Forensic Science International 277 (2017) 133–147

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Improved JPEG anti-forensics with better image visual quality and forensic undetectability

Gurinder Singh*, Kulbir Singh

Department of Electronics and Communication Engineering, Thapar University, Patiala, Punjab, India

ARTICLE INFO

Article history:
Received 23 September 2016
Received in revised form 1 June 2017
Accepted 1 June 2017
Available online 10 June 2017

Keywords:
Digital image forensics
JPEG anti-forensics
Double JPEG compression
Blocking artifacts

ABSTRACT

There is an immediate need to validate the authenticity of digital images due to the availability of powerful image processing tools that can easily manipulate the digital image information without leaving any traces. The digital image forensics most often employs the tampering detectors based on JPEG compression. Therefore, to evaluate the competency of the JPEG forensic detectors, an anti-forensic technique is required. In this paper, two improved JPEG anti-forensic techniques are proposed to remove the blocking artifacts left by the JPEG compression in both spatial and DCT domain. In the proposed framework, the grainy noise left by the perceptual histogram smoothing in DCT domain can be reduced significantly by applying the proposed de-noising operation. Two types of denoising algorithms are proposed, one is based on the constrained minimization problem of total variation of energy and other on the normalized weighted function. Subsequently, an improved TV based deblocking operation is proposed to eliminate the blocking artifacts in the spatial domain. Then, a decalibration operation is applied to bring the processed image statistics back to its standard position. The experimental results show that the proposed anti-forensic approaches outperform the existing state-of-the-art techniques in achieving enhanced tradeoff between image visual quality and forensic undetectability, but with high computational cost.

1. Introduction

The development of multimedia services increases the appearing frequency of doctored images related to the political, advertising and personal attacks. The ease in availability of the photo editing softwares and high quality modern cameras has led to the manipulation of the image information to a great extent. Therefore, maintaining the authenticity of digital images is one of the primary concerns. Digital image forensics provides efficient tools and algorithms which generate the historic information of an image related to the camera or device properties used for shooting. The analysis of this information infers that the considered image is a doctored or not [1]. The objective of the anti-forensic techniques is to fool or mislead the forensics examiners by creating difficulties in forgery detection process. The anti-forensic techniques can further help the examiners to find the flaws in the conventional forensics methods for the enhancement of digital forensics [2].

The majority of digital devices or cameras and various image processing software tools generally utilize the JPEG compression format. Therefore, for the image forensics purposes the research has been vastly concentrated on evaluating whether an image has been JPEG compressed (or doubly JPEG compressed). On the other hand, employing the anti-forensic techniques to hide the traces left in an image during JPEG compression becomes an interesting area for the researchers.

Fan and De Queiroz in Ref. [3] presented two forensic detectors, one to identify the JPEG artifacts in DCT domain and other in the spatial domain. The technique based on the JPEG error analysis mentioned in Ref. [4] revealed the historic information of JPEG compression. Stamm et al. [5] suggested a JPEG anti-forensic technique to hide the history of JPEG compression in digital images. The dithering technique based on the Laplacian model for AC coefficients is proposed in Ref. [5] to perform DCT histogram smoothing. It is generally focused on filling the gaps in the comb-like distribution of DCT coefficients in each subband. The proposed dithering operation [5] was successfully able to fool the detector proposed in Ref. [3], which investigates the DCT domain artifacts. Later on, Stamm et al. in Ref. [6] proposed a deblocking operation based on median filtering to tackle the JPEG blocking artifacts

* Corresponding author.
E-mail addresses: gurinder.singh@thapar.edu (G. Singh), ksingh@thapar.edu (K. Singh).

http://dx.doi.org/10.1016/j.forsciint.2017.06.003
0379-0738/

detector [3] in the spatial domain. The detectors proposed in Refs. [7,8] detect the footprints left by JPEG anti-forensic technique suggested in Ref. [5]. The method proposed in Ref. [5] reduces the visual quality of the image. Thus in Ref. [9], a perceptual anti-forensic technique was opined to improve the image quality.

Image modification detection can be performed by using the DCT coefficients analysis [10] and color information [11]. The detector based on the machine learning [12] utilizes the steganalysis feature vector of [13], which was successfully able to detect the JPEG forgery introduced by Stamm et al. in Ref. [5]. A JPEG anti-forensic technique was proposed in Ref. [14] to fool the existing JPEG forensic detectors with high visual quality. The technique opined in Ref. [14] minimizes the variational energy to eradicate the blocking artifacts in the spatial domain. Moreover, an anti-forensic technique was suggested in Ref. [15] to remove the traces left during the double JPEG compression. Some relative work of JPEG forensic based on the machine learning steganalysis method was also proposed in Ref. [16]. Pasquini and Boato proposed a novel anti-forensic technique [17] in which the original distribution of first significant digits (FSD) of the DCT coefficients are recovered. This anti-forensic technique has strong effect on the reliability of forensic analysis as it conceals the traces of first compression to a great extent. In Ref. [18], a counter-forensic technique is proposed to mask the traces of multiple compressions based on the histograms analysis of quantized DCT coefficients. This counter-forensic scheme can effectively hide the artifacts of double and triple JPEG compression and also provide a better image visual quality. Furthermore, in Ref. [19] a JPEG anti-forensic approach was presented which is based on the dithering, denoising, and deblocking operation. A JPEG anti-forensic technique was presented in Ref. [20], which includes adaptive local dithering model, TV (total variation) based deblocking, and decalibration operation. This technique establishes a good tradeoff between image visual quality and forensic undetectability.

A forger can easily hide the traces that are left behind due to the forgery operation by utilizing the JPEG anti-forensics [6]. The parameters such as quantization table are related to the camera used for shooting can be used as an evidence for the image authentication purposes [21]. The image which is initially processed through the JPEG anti-forensic technique can be compressed again by the forger with different quantization table to counterfeit the original camera used for shooting. A forgery can be created through cut-and-paste or splicing operation by taking a JPEG image compressed with one quantization table and pastes it on to another image that is compressed with another quantization table. Now the resultant image may be saved again using JPEG format. This is the case in which double compression artifacts appear [22].

The double JPEG compression detector is a tool to evaluate whether an image is a doctored or not. It can also localize the part of the image which is tampered or double compressed. Grid shift of second JPEG compression is observed with respect to the first compression while applying DCT. Thus the compression can be classified into aligned double JPEG (A-DJPG) compression and nonaligned double JPEG (NA-DJPG) compression. In Ref. [23], a detector is proposed to detect the A-DJPG compression and it can also be used in steganography. This detector is further used as an attack in the anti-forensic work proposed in Ref. [24]. Bianchi and Piva proposed a detection technique for NA-DJPG compression in Ref. [25]. In this method, integer periodicity map was generated from the DC coefficients. The A-DJPG and NA-DJPG compressions were also regarded in the work [22] where the JPEG artifacts were analyzed by using the block grained image forgery localization method. In Ref. [26], a two-stage technique is proposed to estimate the first quantization matrix from partial double compressed JPEG images. The first stage is dedicated to isolate the double compressed region from partial double compressed image. Subsequently, this isolated region is analyzed in second stage to estimate the first quantization matrix.

The existing techniques mainly concentrated on removing the JPEG blocking artifacts in the spatial domain. In this paper, the approach is modified by considering JPEG artifacts in DCT domain along with the spatial domain. All the existing anti-forensic techniques degrade the visual quality of the processed images. The proposed anti-forensic schemes provide high visual quality of processed image along with the better forensic undetectability. A four-step technique is proposed which includes perceptual DCT histogram smoothing, proposed denoising operation, improved TV-based deblocking, and decalibration operation. In this paper, two denoising algorithms are presented to achieve the high image quality. The robustness of the proposed scheme is verified by its undetectability with the high visual quality of an image against various forensic detection attacks [3,4,7,8,14] in both spatial and the DCT domains. The performance of the proposed schemes is also evaluated by using the machine learning based detectors [12,16]. Furthermore, the double JPEG compression detectors proposed in Refs. [22,23,25] are also considered to analyze the efficacy of the proposed scheme. The forensic undetectability is achieved by disguising both the single and double JPEG compression artifacts with improved visual quality of the processed image.

The structure of paper includes the related work on JPEG compression anti-forensics in Section 2. In Section 3, various JPEG compression detectors are discussed. Section 4 covers the proposed JPEG anti-forensics framework in detail. The experiment results are provided in Section 5 and conclusion is discussed in Section 6.

2. Related work

An efficient anti-forensic technique was proposed in Ref. [5] to conceal the traces left during the JPEG compression. A dithering operation was introduced in Ref. [5] for the DCT histogram smoothing. It removed the traces of JPEG compression by filling the gaps of the comb-like DCT histograms of all the subbands created due to the JPEG compression. Therefore, the resultant image processed through the dithering operation can be recognized as an anti-forensically processed image. In the case of uncompressed image, the AC coefficients of the same subband pursue the Laplacian distribution as follows [5]:

$$P(Y=y)=\frac{\gamma}{2}e^{-\gamma|y|} \tag{1}$$

where Y denotes the DCT coefficient of uncompressed image at the (r,c)th subband, and γ is the Laplacian parameter. The value of r and c lies in the range $\{1,2,3,\ldots,8\}$. When the image is JPEG compressed, the quantized AC coefficients of each subband are distributed according to the discrete Laplacian distribution. For the (r,c)th subband coefficients with quantization step $q_{r,c}$ of the quantization matrix q, the distribution can be modeled as [5]:

$$P(Z=z)=\begin{cases}1-e^{-\frac{\gamma q_{r,c}}{2}} & if z=0\\ e^{-\gamma|z|}\sin\left(\frac{\gamma q_{r,c}}{2}\right) & if z=kq_{r,c}\\ 0 & otherwise\end{cases} \tag{2}$$

where $Z=q_{r,c}.round(Y/q_{r,c})$, and maximum likelihood estimation is used to generate the parameter $(\gamma=\gamma_{mle})$. In the dithering operation proposed in Ref. [5], the comb-like DCT histograms for each subband are produced due to the discreteness of Laplacian distribution. Therefore, to approximately reconstruct the histogram of each subband, noise is added to the AC coefficients of each

subband as follows:

$$D = Z + N \tag{3}$$

Here D is the dithered image coefficient, N is the added noise and the noise distribution for the coefficient Z having zero value at the (r,c)th position which can be modeled as [5]:

$$\mathrm{P}(N=n|Z=0) = \begin{cases} \frac{1}{c_0} e^{-\gamma_{mle}|n|} & if \frac{-q_{r,c}}{2} \le n < \frac{q_{r,c}}{2} \\ 0 & otherwise \end{cases} \tag{4}$$

where $c_0 = 1 - e^{-\gamma_{mle}} q_{r,c}/2$, and the distribution for the non-zero coefficient values is given in Ref. [5] as:

$$\mathrm{P}(N=n|Z=z) = \begin{cases} \frac{1}{c_1} e^{-sgn(z)\gamma_{mle}(n+q/2)} & if \ \frac{-q_{r,c}}{2} \le n < \frac{q_{r,c}}{2} \\ 0 & otherwise \end{cases} \tag{5}$$

where $c_1 = (1/\gamma_{mle})(1 - e^{-\gamma_{mle} q_{r,c}})$.

