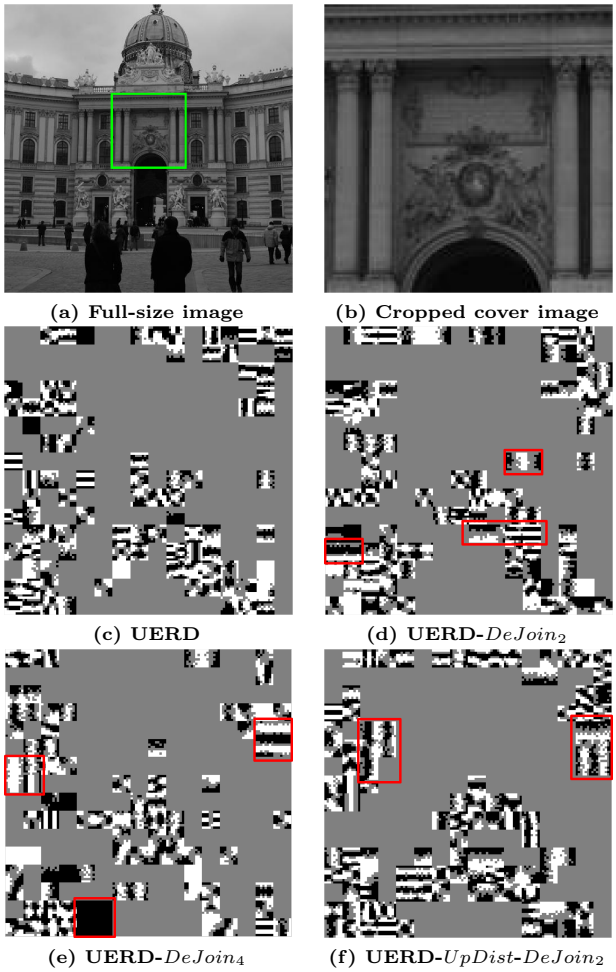
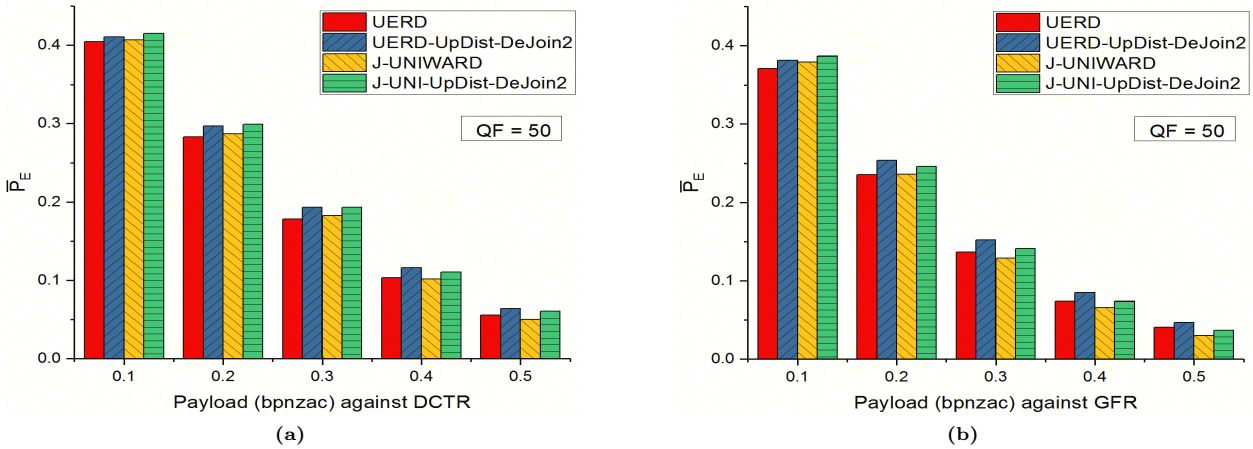