In Ref. [20], a four step JPEG anti-forensic technique is proposed which includes first round TV-based deblocking, perceptual DCT histogram smoothing, second round TV-based deblocking and decalibration operation. The gaps in the DCT histogram are partially filled by the TV-deblocking operation. So, to completely fill the gaps left during the TV-based deblocking, an adaptive local dithering signal model is presented in Ref. [20] based on the Laplace and uniform distribution. Let X be a given image and $(D_{matrix}X)^l_{r,c}$ represents the DCT coefficient value of (r,c)th subband of l-th 8×8 block of a matrix. Here D_{matrix} represents the DCT block matrix and $l = 1, 2, 3, \ldots, L$. The integer variable L represents the total number of non-overlapping 8×8 blocks of an image. For a quantization bin b, the following constrained weighted least-squares fitting problem can be solved for the parameter γ_b as [20]:

$$\gamma_b = \underset{\gamma_b^- \le \gamma \le \gamma_b^+}{arg\ min} \sum_{k=B^-_{r,c}q_{r,c}-\left[\frac{q_{r,c}}{2}\right]}^{B^+_{r,c}q_{r,c}+\left[\frac{q_{r,c}}{2}\right]} w_k \times \left(H^X_{r,c}(k) - P(Z=k)\right)^2 \tag{6}$$

where $B^+_{r,c} = max\left((q_{block}(D_{matrix}X))^l_{r,c}\right)$ and $B^-_{r,c} = min\left((q_{block}(D_{matrix}X))^l_{r,c}\right)$. The operator $q_{block}(.)$ represents the block DCT coefficient quantization and $H^X_{r,c}(k)$ represents the normalized DCT histogram as follows:

$$H^X_{r,c}(k) = \frac{1}{L}\sum_{l=1}^{L} \delta\left(round\left((D_{matrix}X)^l_{r,c}\right) - k\right), k \in \mathbb{Z} \tag{7}$$

Here, $\delta(x) = 1$ if the value of $x = 0$, otherwise $\delta(x) = 0$. To realize the parameter γ_b, weight can be modeled as:

$$w_k = \left(|round\left(\frac{k}{q_{r,c}}\right) - b| + 1\right)^{-1} \tag{8}$$

Based on the discrete Laplacian distribution, the fitting result can be represented as [20]:

$$\mathrm{P}(Z=z) = \begin{cases} 1 - e^{-\gamma/2} & if\ z = 0 \\ e^{-\gamma|z|}\sinh(\gamma/2) & if\ z \in \mathbb{Z}, z \ne 0 \\ 0 & otherwise \end{cases} \tag{9}$$

If γ_b cannot be found by solving the fitting problem in Eq. (6) then dithering signal model of Ref. [5] can be followed. Moreover, the calibration feature based detector K_L can be fooled by the decalibration operation proposed in Ref. [20]. This decalibration operation is based on the following minimization problem which can be solved by using the subgradient method [20].

$$X^* = \underset{X}{arg\ min} \sum_{k=1}^{28} |var(D_kX) - var(D_kX_{cal})| \tag{10}$$

3. JPEG compression detection

Many detection techniques are used to measure the JPEG blocking artifacts induced due to the compression of an image at different quality factors. The JPEG compression creates two types of compression artifacts. These include quantization artifacts in the DCT domain and blocking artifacts in the spatial domain.

The JPEG blocking artifacts detector [3] is based on the fact that the pixel differences across blocks are similar to those within blocks when no JPEG compression is performed. The blocking signature parameter is the energy difference between histograms and it can be represented as [3]:

$$K_F = \sum_n |H_1(n) - H_2(n)| \tag{11}$$

where $H_1(n)$ and $H_2(n)$ are the normalized histograms of pixel values differences across block boundaries, and within the block respectively.

The Gradient aware blockiness detector [27] is concerned with the awareness of gradients along block borders. Instead of using only two adjacent pixel values, the gradient aware blockiness $\left(B^p_{gr}\right)$ uses four adjacent pixel values across block borders to measure the artifacts. Therefore for smooth areas with a gradient, B^p_{gr} is still zero. A family of measures was built to detect JPEG blocking artifacts in Ref. [27] as:

$$K^p_U = |B^p_{gr}(X) - B^p_{gr}(X_{cal})| \tag{12}$$

where B^p_{gr} represents the gradient aware blockiness, which is described as the normalized l_p norm of the weighted gradient.

The Calibration based detector [8] is established by utilizing the ratio of the variance of high-frequency subbands. The image denoted by X, a cropping of X by 4 pixels both horizontally and vertically provides its calibrated version X_{cal}. After this the variance of 28 high frequency subbands is calculated. The calibrated feature K_L is calculated as follows [8]:

$$K_L = \frac{1}{28}\sum_{k=1}^{28} \left|\frac{var(D_kX) - var(D_kX_{cal})}{var(D_kX)}\right| \tag{13}$$

where $var(.)$ returns the sample variance of the input vector, and D_k is a DCT coefficients matrix of the k-th high-frequency subband.

4. Improved JPEG anti-forensic technique

Based on the anti-forensic technique [20], a new JPEG anti-forensic technique is proposed to achieve a better tradeoff between image visual quality and forensic undetectability. The proposed scheme comprises of four stages which include perceptual DCT histogram smoothing, proposed denoising operation, improved TV-based deblocking, and decalibration operation, as shown in Fig. 1. Also, two types of denoising algorithms are presented to

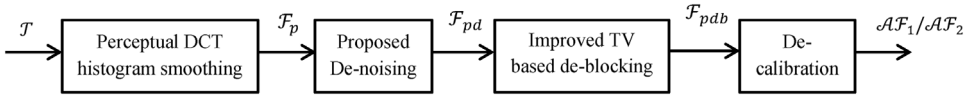

Fig. 1. Sketch of the proposed JPEG anti-forensic technique.

reduce the grainy noise left by the perceptual histogram smoothing; one is based on the constrained minimization problem and the other one on the normalized weighted function. To remove the blocking artifacts left by the JPEG compression in the spatial domain, a new TV-based deblocking operation is proposed considering the combined effect of total variation of energy in horizontal, vertical and diagonal directions (the code of the proposed scheme is available online).[1] The denoising algorithms and deblocking operation are described in the following two sections.

4.1. The proposed denoising schemes

Algorithm 1. Initially, the JPEG compressed image is processed through the perceptual histogram smoothing (adaptive dithering operation) in order to fill the gaps of comb-like DCT histogram of each subband. But, the resultant dithered image contains a significant amount of grainy noise. Therefore, a denoising algorithm is presented based on the total variation [28] to achieve a better image visual quality. Consider the noisy dithered image $F: \Omega \rightarrow V$ of size $T \times T'$, where Ω is bounded open subset of V^2, and U denotes the resultant denoised image. It is desired to solve the following constrained minimization problem based on the total variation of energy:

$$U^* = arg \min_U \sum_{i,j} \{|U_{i,j}| + \lambda \|F_{i,j} - U_{i,j}\|^2\} \tag{14}$$

where $|U_{i,j}|$ denotes the bounded variation semi norm, where $i = 1, 2, 3, \ldots, T$ and $j = 1, 2, 3, \ldots, T'$, and λ represents the adjusting parameter between the two energy terms. The scaling factor λ determines the denoising level. If the value of λ is large, the denoising is minimum, whereas, small values of λ provide maximum denoising. But, the small λ value degrades the image quality. Therefore, to achieve better tradeoff between optimal image quality and forensic undetectability, the value of λ is set at 0.8. The term $\|F_{i,j} - U_{i,j}\|^2$ represents the estimated variance of the unnatural noise added during the perceptual histogram smoothing. By considering $|U_{i,j}| = K(U_{i,j})$ and $\lambda \|F_{i,j} - U_{i,j}\|^2 = J(F_{i,j}, U_{i,j})$, Eq. (14) becomes:

$$U^* = arg \min_U \sum_{i,j} \{K(U_{i,j}) + J(F_{i,j}, U_{i,j})\} \tag{15}$$

where $K(U_{i,j})$ represents the positive convex regularization function and $J(F_{i,j}, U_{i,j})$ is a positive convex fitting function of $U_{i,j}$ for certain value of $F_{i,j}$. This convex nature of $U_{i,j}$ is responsible for the existence of a minimizer U^*. The sub differential of the term $(K(U_{i,j}) + J(F_{i,j}, U_{i,j}))$ which is the combination of two convex functions is zero at U^*, expressed as:

$$\partial_{U_{i,j}} K(U^*) + \partial_{U_{i,j}} J(U^*, F_{i,j}) = 0 \tag{16}$$

The sub differential of the function $J(F_{i,j}, U_{i,j})$ is given as:

$$\partial_{U_{i,j}} J(U_{i,j}, F_{i,j}) = 2\lambda (U_{i,j} - F_{i,j}) \tag{17}$$

Also, the sub differential of the convex function $K(U_{i,j})$ can be written in the form as:

$$\partial_{U_{i,j}} K(U_{i,j}) = -div\left(\frac{\nabla U_{i,j}}{|\nabla U_{i,j}|}\right) \tag{18}$$

where *div* denotes the divergence function and Eq. (16) can be written in the form of Euler–Lagrange differential equation as:

$$-div(\nabla U_{i,j}/|\nabla U_{i,j}|) + 2\lambda (U_{i,j} - F_{i,j}) = 0 \tag{19}$$

Thus, Eq. (19) can be re-written as:

$$U = \sum_{i,j} \left\{ F_{i,j} + \frac{div(\nabla U_{i,j}/|\nabla U_{i,j}|)}{2\lambda} \right\} \tag{20}$$

Hence by solving the constrained minimization problem based on the total variation of energy, the unnatural noise left during the perceptual DCT histogram smoothing is removed to a great extent.

Algorithm 2. To reduce the grainy noise efficiently from the adaptive dithered image, another denoising algorithm is presented exploiting the concept of self-similarity in natural images. It was initially recommended in Ref. [29] that every pixel in an image can be represented as the linear combination of all other pixels, but due to the computational complexity, all the pixels are not considered. Let the noisy dithered image be represented as $F = \{F(i) | i \in I\}$. Then the estimated value of each pixel i can be computed as the weighted average of all the pixels in the image as follows:

$$U[F](i) = \sum_{j \in I} W(i,j) F(j) \tag{21}$$

The weight function $W(i,j)$ depends upon the similarity between the two pixels i and j. The neighboring pixels similar to the pixel under evaluation are allocated with larger weight and the different neighborhood pixels are allocated with small weights.

The function $f_2(x)$ is the estimated function based on the data observations generated for the original function $f_1(x)$. The estimation can be performed by considering a window and the weighted sum of all the pixels falling inside the window will contribute for the value of pixel under process. The estimation of function value $f_2(x_0)$ at point x_0 can be performed by taking the average of all the pixels inside the considered window as shown in orange color.

Basically, two types of windows are considered for the evaluation purposes, one is rectangular and the other is Gaussian window function or inverted parabola as shown in Fig. 2. The Gaussian window provides better estimation as compared to the rectangular window function as shown in Fig. 2. This is due to the fact that in the shifting of rectangular window function to right or left, the estimation shifts up and down in discrete manner. Here the allocated weight is exactly either zero or one, which is the reason behind the rough estimation. On the other hand, Gaussian window function provides zero weight to the points which are too far away and then gives a weight that increases or decreases, thus, providing continuous estimation. There are some regions which have more data points than others, so normalization is needed for the weighted number of data points that are actually inside the prediction window. The normalized weighted function can be represented as:

$$W(i,j) = \frac{e^{\left(-\frac{d(i,j)^2}{2h^2}\right)}}{\sum_j e^{\left(-\frac{d(i,j)^2}{2h^2}\right)}} \tag{22}$$

So, the consequence of this normalization creates the conditions $0 \leq W(i,j) \leq 1$ and $\sum_j W(i,j) = 1$. Where $d(i,j)$ denotes the Euclidean distance between the two points, therefore point out the dissimilarity between the two pixels. The weighting function $W(i,j)$ allocated weights according to this dissimilarity $d(i,j)$. Although, the image is discrete in nature, its continuous

[1] http://bit.ly/guri_antiforensicshttp://bit.ly/guri_antiforensics.

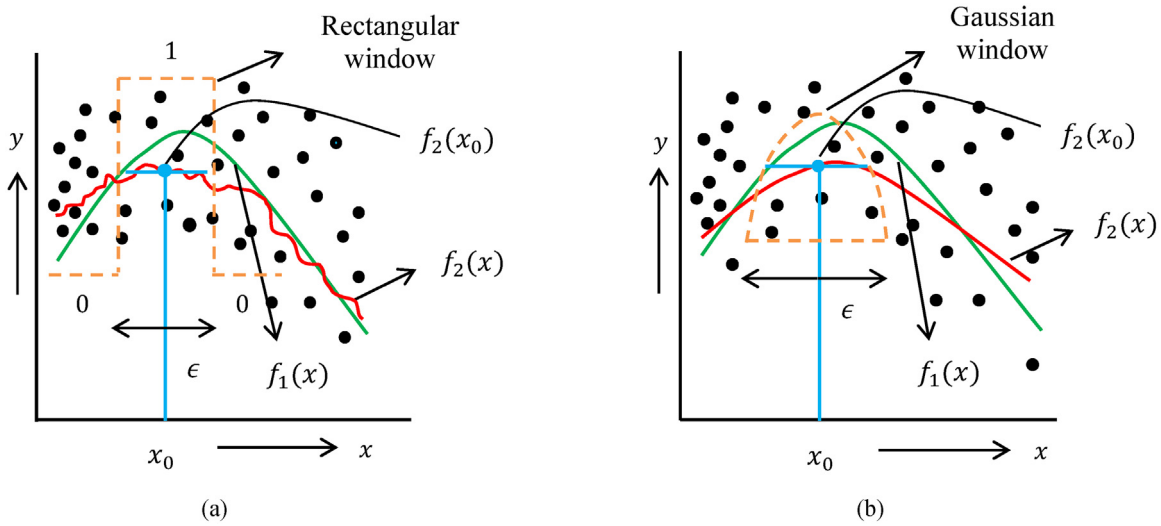

Fig. 2. (a) Estimation of function value $f_2(x)$ at position x_0 using rectangular window, (b) Estimation of function value $f_2(x)$ at position x_0 using Gaussian window.

formulation can be written as:

$$d(i,j) = \int_{R^2} [F(i+t) - F(j+t)]^2 G_a(t) dt \tag{23}$$

where $G_a(t)$ represents the Gaussion window function with standard deviation (a) and i, j and t belong to the 2-dimentional vector R^2. The integral in Eq. (23) depicts that the similarity between the pixels i and j which depends not only on the pixels i and j but also on their surrounding pixels. The Gaussian window function only decays to zero but never becomes quite zero. So, better results can be obtained by using the b-spline function which becomes exactly zero.

4.2. The improved TV-based deblocking operation

The researchers utilize the effective methods to investigate and remove the JPEG blocking artifacts. But these methods focus only on improving the visual quality of the processed image. It is necessary for the anti-forensics to maintain a proper tradeoff between the image visual quality and forensic undetectability. The purpose of this improved deblocking scheme is to minimize the total variation (TV) based energy. The energy term includes TV term and blocking measurement term based on the total variation. By considering an image X of size $T \times T'$, the TV term can be represented as follows [20]:

$$TV(X) = \sum_{1 \le i \le T, 1 \le j \le T'} t_{i,j} \tag{24}$$

For the evaluation of total variation term proposed in Ref. [20], only the horizontal and vertical directions were considered. So, a new definition is proposed in this paper for the total variation term $t_{i,j}$ by considering the combined effect of energy variation along horizontal, vertical and diagonal directions; modeled as:

$$t_{i,j} = \frac{t'_{i,j} + t''_{i,j}}{2} \tag{25}$$

where $t'_{i,j}$ represents the total variation along horizontal and vertical direction and $t''_{i,j}$ denotes the variation along diagonal direction as shown in Fig. 3. This proposed definition of TV-term provides better results than the TV-term proposed in Ref. [20] by providing a better tradeoff between image visual quality and forensic undetectability.

$$t'_{i,j} = \left((X_{i-1,j} + X_{i+1,j} - 2X_{i,j})^2 + (X_{i,j-1} + X_{i,j+1} - 2X_{i,j})^2 \right)^{1/2} \tag{26}$$

$$t''_{i,j} = \left((X_{i-1,j-1} + X_{i+1,j+1} - 2X_{i,j})^2 + (X_{i-1,j+1} + X_{i+1,j-1} - 2X_{i,j})^2 \right)^{1/2} \tag{27}$$

where $X_{i,j}$ represents the value of the pixel at the (i,j)th location. The statistical traces of the JPEG blocking artifacts are removed by using the TV-based blocking measurement term. This term is based on the concept that the energy sum of the pixel value variation along the block borders should be close to the energy sum of pixel value variation within the block. Statistically in the image matrix the energy sum does not depend upon on the starting point of 8×8 DCT block. Therefore, depending upon the pixel positions in the block, all the pixels in the image are divided into two sets as shown in Fig. 4. The shaded pixels location are considered in set A, while the others are put into the set B. The second term is based on

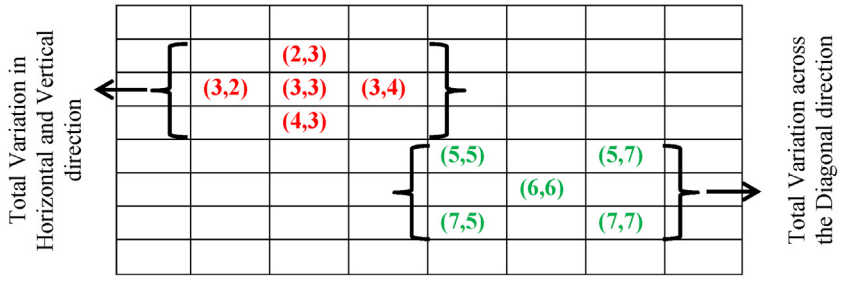

Fig. 3. Combined effect of total variation in horizontal, vertical and diagonal direction.

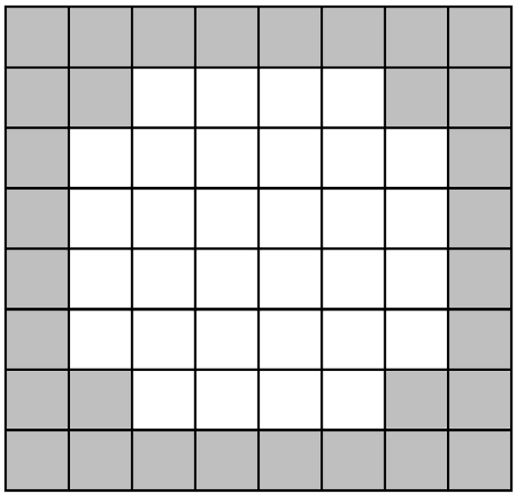

Fig. 4. Classification of pixels into two sets A (shaded) and B (whiten).

this concept and it can be modeled as:

$$E(X) = |\sum_{X_{i,j}\in A} t_{i,j} - \sum_{X_{i,j}\in B} t_{i,j}| \tag{28}$$

To achieve a good quality of the processed image with forensic undetectability, the constraint image space (S) can be defined as:

$$S = \left\{X \in M^{T\times T'} | (D_{matrix}X)^l_{r,c} \in \left[\left(h^l_{r,c} - \beta\right)q_{r,c}, \left(h^l_{r,c} + \beta\right)q_{r,c}\right]\right\} \tag{29}$$

where M denotes the set of integers within the range $[0, 255]$. Here the constraint image space (S) can be controlled by using the parameter β which is a small positive number and $h^l_{r,c} = (q_{block}(D_{matrix}\boldsymbol{I}))^l_{r,c}$ represents the quantized DCT coefficients for the original uncompressed image $\boldsymbol{I}$. This constraint space keeps the DCT coefficients of the processed image within the same quantization bins or in the neighboring bins as those of the JPEG image which is compressed from the original uncompressed image. The TV-based minimization problem can be defined as:

$$X^* = arg\min_{X\in S} T(X) = arg\min_{X\in S} (TV(X) + \alpha E(X)) \tag{30}$$

To balance the two energy terms, α acts as a regularization parameter with positive value. It is easily demonstrated that $T(X)$ is a convex function and it is not differentiable, whereas S is a convex set [30,31]. The projected subgradient method is used to solve the energy minimization problem as:

$$X^{(h+1)} = O_S\left(X^{(h)} - t_{step} \times G\left(X^{(h)}\right)\right) \tag{31}$$

where $X^{(h)}$ represents the processed image at the h-th iteration such that $X^{(0)}$ denotes the given JPEG image, O_S is the projection operator [14], $G(X)$ is the subgradient of $T(X)$, and t_{step} denotes the positive step size. The step size t_{step} and the regularization parameter α can be adjusted to achieve a better tradeoff between the visual quality of the processed image and the restored DCT histogram quality. The optimized value of α is 1.5 and the step size $t_{step} = h$ is set at the h-th iteration. Also, the optimized value of β is set to 0.5, which helps to constrain the processed DCT coefficients to remain in the same quantization bin as its original value. Therefore, the proper adjustment of these parameters provides better image visual quality and forensic undetectability.

When the processed DCT coefficients fall outside the original quantization bins, then these coefficients will be mapped back to the original quantization bins with the help of the projection operator O_S. The strategy used to select the candidate deblocked image is based on the blocking signature measure K_F. The value of K_F has lower standard deviation as compared to another blocking signature K^p_U in the case of uncompressed images. Also, the detection strength of K_F parameter is more than the parameter K^p_U. Therefore, parameter K_F is used for the selection of deblocked image. In the experiment 50 iterations are performed and the image with smallest K_F value is selected as the resultant deblocked image.