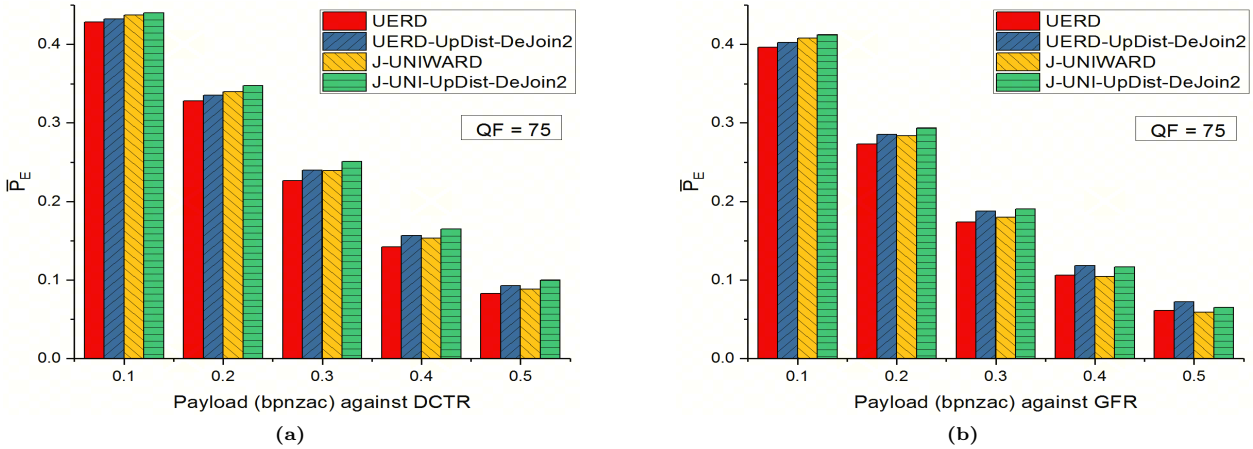
Figure 10: The spatial changes (c)-(f) of the cropped cover image (b) with payload 0.5 bpnzac, QF=75, using UERD and the corresponding BBC-based schemes respectively, where white represents increase (≥ 1) and dark represents decrease (≤ -1).

under $QF = 50, 75, 90$. For $QF = 75$ in Figure 12, the BBC-based schemes perform better than the initial distortions in all cases, and improvements are larger than 1% at high payloads ($\geq$0.3bpnzac) both for UERD and J-UNIWARD. As shown in Figure 11, improvements are more outstanding for $QF = 50$, even with small payloads. However for $QF = 90$ in Figure 13, promotions from BBC are relatively mild. We attribute this phenomenon to the degree of blocking artifact caused by modifications on inter-block neighbors with different QFs. It is clear in (9) that with the QF becoming smaller and equivalently the quantization matrix $\mathbf{Q}$ becoming larger, modifying a DCT coefficient is creating a larger impact on spatial block, so unreasonable joint modifications would inevitably lead to more severe blocking artifacts. Although the magnitudes of spatial changes are neglected in this paper,



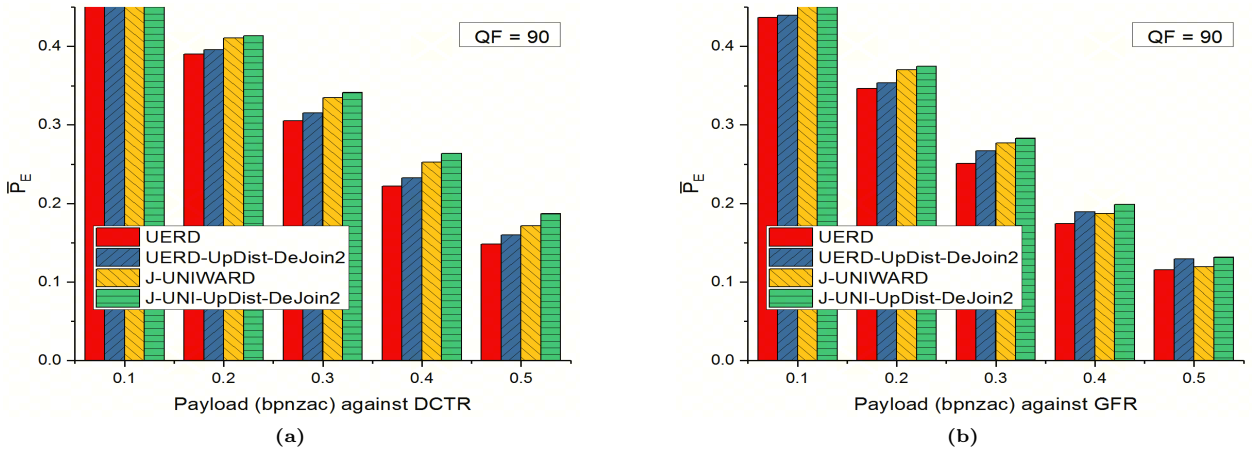
(a) (b)

Figure 11: Detection errors for UERD, J-UNIWARD and their corresponding BBC-based $UpDist$-$DeJoin_2$ against two steganalysis features under QF=50. (a) DCTR. (b) GFR.