Algorithm for the projection operator:

Input:
$\boldsymbol{I}$: Original uncompressed image.
$C_{\mathcal{T}}$: DCT coefficients from JPEG image ($\mathcal{T}$).
q: Quantization matrix.
Output:
Projected Image: Image after the application of projection operator (O_S).
Parameters:
$h^l_{r,c}$: Quantized DCT coefficients obtained from the original uncompressed image ($\boldsymbol{I}$).
Outliers: Represents the coefficients falling outside the original quantization bins.

begin
$[H_{row}, W_{coln}] = size(\boldsymbol{I})$;
$C_{proj} = dct(\boldsymbol{I})$;
$Q_{mat} = repmat(q, H_{row}, W_{coln})$;
$Outliers = round(C_{proj}/Q_{mat}) \sim = h^l_{r,c}$;
i) Noise generation.
$rand_{pos} = (2 * abs(rand(H_{row}, W_{coln}) - 0.5) - 0.5) * Q_{mat};\ for\ [-q_{r,c}/2, q_{r,c}/2)$
$rand_{zero} = rand(H_{row}, W_{coln}) * Q_{mat}/2;\ for\ (-q_{r,c}/2, q_{r,c}/2)$
$rand_{neg} = (-2 * abs(rand(H_{row}, W_{coln}) - 0.5) + 0.5) * Q_{mat};\ for\ (-q_{r,c}/2, q_{r,c}/2]$
ii) Randomly project the outliers to the original quantization bins by adding the random noise to the DCT coefficients.
$C_{proj}(Outliers\ \&\ C_{\mathcal{T}} > 0) = C_{\mathcal{T}}(Outliers\ \&\ C_{\mathcal{T}} > 0) + rand_{pos}(Outliers\ \&\ C_{\mathcal{T}} > 0)$;
$C_{proj}(Outliers\ \&\ C_{\mathcal{T}} == 0) = C_{\mathcal{T}}(Outliers\ \&\ C_{\mathcal{T}} == 0) + rand_{zero}(Outliers\ \&\ C_{\mathcal{T}} = 0)$;
$C_{proj}(Outliers\ \&\ C_{\mathcal{T}} < 0) = C_{\mathcal{T}}(Outliers\ \&\ C_{\mathcal{T}} < 0) + rand_{neg}(Outliers\ \&\ C_{\mathcal{T}} < 0)$;
$Projected\ Image = idct(C_{proj})$;
end

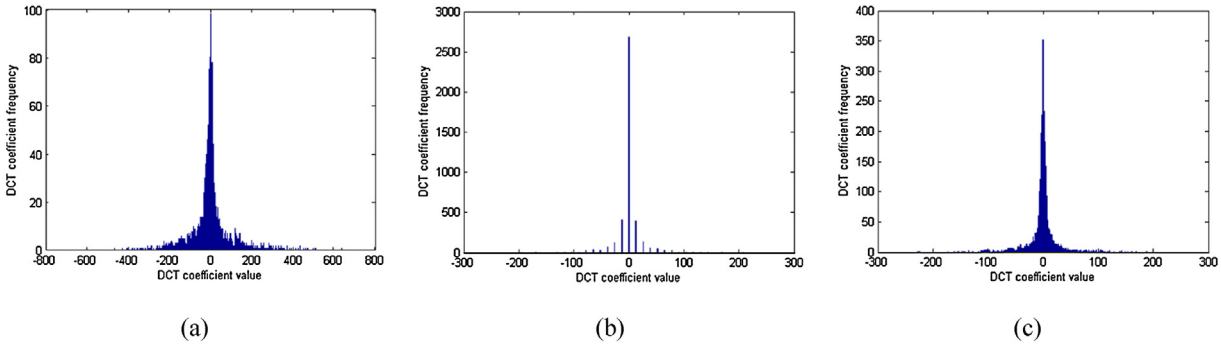

Fig. 5. (a) DCT coefficients histogram of (3, 3) subband of an uncompressed image, (b) Histogram of JPEG compressed the image with quality factor 50, (c) Histogram of the subband (3, 3) after proposed anti-forensic approach.

Fig. 6. (a) represents the compressed Lena image $\mathcal{T}$ with quality factor 50 having PSNR value 35.8085 and SSIM value is 0.9809, (b) denotes JPEG forgery AF_{Fan} with PSNR value is 35.5471 and SSIM value is 0.9753, (c) denotes the proposed JPEG forgery AF_1 with PSNR value is **35.7736** and SSIM value is **0.9770**, (d) denotes the proposed JPEG forgery AF_2 with PSNR value is **35.7938** and SSIM value is **0.9778**.

5. Experiment results

To confirm the adequacy of the proposed JPEG anti-forensic approach, several tests have been conducted by considering standard dataset (UCID) Uncompressed Color Image Database [32]. Originally, the UCID (v2) contains 1338 uncompressed TIFF images with a sure disparity in terms of scene contents. The dataset of single and double compressed images is developed by compressing the UCID dataset images with different quality factors ranging from 50 to 95. The original uncompressed image $\boldsymbol{I}$ is JPEG compressed to yield an image $\mathcal{T}$ and various JPEG anti-forensic techniques are as follows:

- $\mathcal{AF}_{S_q}$, represents the DCT histogram smoothing anti-forensic method [5].
- $\mathcal{AF}_{S_qS_b}$, denotes the anti-forensic method based on dithering and deblocking operation [6].
- $\mathcal{AF}_V$, represents the perceptual anti-forensic dithering method [9].
- $\mathcal{AF}_{S_u}$, denotes anti-forensic technique with SAZ attack [15].

Table 1
Different parameter values attained by $\mathcal{T}$, AF_{Fan}, AF_1 and AF_2, with an uncompressed image as the reference, when tested on different kind of images compressed with quality factor 50.

	PSNR	SSIM	K_L	K_F	K_U^1	K_U^2	K_F^q
Lena $\mathcal{T}$	35.8085	0.9809	67.3822	1.0350	8.4993	349.4780	38.0000
Lena AF_{Fan}	35.5471	0.9753	0.0470	0.0816	0.1061	3.8860	0.0000
Lena AF_1	**35.7736**	**0.9770**	**0.0298**	**0.0756**	**0.0899**	**1.3455**	**0.0000**
Lena AF_2	**35.7938**	**0.9778**	**0.0276**	**0.0456**	**0.0795**	**2.0702**	**0.0000**
Peppers $\mathcal{T}$	34.7507	0.9791	43.5405	1.1595	8.4391	298.8035	41.0000
Peppers AF_{Fan}	34.5238	0.9738	0.0403	0.1947	0.7018	18.8809	1.0000
Peppers AF_1	**34.6758**	**0.9747**	**0.0305**	**0.1036**	**0.2783**	**7.0140**	**0.0000**
Peppers AF_1	**34.6921**	**0.9743**	**0.0363**	**0.0749**	**0.1818**	**5.7105**	**0.0000**
Barbara $\mathcal{T}$	32.8856	0.9821	88.4474	0.8188	10.4921	633.8459	48.0000
Barbara AF_{Fan}	31.6186	0.9709	0.0517	0.2109	0.5684	118.4605	0.0000
Barbara AF_1	**32.4339**	**0.9774**	**0.0375**	**0.1930**	**0.4032**	**1.3633**	**0.0000**
Barbara AF_2	**32.6269**	**0.9766**	**0.0339**	**0.1040**	**0.1035**	**1.1039**	**0.0000**
Baboon $\mathcal{T}$	28.2279	0.9803	44.3233	0.5921	17.3929	1845.2121	60.0000
Baboon AF_{Fan}	27.0367	0.9713	0.0506	0.2184	0.8060	224.1467	0.0000
Baboon AF_1	**28.0614**	**0.9761**	**0.0378**	**0.1782**	**0.6896**	**185.0503**	**0.0000**
Baboon AF_2	**28.8236**	**0.9758**	**0.0323**	**0.1983**	**0.4151**	**124.3153**	**0.0000**

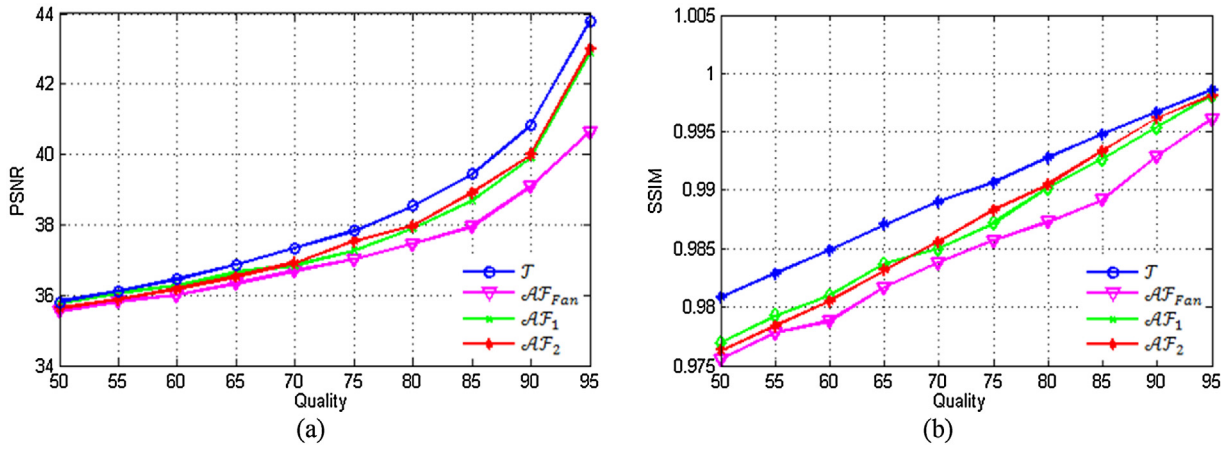

Fig. 7. (a) PSNR values attained by $\mathcal{T}$,$\mathcal{AF}_{Fan}$, AF_1 and AF_2, with an uncompressed image as the reference, when tested on Lena image, (b) SSIM values attained by $\mathcal{T}$, AF_{Fan}, AF_1 and AF_2, with an uncompressed image as the reference, when tested on Lena image.

- $\mathcal{AF}_F$, represents the TV-based anti-forensic method [14].
- $\mathcal{AF}_{Fan}$, demonstrates a four-step JPEG anti-forensic method based on the adaptive dithering model [20].
- $\mathcal{AF}_1$, represents the proposed anti-forensic technique with an improved TV-based deblocking operation and proposed denoising Algorithm 1 based on constrained minimization problem.
- $\mathcal{AF}_2$, represents the proposed anti-forensic technique with an improved TV-based deblocking operation and proposed denoising Algorithm 2 based on the normalized weighted function.

The JPEG forensic detectors used as attacks to validate the adequacy of the proposed anti-forensic technique are as follows:

- K_F, denotes the JPEG compression blocking artifacts detector [3].
- K_F^q, represents the detector based on the quantization table estimation [3].
- K_{Weiqi}, denotes the JPEG identifying detector [4].
- K_{Weiqi}^q, represents the detector based on the quantization step estimation [4].
- K_V, represents the JPEG forensic detector based on total variation [33].
- K_L, JPEG detector based on the calibration feature [8].
- K_U^1 and K_U^2, JPEG blocking artifacts detectors [14].
- K_{Li}^{S100}, represents the detector based on 100-D intra and inter block correlation feature [12,13].
- K_P^{S162}, detector based on 162-D SPAM feature [16].

Table 2
Different parameter values attained after the denoising (Algorithm 1) step F_{pd} of the proposed anti-forensic technique, with an uncompressed image as the reference, when tested on the classical Lena image.

Quality	Parameters						
	PSNR	SSIM	K_L	K_F	K_U^1	K_U^2	K_F^q
50	33.9269	0.9666	0.9758	0.2696	1.0713	41.5550	38.0000
55	34.0321	0.9679	0.9013	0.1950	0.9654	36.4526	38.0000
60	34.1316	0.9687	0.8020	0.2212	0.8387	30.5214	38.0000
65	34.2330	0.9698	0.7124	0.1678	0.7537	25.1501	37.0000
70	34.3444	0.9707	0.6084	0.2101	0.6362	18.3118	36.0000
75	34.4604	0.9715	0.5256	0.1930	0.5723	17.9463	35.0000
80	34.5672	0.9724	0.4139	0.2041	0.4611	13.0580	36.0000
85	34.6862	0.9730	0.2819	0.1744	0.3870	10.7470	33.0000
90	34.8094	0.9736	0.1948	0.1895	0.3044	5.3094	31.0000
95	34.8890	0.9739	0.1411	0.1784	0.2322	5.8981	26.0000

The Fig. 5 represents the histogram of the DCT coefficients for the (3, 3) subband of an uncompressed classical Lena image, a histogram of JPEG compressed Lena image with quality factor 50, and the histogram of the (3, 3) subband after the proposed anti-forensic technique. It is evident from Fig. 5(c) that the periodic gaps left during the JPEG compression are properly filled by the proposed anti-forensic technique without any grainy noise. Moreover, the effectiveness of the proposed anti-forensic techniques $\mathcal{AF}_1$ and $\mathcal{AF}_2$ is evident from Fig. 6. The proposed techniques $\mathcal{AF}_1$ and $\mathcal{AF}_2$ outperforms the existing $\mathcal{AF}_{Fan}$ technique in term of PSNR and SSIM values.

The proposed anti-forensic techniques $\mathcal{AF}_1$ and $\mathcal{AF}_2$ are evaluated based on the PSNR, SSIM, and the various existing forensic detectors including K_L, K_F, K_U^1, K_U^2, K_F^q with different kind of images compressed with quality factor 50. It can be seen from Table 1 that the proposed anti-forensic techniques $\mathcal{AF}_1$ and $\mathcal{AF}_2$ outperform the existing $\mathcal{AF}_{Fan}$ technique by providing smaller values for all the forensic detector parameters and larger PSNR and SSIM values for all the considered images.

The original uncompressed image is taken as a reference to calculate the PSNR and SSIM values by conducting a test on the classical Lena image. Fig. 7 shows the PSNR and SSIM values of the final image processed through the existing JPEG anti-forensic technique $\mathcal{AF}_{Fan}$ and the proposed anti-forensic techniques $\mathcal{AF}_1$ and $\mathcal{AF}_2$. The results of existing anti-forensic techniques confirm that it is not easy to hide the traces of JPEG compression without serious quality loss. So, a better tradeoff is required between the image visual quality and forensic undetectability. By studying the trend of the curves in Fig. 7, it is evident that the proposed anti-forensic methods $\mathcal{AF}_1$ and $\mathcal{AF}_2$ shows better results as compared to the existing anti-forensic technique $\mathcal{AF}_{Fan}$ in terms of PSNR and SSIM values.

The average PSNR value for the proposed JPEG anti-forensic methods $\mathcal{AF}_1$ and $\mathcal{AF}_2$ is **35.9157** dB and **35.9387** dB respectively, when a large-scale test is conducted on the UCID dataset images. Similarly, the average SSIM value for the proposed JPEG anti-forensic methods $\mathcal{AF}_1$ and $\mathcal{AF}_2$ is **0.9884** and **0.9893** respectively. It demonstrates that the proposed anti-forensic methods $\mathcal{AF}_1$ and $\mathcal{AF}_2$, help to achieve higher image quality with better forensic undetectability when compared to the state-of-the-art anti-forensic methods.

Different parameter values are shown in Tables 2 and 3 which includes PSNR, SSIM, and the forensic detectors corresponding to the different quality factors after the denoising stage F_{pd} of the proposed anti-forensic techniques $\mathcal{AF}_1$ and $\mathcal{AF}_2$. Tables 2 and 3

Table 3
Different parameter values attained after the denoising (Algorithm 2) step F_{pd} of the proposed anti-forensic technique, with an uncompressed image as the reference, when tested on the classical Lena image.

Quality	Parameters						
	PSNR	SSIM	K_L	K_F	K_U^1	K_U^2	K_F^q
50	35.1292	0.9637	0.4854	0.0579	0.4675	34.9369	34.0000
55	35.3078	0.9646	0.4119	0.0428	0.3978	32.2444	34.0000
60	35.4509	0.9651	0.2951	0.0670	0.2267	13.5201	34.0000
65	35.6113	0.9658	0.2686	0.0579	0.1675	4.3651	34.0000
70	35.7422	0.9662	0.2288	0.0443	0.1242	0.7861	33.0000
75	35.9065	0.9668	0.2033	0.0726	0.0702	1.4872	32.0000
80	36.0984	0.9675	0.1774	0.0942	0.0209	5.6279	33.0000
85	36.2536	0.9677	0.1316	0.0842	0.0039	5.7050	32.0000
90	36.3805	0.9679	0.1015	0.0595	0.0143	9.4677	33.0000
95	36.4411	0.9680	0.0982	0.0574	0.0206	4.8654	32.0000

Table 4
KL divergence values difference between the successive result of the second round TV-deblocking operation of anti-forensic technique AF_{Fan} and the third step of AF_1 technique respectively for all 64 DCT subbands.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	−0.0212	0.0505	0.1059	−0.0740	0.0635	0.2403	0.2927	0.4409
2	0.1072	0.0838	0.0716	0.0239	0.0555	0.2218	0.2085	0.3271
3	0.1089	0.0701	0.0638	−0.0238	0.2922	0.3485	0.4604	0.5154
4	0.0227	0.0392	−0.0121	0.0874	0.3361	0.4228	0.5292	0.4643
5	0.1582	0.1229	0.2998	0.3430	0.4140	0.5910	0.6108	0.6019
6	0.3177	0.2895	−0.3932	0.4883	0.5076	0.6550	0.6582	0.6515
7	0.4628	0.3911	0.4560	0.6762	0.5745	0.6832	0.7387	0.5628
8	0.7031	−0.5795	0.6481	0.6926	0.7137	0.6261	0.5982	0.2761

Table 5
KL divergence values difference between the after the second round TV-deblocking operation of anti-forensic technique AF_{Fan} and after the third step of AF_2 respectively for all 64 DCT subbands.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	−0.0129	0.0492	0.0736	−0.0627	0.0614	0.2485	0.3199	0.5170
2	0.0703	0.0639	0.0658	0.0345	0.0571	0.2337	0.2374	0.3260
3	0.0986	0.0755	0.0367	−0.0443	0.3068	0.3873	0.4706	0.5321
4	0.0232	0.0218	−0.0342	0.0666	0.3406	0.4388	0.5258	0.4660
5	0.1493	0.1153	0.2686	0.3452	0.4318	0.5846	0.5952	0.5818
6	0.3415	0.3143	−0.3974	0.4850	0.5279	0.6235	0.6496	0.6328
7	0.4762	0.4037	0.4514	0.6734	0.5662	0.6637	0.7325	0.5550
8	0.7765	−0.5692	0.6920	0.6893	0.7118	0.6079	0.5918	0.3248

depicts that the denoising Algorithm 2 provides better results as compared to the Algorithm 1 by providing high PSNR with a small loss in SSIM values, and lower values for the forensic detection parameters. This may happen because of the consideration of all the pixel values of the image for the evaluation of the new pixel value at the particular position in the case of Algorithm 2. The comparative analysis of results is shown in Tables 4 and 5 based on the Kullback Leibler (KL) divergence values difference obtained after the third step of the existing anti-forensic technique $\mathcal{A}F_{Fan}$ as well as the proposed anti-forensic techniques $\mathcal{A}F_1$ and $\mathcal{A}F_2$. This KL divergence measures the difference between two probability distributions. The presented denoising operations reduce the unnatural noises in the image which is processed through the perceptual histogram smoothing. Subsequently, the improved TV-based deblocking operation is applied which removes the compression blocking artifacts to a great extent. Hence, it can be observed from the Tables 4 and 5 that the third stage of proposed anti-forensic techniques $\mathcal{A}F_1$ and $\mathcal{A}F_2$ provide better results as compared to the third stage of the existing $\mathcal{A}F_{Fan}$ technique by providing smaller KL-divergence values for most of the DCT coefficient subbands.

To confirm the efficacy of the proposed TV-based deblocking, the performance analysis is carried out by considering the deblocking operation only. The performance of the improved

Table 6
Comparison between the TV-deblocking operation proposed in $\mathcal{A}F_{Fan}$ and the proposed TV-deblocking operation based on different parameter values, with an uncompressed image as the reference, when tested on the classical Lena image.

	PSNR	SSIM	K_L	K_F	K_U^1	K_U^2	K_F^q
$\mathcal{T}$	35.8085	0.9809	67.3822	1.0350	8.4993	349.4780	38.0000
TV-deblocking [20]	**35.7489**	**0.9765**	0.6704	0.1169	0.4362	21.6955	0.0000
Proposed TV-deblocking	35.6977	0.9756	**0.2607**	**0.0937**	**0.4170**	**0.3579**	**0.0000**

Table 7
KL divergence values difference between the TV-deblocking operation proposed in AF_{Fan} and the proposed TV-deblocking operation for all 64 DCT subbands, when tested on classical Lena image.

	1	2	3	4	5	6	7	8
1	−0.0019	0.0519	0.0404	0.0358	0.0222	0.0629	0.2132	0.4080
2	0.0448	0.0827	0.0024	0.0148	0.0385	0.0994	0.1568	0.2933
3	0.0279	0.0137	0.0030	0.0294	0.0698	0.1647	0.3522	0.5435
4	0.0217	0.0131	0.0213	0.0798	0.1077	0.2371	0.3897	0.4608
5	0.0995	0.0351	0.0893	0.1075	0.1999	0.3479	0.4029	0.5066
6	0.1633	0.1158	0.1992	0.2483	0.3464	0.4633	0.3902	0.4939
7	0.4196	0.2632	0.3098	0.3948	0.3993	0.5070	0.4871	0.3086
8	0.7497	0.4837	0.5743	0.6294	0.6113	0.5274	0.3590	0.1035

TV-based deblocking is analyzed by evaluating various forensic detector parameters along with the existing TV-based deblocking as shown in Table 6. It can be observed that the improved TV-based deblocking provides better results for all the considered forensic detection parameters with a small effect on the image quality. It can be noticed from Table 6 that the value of detection parameter K_U^2 decreases significantly from **21.6955** to **0.3579**. This is generally because of the proposed definition of TV-term considers the combined effect of total variation of energy in horizontal, vertical and diagonal directions.

To further confirm the efficacy of the proposed TV-based deblocking, the KL divergence values (with a reference uncompressed image) of existing TV-based deblocking in $\mathcal{A}F_{Fan}$ and the proposed TV-based deblocking are evaluated. Table 7 depicts the difference between the KL divergence values of existing TV-based deblocking and the proposed TV-based deblocking. It can be inferred from Table 7 that except the DC subband, the proposed TV-deblocking operation outperforms the existing deblocking operation with smaller KL divergence value for all the 63 AC subbands.