(a) (b)

Figure 12: Detection errors for UERD, J-UNIWARD and their corresponding BBC-based $UpDist$-$DeJoin_2$ against two steganalysis features under QF=75. (a) DCTR. (b) GFR.



(a) (b)

Figure 13: Detection errors for UERD, J-UNIWARD and their corresponding BBC-based $UpDist$-$DeJoin_2$ against two steganalysis features under QF=90. (a) DCTR. (b) GFR.

Table 3: Detection errors at 0.3 bpnzac for three novel additive schemes (IUERD, HDS and RBV) and their corresponding BBC-based schemes on BOSSbase1.01 using the FLD ensemble classifier with two feature sets under quality factor 75.

Embedding Method	DCTR	GFR
IUERD	.2287 ± .0019	.2011 ± .0020
IUERD-$UpDist$-$DeJoin_2$	.2447 ± .0017	.2198 ± .0025
HDS	.2292 ± .0022	.1841 ± .0024
HDS-$UpDist$-$DeJoin_2$	.2420 ± .0018	.1993 ± .0023
RBV	.2420 ± .0019	.1996 ± .0027
RBV-$UpDist$-$DeJoin_2$	.2541 ± .0014	.2171 ± .0021

the principle of BBC is in reality designed to utilize blocking artifacts, so the BBC-based schemes can work better with small QFs and correspondingly the highly compressed JPEG images. Table 4 reports the security performances of the involved schemes against two steganalysis features under quality factor 50, 75 and 90.

We also test three novel additive distortions IUERD, HDS, RBV and their corresponding BBC-based $UpDist$-$DeJoin_2$ at 0.3bpnzac under $QF = 75$, to verify whether the BBC principle can be generalizable to other additive schemes. It confirms in Table 3 that deploying BBC into non-additive scheme is beneficial to steganographic security. To date, RBV-$UpDist$-$DeJoin_2$ achieves the state-of-the-art security performance for JPEG steganography when resisting DCTR-8,000D, while IUERD-$UpDist$-$DeJoin_2$ and RBV-$UpDist$-$DeJoin_2$ receive the same best security against GFR-17,000D, of which the improvements to the initial distortions are apparent by about 1.8%.

6 CONCLUSION

Nowadays, non-additive distortion schemes with the principle of SMD have been proved to be tremendously beneficial for spatial image steganography. However, experimental results show that SMD could not be directly applied to JPEG steganography, and thus finding some principles appropriate for JPEG steganography has become an essential and interesting research problem.

In this paper, we introduce a principle of Block Boundary Continuity (BBC) for defining JPEG joint distortion, which tactfully and initiatively inspects the combined embedding changes on adjacent blocks from the spatial domain. According to BBC, the changed directions of some inter-block neighbors may be the same while others should be the opposite, which is related to the DCT mode and the adjacent direction of inter-block neighbors (horizontal or vertical). The principle aims at preserving spatial continuity at block boundaries through restraining blocking artifacts caused by joint modifications, so the neighboring relativity in natural images is maintained and thus modifications in DCT domain would be more secure. Experiments demonstrate that when configured into the model of $DeJoin$, BBC does help improve state-of-the-art JPEG additive schemes in terms of relatively large embedding payloads against modern JPEG steganalyzers.

Since the magnitude of spatial change is neglected in this paper, we only exploit the spatial continuity at block boundaries in a simplified and rough way. How to precisely measure the degree of continuity or blocking artifact to enhance the security of JPEG steganography will be further explored in the future.

ACKNOWLEDGMENTS

This work was supported in part by the Natural Science Foundation of China under Grant U1636201 and 61572452. The authors would like to thank DDE Laboratory of SUNY Binghamton for sharing the source code of steganography, steganalysis and ensemble classifier on the webpage (http://dde.binghamton.edu/download/).