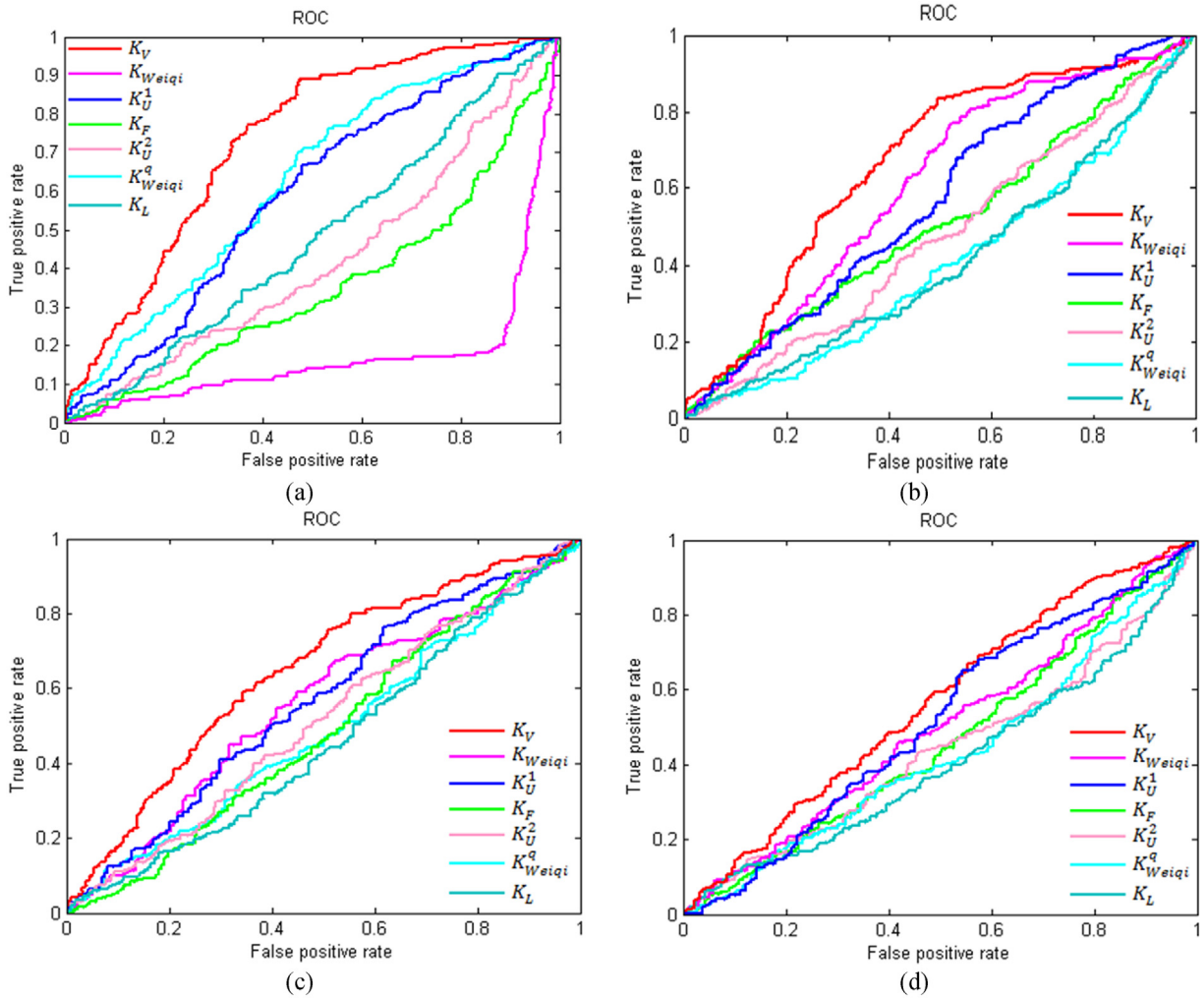

Fig. 8. ROC curves of (a) $\mathcal{A}F_{S_qS_b}$, (b) AF_{Fan}, (c) AF_1 and (d) AF_2 against various forensic detectors. The detectors are fooled better, when the ROC curves approaches to the diagonal (random guess).

Table 8
Average minimum decision error for all the JPEG anti-forensic approaches against various forensic detectors, with an uncompressed image as the reference, when tested on different kind of images.

	K_F	K_{Weiqi}	K^q_{Weiqi}	K_V	K_L	K^1_U	K^2_U
$\mathcal{T}$	0.0061	0	0.0043	0.0136	0.0295	0.0325	0.1687
$\mathcal{AF}_{S_q}$	0.1248	0.4159	0.2197	0.0168	0.0348	0.0723	0.1189
$\mathcal{AF}_{S_qS_b}$	0.3364	**0.4358**	0.3067	0.2176	0.4531	0.4048	0.4775
$\mathcal{AF}_V$	0.0469	0.4078	0.2164	0.0436	0.0235	0.0378	0.1439
$\mathcal{AF}_{S_u}$	0.1538	0.0724	0.0762	**0.4975**	0.0725	0.3457	**0.4924**
$\mathcal{AF}_F$	0.3465	0.3578	0.4697	0.3247	0.4825	0.3723	0.4421
AF_{Fan}	0.3724	0.3726	0.4721	0.3854	0.4982	**0.4128**	0.4627
AF_1	**0.4786**	0.3978	**0.4891**	0.4153	**0.5000**	0.0899	0.4795
AF_2	**0.4956**	0.4297	**0.4974**	0.4359	**0.5000**	0.0795	0.4831

Table 9
Average PSNR and SSIM values for all the JPEG anti-forensic approaches, with an uncompressed image as the reference, when tested on different kinds of image.

	$\mathcal{T}$	$\mathcal{AF}_{S_q}$	$\mathcal{AF}_{S_qS_b}$	$\mathcal{AF}_V$	$\mathcal{AF}_{S_u}$	$\mathcal{AF}_F$	AF_{Fan}	AF_1	AF_1
PSNR	37.1085	33.1438	30.1137	32.1268	30.3694	34.9217	35.2471	**35.9157**	**35.9387**
SSIM	0.9902	0.9713	0.9424	0.9646	0.9627	0.9712	0.9783	**0.9884**	**0.9893**

In steganalysis processes [34,35], the minimum decision error (P_e) is a frequently used measure to evaluate the performance of forgeries against various forensic detectors. Firstly, the receiver operating characteristic (ROC) curve is evaluated for the various forensic detectors. For this SVM classifier is used for the classification purpose by considering the anti-forensically processed images as positive cases, and genuine, uncompressed images as negative cases. The point on the ROC curve with minimum number of wrongly classified images represents the minimum decision error (P_e).

The forensic undetectability of UCID dataset images processed through different anti-forensic techniques is tested by using various JPEG forensic detectors. Most of the DCT coefficients corresponding to the high frequency subbands are quantized to zero for highly compressed JPEG images. Therefore, it is difficult to determine the quantization steps. Consequently, the estimated quantization steps greater than 1 are utilized as the feature value for the forensic detector K^q_{Weiqi}. For a given test image, the output of all the forensic detectors $K_F, K_{Weiqi}, K^q_{Weiqi}, K_V, K_L, K^1_U, K^2_U$ is a one feature value. Based on the threshold value, the detector can classify that whether the considered image is an uncompressed or compressed one. If all the estimated quantization steps corresponding to all the 64 subbands are equal to one or not defined, then the image is classified as uncompressed [3,5]. For testing purpose, 560 images are randomly selected from the UCID dataset, and are compressed with different quality factors. Afterwards, these images are anti-forensically processed to create a database for SVM classifier testing.

The main objective of the proposed scheme is to hide the compression artifacts. Thus, the SVM classifier involves two classes compressed and uncompressed. The true positive rate (TPR) represents the number of anti-forensically processed images that were correctly classified under the compressed image class. Similarly, the false positive rate (FPR) represents the anti-forensically processed images that were wrongly classified under the uncompressed image class. The proposed algorithm is said to be undetectable to the forensic detectors when FPR is greater than TPR. This characteristic can be observed in Fig. 8 as the curve approaches the diagonal (random guess). More the curve is close to the diagonal; higher is the forensic undetectability against the considered detector.

The ROC curve of the JPEG forgery created by the proposed methods is closer to the diagonal (random guess) as compared to the existing JPEG anti-forensic approaches [6,20] against various forensics detectors as shown in Fig. 8. Therefore, the proposed anti-

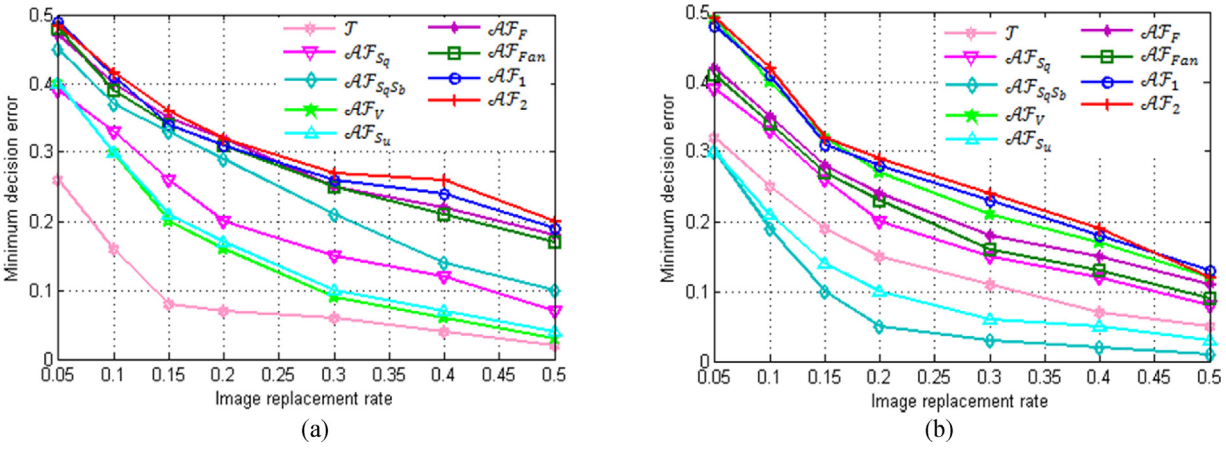

Fig. 9. Minimum decision error as the function of image replacement rate tested against SVM-based forensic detectors (a) K^{S100}_{Li}, (b) K^{S162}_{P}.

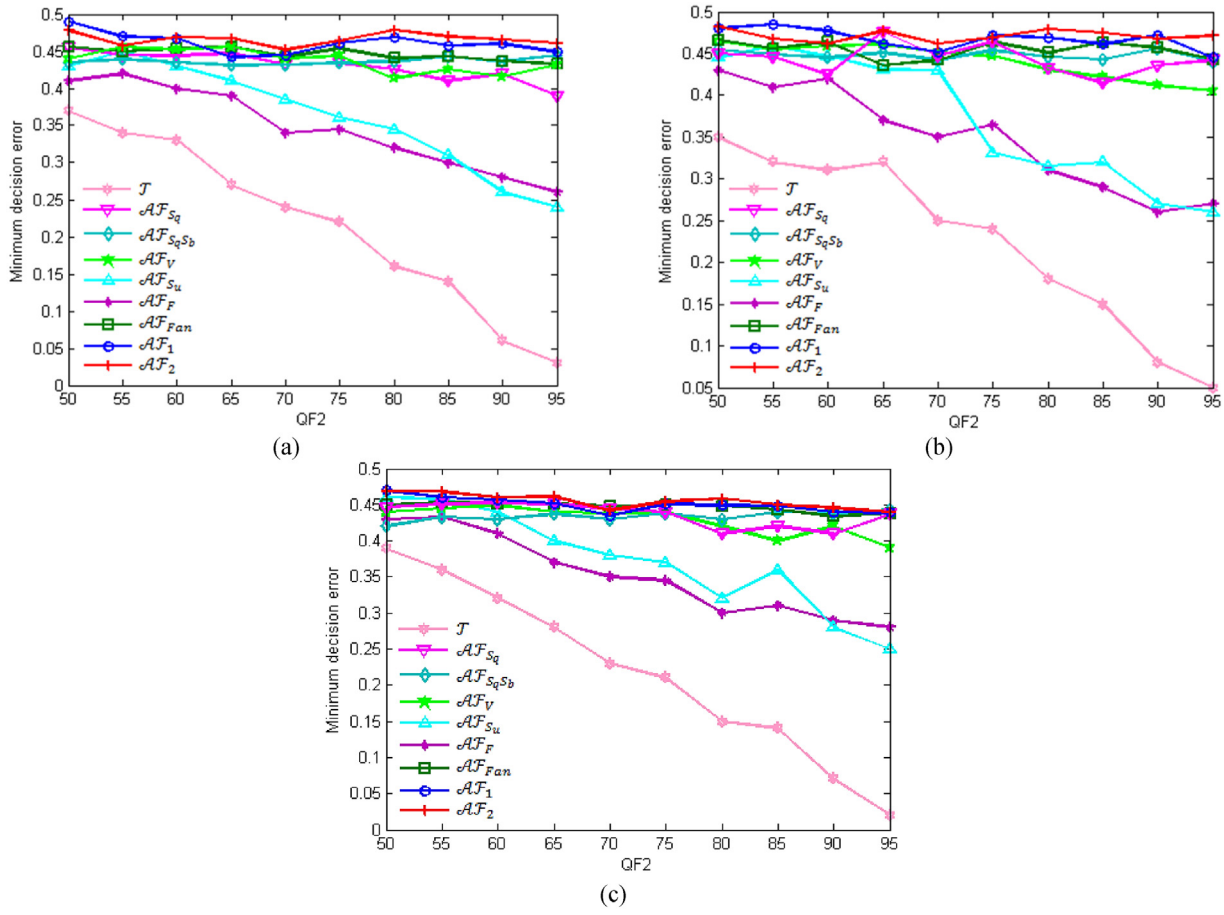

Fig. 10. Minimum decision error as the function of quality factor (*QF*2) for double compressed JPEG images, when tested against various forensic detectors proposed in Refs. [22,23,25].

Table 10
Average PSNR and SSIM values for all the JPEG anti-forensic approaches, with an uncompressed image as the reference, when tested on double JPEG compressed images.

	$\mathcal{T}$	$\mathcal{AF}_{S_q}$	$\mathcal{AF}_{S_qS_b}$	$\mathcal{AF}_V$	$\mathcal{AF}_{S_u}$	$\mathcal{AF}_F$	AF_{Fan}	AF_1	AF_2
PSNR	34.3289	31.6892	29.1573	32.6548	30.3154	33.6214	33.7584	**33.8524**	**33.9254**
SSIM	0.9415	0.8245	0.7958	0.8824	0.8674	0.9148	0.9254	**0.9324**	**0.9345**

forensic techniques are much better in fooling the various forensic detectors as compared to the existing anti-forensic techniques. The proposed anti-forensic algorithms are able to fool the forensic detectors, hence providing higher minimum decision error in comparison to the existing anti-forensic techniques as evident in Table 8.

The motivation behind the development of the proposed anti-forensic technique is that the PSNR and SSIM of the processed image should approach the PSNR and SSIM of the reference

Table 11
Different parameter values attained by the proposed anti-forensic scheme $\mathcal{AF}_1$ by varying the value of scaling factor λ in Eq. (14) exploring the nature of tradeoff between image quality and forensic undetectability based on the different images.

	PSNR	SSIM	K_L	K_F	K_U^1	K_U^2	K_F^q
Lena $\mathcal{T}$	35.8085	0.9809	67.3822	1.0350	8.4993	349.4780	38.0000
Lena (highest quality at $\lambda = 0.98$)	**35.7976**	**0.9782**	0.2505	0.1880	0.6368	4.5646	0.0000
Lena (highest forensic undetectability at $\lambda = 0.91$)	35.7058	0.9766	**0.0102**	**0.0421**	**0.0373**	**0.4496**	**0.0000**
Peppers $\mathcal{T}$	34.7507	0.9791	43.5405	1.1595	8.4391	298.8035	41.0000
Peppers (highest quality at $\lambda = 0.98$)	**34.7015**	**0.9749**	0.3304	0.3854	0.3222	10.7726	0.0000
Peppers (highest forensic undetectability at $\lambda = 0.91$)	34.5857	0.9721	**0.0195**	**0.0954**	**0.1721**	**5.7495**	**0.0000**
Barbara $\mathcal{T}$	32.8856	0.9821	88.4474	0.8188	10.4921	633.8459	48.0000
Barbara (highest quality at $\lambda = 0.99$)	**32.5209**	**0.9776**	0.3300	0.4520	1.7230	14.4283	0.0000
Barbara (highest forensic undetectability at $\lambda = 0.91$)	32.2947	0.9719	**0.0214**	**0.1472**	**0.2357**	**0.9547**	**0.0000**
Baboon $\mathcal{T}$	28.2279	0.9803	44.3233	0.5921	17.3929	1845.2121	60.0000
Baboon (highest quality at $\lambda = 0.97$)	**28.1704**	**0.9765**	0.2705	0.4867	4.7869	196.5867	0.0000
Baboon (highest forensic undetectability at $\lambda = 0.9$)	27.8963	0.9712	**0.0215**	**0.0985**	**0.3145**	**112.2514**	**0.0000**

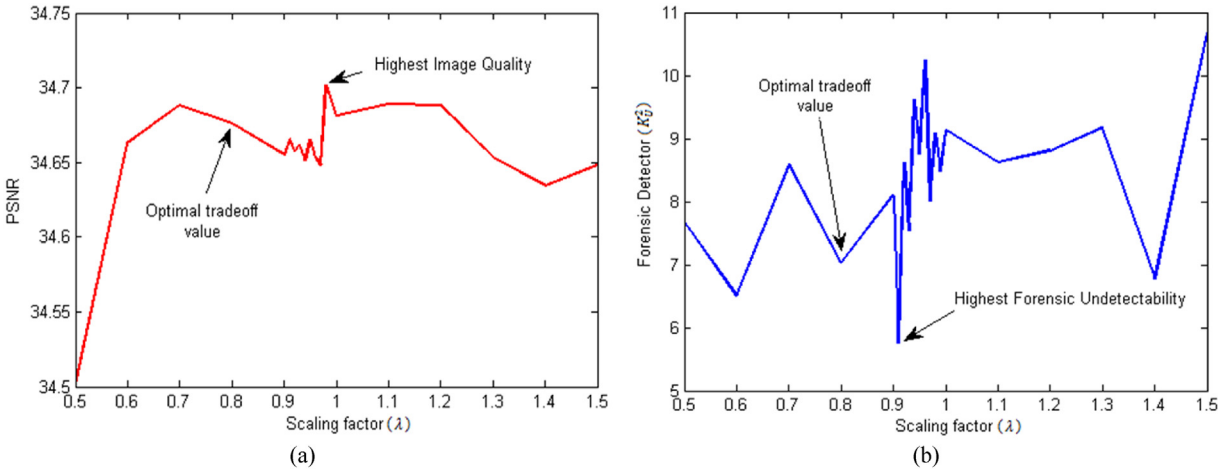

Fig. 11. (a) and (b) represent the PSNR and K_U^2 values respectively obtained by the proposed anti-forensic scheme AF_1 considering the various values of scaling parameter (λ) in Eq. (14), when tested on the Peppers test image.

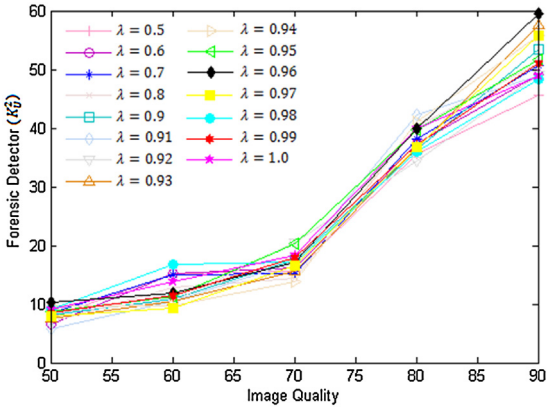

Fig. 12. Represents the forensic parameter values $\left(K_U^2\right)$ obtained by the proposed anti-forensic scheme AF_1 based on the different values of scaling parameter (λ) considering the images obtained by compressing the same Peppers test image with different quality factors ranging from 50 to 90.

uncompressed image. Table 9 lists the average PSNR and SSIM values for all the anti-forensic techniques based on various images by taking the original uncompressed image as a reference. Table 9 demonstrates that the proposed anti-forensic methods AF_1 and AF_2 provides better PSNR and SSIM values.

Table 8 depicts that the forensic detectors can detect the JPEG forgery created by the anti-forensic technique $\mathcal{A}F_{S_q}$. The anti-forensic technique $\mathcal{A}F_V$ based on the perceptual dithering scheme achieves a high SSIM value but its forensic undetectability is comparable to the $\mathcal{A}F_{S_q}$ technique. The JPEG forgery $\mathcal{A}F_{S_qS_b}$ uses the median filtering to improve the forensic undetectability against various detectors but with a loss of 6.9 dB in PSNR value.

The proposed JPEG anti-forensic methods AF_1 and AF_2 have the capacity to fool even the advanced forensic detector K_V. This happens because of the minimization of proposed TV-term in Eq. (25) which is responsible for the suppression of unnatural noises. Moreover, the decalibration operation is applied to defeat the calibration-based detector K_L. When considering the JPEG forgeries created by the proposed anti-forensic methods, 96.72% of the images can be classified as never JPEG compressed. So, the proposed anti-forensic methods outperform the existing anti-forensic techniques in terms of image quality and forensic undetectability.

The SVM based forensic detectors K_{Li}^{S100} and K_P^{S162} which are widely used in steganalysis, exhibit high detection accuracy and provide minimum decision error, which is less than 0.1. If these SVM-based detectors are trained with the existing JPEG anti-forensic images, then most of the images are often detectable. To conceal the JPEG compression artifacts, large number of image DCT coefficients is modified. In image steganalysis, this DCT coefficients modification provides high modification rate (bits per pixel).

To further analyze the performance of the proposed anti-forensic approaches, JPEG forgeries are now created by using the experimental setup of steganography work [35]. The central part of a given uncompressed image is substituted by the JPEG anti-forensic image. In the substitution process, the replacement rate varies from 0.05 to 0.5. Both the processed images and uncompressed images are used for forensic testing. The Fig. 9 shows the minimum decision error as the function of image replacement rate. The proposed forgery techniques outperform the existing anti-forensic techniques by providing higher minimum decision error

Table 12
Average time elapsed (s) to create different types of JPEG forgeries considering the images of different resolution based on the different quality factors.

	$\mathcal{A}F_{S_q}$	$\mathcal{A}F_{S_qS_b}$	$\mathcal{A}F_V$	$\mathcal{A}F_{S_u}$	$\mathcal{A}F_F$	AF_{Fan}	AF_1	AF_2
Cameraman[2] (256 × 256)	0.0398	0.0512	0.2250	0.2458	4.5024	58.9785	**210.5872**	**255.3478**
Table [32] (512 × 384)	0.0750	0.1252	0.5625	0.6257	14.2034	145.5485	**629.1456**	**766.7712**
Boat[2] (512 × 512)	0.1583	0.2473	0.9247	1.0245	18.2456	232.2475	**840.2142**	**1020.2546**
Cat[2] (490 × 733)	0.1510	0.2603	1.0112	1.2315	23.1289	305.2540	**1190.2581**	**1448.2149**
Building[3] (768 × 512)	0.1925	0.3125	1.2349	1.4209	24.3254	350.0124	**1260.3258**	**1575.2321**

when tested against the forensic detectors K_{Li}^{S100} and K_{P}^{S162} as shown in Fig. 9.