REFERENCES

[1] Patrick Bas, Tomáš Filler, and Tomáš Pevnỳ. 2011. Break Our Steganographic System: The Ins and Outs of Organizing BOSS. In *Information Hiding*. Springer, 59–70.
[2] Tomáš Denemark and Jessica Fridrich. 2015. Improving steganographic security by synchronizing the selection channel. In *Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security*. ACM, 5–14.
[3] Tomáš Filler and Jessica Fridrich. 2010. Gibbs construction in steganography. *IEEE Transactions on Information Forensics and Security* 5, 4 (2010), 705–720.
[4] Tomáš Filler, Jan Judas, and Jessica Fridrich. 2011. Minimizing additive distortion in steganography using syndrome-trellis codes. *IEEE Transactions on Information Forensics and Security* 6, 3 (2011), 920–935.
[5] Jessica Fridrich. 2009. *Steganography in digital media: principles, algorithms, and applications*. Cambridge University Press.
[6] Jessica Fridrich and Tomas Filler. 2007. Practical methods for minimizing embedding impact in steganography. In *Electronic Imaging 2007*. International Society for Optics and Photonics, 650502–650502.
[7] Linjie Guo, Jiangqun Ni, and Yun Qing Shi. 2014. Uniform embedding for efficient JPEG steganography. *IEEE transactions on Information Forensics and Security* 9, 5 (2014), 814–825.
[8] Linjie Guo, Jiangqun Ni, Wenkang Su, Chengpei Tang, and Yun-Qing Shi. 2015. Using statistical image model for JPEG steganography: uniform embedding revisited. *IEEE Transactions on Information Forensics and Security* 10, 12 (2015), 2669–2680.
[9] Vojtěch Holub and Jessica Fridrich. 2015. Low-complexity features for JPEG steganalysis using undecimated DCT. *IEEE Transactions on Information Forensics and Security* 10, 2 (2015), 219–228.
[10] Vojtěch Holub, Jessica Fridrich, and Tomáš Denemark. 2014. Universal distortion function for steganography in an arbitrary domain. *EURASIP Journal on Information Security* 2014, 1 (2014), 1.
[11] Andrew D Ker, Patrick Bas, Rainer Böhme, Rémi Cogranne, Scott Craver, Tomáš Filler, Jessica Fridrich, and Tomáš Pevnỳ. 2013. Moving steganography and steganalysis from the laboratory into the real world. In *Proceedings of the first ACM workshop on Information hiding and multimedia security*. ACM, 45–58.
[12] Jan Kodovsky, Jessica Fridrich, and Vojtěch Holub. 2012. Ensemble classifiers for steganalysis of digital media. *IEEE Transactions on Information Forensics and Security* 7, 2 (2012), 432–444.
[13] Bin Li, Junhui He, Jiwu Huang, and Yun Qing Shi. 2011. A survey on image steganography and steganalysis. *Journal of Information Hiding and Multimedia Signal Processing* 2, 2 (2011), 142–172.
[14] Bin Li, Ming Wang, Xiaolong Li, Shunquan Tan, and Jiwu Huang. 2015. A strategy of clustering modification directions in spatial image steganography. *IEEE Transactions on Information Forensics and Security* 10, 9 (2015), 1905–1917.

Table 4: Detection errors for UERD, J-UNIWARD and their corresponding BBC-based schemes on BOSSbase1.01 using the FLD ensemble classifier with two feature sets under quality factor 50, 75, 90.