To further confirm the performance of the proposed anti-forensic techniques, the test is then conducted on the double JPEG compressed images by considering the double compression artifacts detectors [22,23,25]. It is evident from Fig. 10 that the proposed anti-forensic techniques outperform the existing techniques by providing the higher minimum decision error, which approaches 0.5 for all the considered quality factors when tested against the forensic detectors [22,23,25]. Moreover, the proposed anti-forensic approaches provide higher image visual quality for double compressed images as compared to the existing techniques as shown in Table 10.

The further analysis has been carried out to investigate the nature of the tradeoff between the image quality and forensic undetectability. In this analysis, highest image quality and forensic undetectability is achieved by varying the value of scaling factor (λ) in Eq. (14). It can be inferred from Table 11 that highest image quality is achieved at the cost of decrease in forensic undetectability. On the other hand, high forensic undetectability is attained with the loss in image quality.

Moreover, it can also be perceived from Fig. 11(a) and (b), that highest PSNR of 34.7015 is obtained at the value of scaling factor $\lambda = 0.98$ but with small forensic undetectability, whereas, highest forensic undetectability in terms of forensic parameter value $K_U^2 = 5.7495$ is obtained at the value of scaling factor $\lambda = 0.91$ but with smaller PSNR value, when tested on the peppers test image. The optimal tradeoff obtained by the proposed scheme AF_1, is around the scaling factor value of 0.8, observed from Fig. 11(a) and (b). The values of PSNR and K_U^2 obtained at 0.8 are 34.6758 and 7.0140 respectively. Note that smaller the value of forensic detector parameter K_U^2, more is the forensic undetectability. The forensic undetectability is considered in terms of parameter K_U^2 for the evaluation of the results in Figs. 11 and 12. Similarly, evaluation can also be carried out by considering the other discussed forensic detectors. Fig. 12 also demonstrates the nature of this tradeoff by considering the forensic parameter values $\left(K_U^2\right)$. This is performed on the images obtained by compressing the same Peppers test image with different quality factors ranging from 50 to 90. It can be understood from Fig. 12 that forensic undetectability decreases with the increase in the image quality based on the various values of scaling factor (λ). Moreover, the forensic undetectability decreases significantly after the quality factor 70 as shown in Fig. 12. The proposed JPEG anti-forensic techniques help to achieve the better tradeoff between image quality and forensic undetectability.

In this paper, the anti-forensic techniques are created by using MATLAB R2016a software on a PC with 2.13 GHz CPU and 3 GB RAM. The average execution time required to create different JPEG forgeries by considering the images of different resolutions based on the different quality factors is shown in Table 12. The execution time depends on the number of pixels in an image. It can be observed from Table 12 that the time elapsed to create a JPEG forgery increases with the increase in the number of pixels in an image. The execution time per pixel for AF_1 and AF_2 is in the range [0.0031–0.0033 s/pixel] and [0.0038–0.0040 s/pixel] respectively. The proposed anti-forensic techniques AF_1 and AF_2 require around 14 and 17 min respectively to create the JPEG forgery of size 512×512. The bottleneck of the computation cost lies in the perceptual DCT histogram smoothing. The proposed anti-forensic techniques are computationally intensive as compared to the state-of-the-art JPEG anti-forensic approaches. However, the proposed methods of anti-forensics provide better tradeoff between image visual quality and forensic undetectability. Moreover, the forger does not need to create a large number of forgeries. Therefore, it is acceptable to create a JPEG forgery in around 14 min.

6. Conclusions

In this paper, a JPEG anti-forensic technique is proposed to fool the existing forensic detectors by hiding the compression artifacts in both spatial and DCT domain. The suggested denoising algorithms improve the image quality by removing the unnatural noise left during the perceptual histogram smoothing. Furthermore, the proposed TV-based deblocking operation helps to achieve better forensic undetectability by reducing the blocking artifacts. The evaluation is performed on the UCID dataset. A better tradeoff is obtained between image visual quality and forensic undetectability by the proposed JPEG anti-forensic techniques as compared to the state-of-the-art techniques. Though the proposed technique is better than the previous state-of-the-art anti-forensic techniques, but in terms of computational complexity, it takes more execution time to create a JPEG forgery. The further work can be devoted to make a universal framework for the anti-forensics by achieving a better histogram restoration in DCT domain.

Acknowledgements

The authors would like to thank the anonymous reviewers for their valuable comments which have helped to improve the quality of the paper. This work was supported by the Department of Electronics and Information Technology, Ministry of Communications and Information Technology, Government of India (grant PhD-MLA/4(33)/2015-16/01).

References

[1] H. Farid, A survey of image forgery detection, IEEE Signal Process. Mag. 2 (26) (2009) 16–25.
[2] R. Böhme, M. Kirchner, Counter-forensics: attacking image forensics, in: H.T. Sencar, N. Memon (Eds.), Digital Image Forensics, Springer-Verlag, New York, 2013, pp. 327–366.
[3] Z. Fan, R.L. De Queiroz, Identification of bitmap compression history: JPEG detection and quantizer estimation, IEEE Trans. Image Process. 12 (2) (2003) 230–235.
[4] W. Luo, J. Huang, G. Qiu, JPEG error analysis and its applications to digital image forensics, IEEE Trans. Inf. Forensics Secur. 5 (3) (2010) 480–491.
[5] M. Stamm, S. Tjoa, W.S. Lin, K.J.R. Liu, Anti-forensics of JPEG compression, Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing (2010) 1694–1697.
[6] M. Stamm, S. Tjoa, W.S. Lin, K.J.R. Liu, Undetectable image tampering through JPEG compression anti-forensics, Proceedings of 17th IEEE International Conference on Image Processing (2010) 2109–2112.
[7] G. Valenzise, V. Nobile, M. Tagliasacchi, S. Tubaro, Countering JPEG anti-forensics, Proceedings of 18th IEEE International Conference on Image Processing (2011) 1949–1952.
[8] S. Lai, R. Böhme, Countering counter-forensics: the case of JPEG compression, Proceedings of IEEE International Conference on Information Hiding (2011) 285–298.
[9] G. Valenzise, M. Tagliasacchi, S. Tubaro, The cost of JPEG compression anti-forensics, Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing (2011) 1884–1887.
[10] Liyang Yu, Qi Han, Xiamu Niu, S.M. Yiu, Junbin Fang, Ye Zhang, An improved parameter estimation scheme for image modification detection based on DCT coefficient analysis, Forensic Sci. Int. 259 (2016) 200–209.
[11] Haoyu Zhou, Yue Shen, Xinghui Zhu, Bo Liu, Zigang Fu, Na Fan, Digital image modification detection using color information and its histograms, Forensic Sci. Int. 266 (2016) 379–388.
[12] H. Li, W. Luo, J. Huang, Countering anti-JPEG compression forensics, Proceedings of IEEE International Conference on Image Processing (2012) 241–244.

[2] https://homepages.cae.wisc.edu/~ece533/images/.
[3] http://r0k.us/graphics/kodak/.

[13] C. Chen, Y.Q. Shi, JPEG image steganalysis utilizing both intrablock and interblock correlations, Proceedings of IEEE International Symposium of Circuits and Systems (2008) 3029–3032.
[14] W. Fan, K. Wang, F. Cayre, Z. Xiong, A variational approach to JPEG anti-forensics, Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing (2013) 3058–3062.
[15] P. Sutthiwan, Y.Q. Shi, Anti-forensics of double JPEG compression detection, Proceedings of International Workshop on Digital Forensics Watermarking (2011) 411–424.
[16] T. Pevny, P. Bas, J. Fridrich, Steganalysis by subtractive pixel adjacency matrix, IEEE Trans. Inf. Forensics Secur. 5 (2) (2010) 215–224.
[17] C. Pasquini, G. Boato, JPEG compression anti-forensics based on first significant digit distribution, Proceedings of IEEE International Workshop on Multimedia Signal Processing (2013) 500–505.
[18] M. Barni, M. Fontani, B. Tondi, Universal counterforensics of multiple compressed JPEG images, Proceedings of International Workshop on Digital-Forensics and Watermarking (2015) 31–46.
[19] Z. Qian, X. Zhang, Improved anti-forensics of JPEG compression, J. Syst. Softw. 91 (2014) 100–108.
[20] W. Fan, K. Wang, F. Cayre, Z. Xiong, JPEG anti-forensics with improved tradeoff between forensic undetectability and image quality, IEEE Trans. Inf. Forensics Secur. 9 (8) (2014) 1211–1226.
[21] E. Kee, M.K. Johnson, H. Farid, Digital image authentication from JPEG headers, IEEE Trans. Inf. Forensics Secur. 6 (3) (2011) 1066–1075.
[22] T. Bianchi, A. Piva, Image forgery localization via block-grained analysis of JPEG artifacts, IEEE Trans. Inf. Forensics Secur. 7 (3) (2012) 1003–1017.
[23] T. Pevny, J. Fridrich, Detection of double-compression in JPEG images for applications in steganography, IEEE Trans. Inf. Forensics Secur. 3 (2) (2008) 247–258.
[24] P. Comesaña-Alfaro, F. Pérez-González, Optimal counterforensics for histogram-based forensics, Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing (2013) 3048–3052.
[25] T. Bianchi, A. Piva, Detection of nonaligned double JPEG compression based on integer periodicity maps, IEEE Trans. Inf. Forensics Secur. 7 (2) (2012) 842–848.
[26] G. Singh, K. Singh, Forensics for partially double compressed doctored JPEG images, Multimed. Tools Appl. 75 (24) (2016) 1–18.
[27] C. Ullerich, A. Westfeld, Weaknesses of MB2, Proceedings of International Workshop on Digital Watermarking (2008) 127–142.
[28] Chambolle, An algorithm for total variation minimization and applications, J. Math. Imaging Vis. 20 (1) (2004) 89–97.
[29] A. Efros, T.K. Leung, Texture synthesis by non-parametric sampling, Proceedings of IEEE International Conference on Computer Vision (1999).
[30] F. Alter, S. Durand, J. Froment, Adapted total variation for artifact free decompression of JPEG images, J. Math. Imaging Vis. 23 (2) (2005) 199–211.
[31] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge Univ. Press, Cambridge, U.K, 2004.
[32] G. Schaefer, M. Stich, UCID-an uncompressed color image database, Proceedings of SPIE (2004) 472–480.
[33] G. Valenzise, M. Tagliasacchi, S. Tubaro, Revealing the traces of JPEG compression anti-forensics, IEEE Trans. Inf. Forensics Secur. 8 (2) (2013) 335–349.
[34] T. Pevny, T. Filler, P. Bas, Using high-dimensional image models to perform highly undetectable steganography, Proceedings of International Workshop on Information Hiding (2010) 161–177.
[35] V. Holub, J. Fridrich, Digital image steganography using universal distortion, Proceedings of ACM International Workshop on Information Hiding and Multimedia Security (2013) 59–68.