QF	Feature	Embedding Method	0.1 bpnzac	0.2 bpnzac	0.3 bpnzac	0.4 bpnzac	0.5 bpnzac
50	DCTR	UERD	.4048 ± .0026	.2830 ± .0013	.1782 ± .0017	.1036 ± .0010	.0559 ± .0010
		UERD-$UpDist$-$DeJoin_2$	.4104 ± .0024	.2972 ± .0016	.1932 ± .0017	.1162 ± .0011	.0639 ± .0010
		J-UNIWARD	.4069 ± .0019	.2869 ± .0015	.1825 ± .0015	.1019 ± .0014	.0500 ± .0005
		J-UNI-$UpDist$-$DeJoin_2$	.4153 ± .0025	.2992 ± .0025	.1934 ± .0024	.1105 ± .0017	.0607 ± .0011
	GFR	UERD	.3705 ± .0038	.2353 ± .0028	.1369 ± .0015	.0739 ± .0009	.0408 ± .0009
		UERD-$UpDist$-$DeJoin_2$	.3813 ± .0032	.2534 ± .0035	.1524 ± .0018	.0849 ± .0010	.0468 ± .0011
		J-UNIWARD	.3793 ± .0026	.2360 ± .0034	.1287 ± .0014	.0654 ± .0011	.0302 ± .0008
		J-UNI-$UpDist$-$DeJoin_2$	.3868 ± .0026	.2462 ± .0028	.1413 ± .0012	.0739 ± .0007	.0370 ± .0009
75	DCTR	UERD	.4284 ± .0029	.3280 ± .0024	.2263 ± .0021	.1424 ± .0010	.0829 ± .0012
		UERD-$DeJoin_2$	.4327 ± .0017	.3316 ± .0021	.2315 ± .0026	.1476 ± .0015	.0888 ± .0009
		UERD-$DeJoin_4$	.4333 ± .0029	.3359 ± .0015	.2352 ± .0024	.1536 ± .0020	.0897 ± .0016
		UERD-$UpDist$-$DeJoin_2$	.4326 ± .0028	.3351 ± .0038	.2401 ± .0015	.1566 ± .0020	.0927 ± .0016
		J-UNIWARD	.4375 ± .0011	.3399 ± .0023	.2392 ± .0017	.1535 ± .0027	.0883 ± .0014
		J-UNI-$UpDist$-$DeJoin_2$	.4402 ± .0021	.3476 ± .0021	.2511 ± .0016	.1651 ± .0020	.0998 ± .0022
	GFR	UERD	.3962 ± .0031	.2729 ± .0023	.1739 ± .0012	.1059 ± .0011	.0611 ± .0008
		UERD-$UpDist$-$DeJoin_2$	.4024 ± .0027	.2856 ± .0021	.1877 ± .0019	.1183 ± .0016	.0721 ± .0011
		J-UNIWARD	.4081 ± .0024	.2836 ± .0014	.1797 ± .0013	.1043 ± .0015	.0587 ± .0008
		J-UNI-$UpDist$-$DeJoin_2$	.4121 ± .0027	.2935 ± .0024	.1906 ± .0020	.1168 ± .0012	.0651 ± .0009
90	DCTR	UERD	.4611 ± .0025	.3902 ± .0031	.3054 ± .0022	.2219 ± .0022	.1484 ± .0010
		UERD-$UpDist$-$DeJoin_2$	.4634 ± .0022	.3955 ± .0019	.3154 ± .0033	.2327 ± .0028	.1600 ± .0020
		J-UNIWARD	.4728 ± .0024	.4106 ± .0033	.3348 ± .0022	.2524 ± .0021	.1715 ± .0014
		J-UNI-$UpDist$-$DeJoin_2$	.4737 ± .0014	.4135 ± .0016	.3412 ± .0030	.2638 ± .0025	.1869 ± .0015
	GFR	UERD	.4366 ± .0028	.3464 ± .0042	.2507 ± .0015	.1742 ± .0013	.1154 ± .0014
		UERD-$UpDist$-$DeJoin_2$	.4396 ± .0015	.3533 ± .0046	.2668 ± .0031	.1891 ± .0008	.1297 ± .0024
		J-UNIWARD	.4523 ± .0025	.3703 ± .0020	.2770 ± .0031	.1868 ± .0021	.1196 ± .0020
		J-UNI-$UpDist$-$DeJoin_2$	.4536 ± .0023	.3748 ± .0025	.2830 ± .0027	.1990 ± .0017	.1315 ± .0018

[15] Yuanfeng Pan, Jiangqun Ni, and Wenkang Su. 2016. Improved Uniform Embedding for Efficient JPEG Steganography. In *International Conference on Cloud Computing and Security*. Springer, 125–133.

[16] Tomás Pevnỳ and Jessica J Fridrich. 2008. Benchmarking for Steganography.. In *Information Hiding*, Vol. 5284. Springer, 251–267.

[17] Xiaofeng Song, Fenlin Liu, Chunfang Yang, Xiangyang Luo, and Yi Zhang. 2015. Steganalysis of adaptive JPEG steganography using 2D Gabor filters. In *Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security*. ACM, 15–23.

[18] Zichi Wang, Xinpeng Zhang, and Zhaoxia Yin. 2016. Hybrid distortion function for JPEG steganography. *Journal of Electronic Imaging* 25, 5 (2016), 050501–050501.

[19] Qingde Wei, Zhaoxia Yin, Zichi Wang, and Xinpeng Zhang. 2017. Distortion function based on residual blocks for JPEG steganography. *Multimedia Tools and Applications* (2017), 1–14.

[20] Weiming Zhang, Zhuo Zhang, Lili Zhang, Hanyi Li, and Nenghai Yu. 2017. Decomposing joint distortion for adaptive steganography. *IEEE Transactions on Circuits and Systems for Video Technology* 27, 10 (2017), 2274–2280.