Reversible Data Hiding in Encrypted Images
Using MSBs Modulation and Histogram Modification

Ammar Mohammadi, Mohammad Ali Akhaey and Mansor Nakhkash

Abstract—This paper presents a reversible data hiding in encrypted image that employs based notions of the RDH in plain-image schemes including histogram modification and prediction-error computation. In the proposed method, original image may be encrypted by desire encryption algorithm. Most significant bit (MSB) of encrypted pixels are modulated to vacate room for embedding data bits. Modulated ones will be more resistant against failure of reconstruction if they modify for embedding data bits. At the recipient, we employ chess-board predictor for lossless reconstruction of the original image by the aim of prediction-error analysis. Comparing to existent RDHEI algorithms, not only we propose a separable method to extract data bits, but also content-owner may attain a perfect reconstruction of the original image without having data hider key. Experimental results confirm that the proposed algorithm outperforms state of the art ones.

Index Terms—Histogram modification, prediction-errors, reversible data hiding, vacating room after encryption.

I. INTRODUCTION

Reversible data hiding in plain-image (RDHPI) is drastically developed by many researchers in recent years [1]. In RDHPI secret data imperceptibly is embedded in a plain-image in a way that original image can be losslessly recovered after secret data extraction. More developed papers in RDHPI drive from three main notions that are difference expansion, histogram modification and lossless compression respectively pioneered by [2], [3] and [4]. These papers also employ prediction-error to improve hiding capacity in a determined level of distortion. Whatever more accurate prediction is attained, there is sharper histogram of the errors that can be modified to embed secret data. For example, [5] and [6] exploit gradient-adjusted prediction (GAP) and median edge detector (MED) predictors that firstly are introduced in [7] and [8], respectively. Tsai et al. [9] present a predictor, may be denoted local difference (LD) predictor, that computes difference between pixels intensities in a local area of the image and most central one in the area to bring out prediction-errors. They embed secret data via histogram modification of the prediction-errors. Sachnev et al. [10] further present cross-dot predictor that divides image into two “cross” and “dot” sets. Dot set may be predicted using cross one and vice versa. Cross-dot predictor is also represented as chess-board (CB) predictor in [11].

Besides RDHPI, reversible data hiding in encrypted image (RDHEI) is a new need of cloud computing attracted intensive attention of the researchers. In RDHEI there are three parties: image-owner, data hider and recipient. Image-owner may not trust a channel administrator (or inferior assistant) so the image-owner encrypts the image before uploading to cloud server while he/she is not motivated to compress image before encryption. On the other side, data hider, i.e. the channel administrator, is not allowed to have original-content but authorized to embed some handy data in the encrypted image. The approach should guarantee lossless original image reconstruction and error-free data extraction at the recipient. These challenges are the cause of developing RDHEI.

Schemes introduced in RDHEI may be classified into three categories, namely reserving room before encryption (RRBE) [12-19], vacating room by encryption (VRBE) [20-26] and vacating room after encryption (VRAE) [27-32]. In RRBE there is a pre-processing before encryption that enables data hider to embed data bits in the encrypted image. Most notions in VRBE are attained by encrypting some pixels intensities in a local area of the image using same cipher bytes. The approach preserves correlation of the pixels employed to embed data bits by data hider. Thus, some information remains disclosed in VRBE.

The schemes presented in RRBE and VRBE are separable means extraction of the data bits at the recipient are not tied to decrypted information. On the other hand, in joint ones data extraction can be realized just using decrypted marked image.

In VRAE schemes, data hider is completely blind to original information. After image encryption, data hider vacates rooms to embed data bits without any knowledge of original-content. The approach provides more level of security that protects image-owner privacy more than two other ones. Some methods in VRAE are joint [28, 30, 31] and some others are separable [27, 29]. In [32], two different joint and separable procedure are introduced. The separable approaches in VRAE are more functional than joint ones and further they are more functional than RRBE and VRBE. However, achieving high embedding capacity is more challengeable in VRAE than RRBE and VRBE.

In general kind of view, similar to RDHPI, most schemes in RDHEI employ correlation of neighboring pixels in an image and further they employ based idea or procedure of RDHPI such as histogram modification, difference expansion and prediction-error computation.

On the security term in using secret keys, Chen et al. [15] classify three different schemes: namely share independent secret keys (SIK), shared one key (SOK) [15] and share no secret keys (SNK) [12, 17]. In SIK there are two data hider ($K_d$) and image-owner ($K_e$) keys that are shared with recipient independently while in SOK there is just one key and in SNK there is no key to be shared. Most schemes in RDHEI, including presented VRBE and VRAE schemes, employ SIK to manage
secret keys. Schemes [15, 17] exploit difference expansion and further [12] employs histogram modification of the prediction-errors to reserve room before encryption.

Huang et al. [24] present new framework in RDHEI that makes possible using most notions of RDH in plain-image for encrypted one. In this self-contained scheme any kind of prediction technique, including GAP, MED, CB and LD may be used to bring out prediction-errors.

Schemes [19] and [20] use idea of LD predictor to embed data in the encrypted image. They realize lossless reconstruction (LR) of the original image and error-free extraction of data bits. Also, scheme [29] employs local correlation of neighboring pixels to reconstruct original one at the recipient. Using MED predictor, Yin et al. [14] bring out some labels for almost all pixels before encryption. These labels then are compressed via Huffman coding to be embedded along with data bits. They improve [19] and [20] naturally with the aim of source coding.

Fallahpour and Sedaaghi [33] present a RDHPI method that apply scheme [3] in non-overlapped blocks of the plain-image to embed data using histogram modification of pixels in a block. In the same approach, Ge et al. [25] introduce a RDHEI method that employs histogram modification ([3]) in non-overlapped blocks of the encrypted one to vacate room for embedding data bits.

As discussed, there are several schemes in RDHEI that employ based notions of RDHPI to embed data in the encrypted image. As well, we present a new scheme in RDHEI that uses based idea of the histogram modification to vacate room after encryption. Data bits just are embedded in MSBs of target pixels that are modulated to be prepared for histogram modification. Target pixels are the ones used to embed data bits. Regarding another notion that is employed in RDHPI methods, we use CB predictor for LR of the original image at the recipient. We use independent secret keys, \( k_c \) and \( k_d \) to encrypt original image and data bits, respectively.

In our scheme, not only data bits perfectly are extracted but also it is separated from original image reconstruction. This scheme can be considered as an applicable one that improves state of the art RDHEI schemes.

II. RELATED WORKS

Proposed method is inspired from histogram modification. Besides, we employ CB predictor to reconstruct original image at the recipient by analyzing prediction-error. They are discussed in more details in this section.

A. Histogram modification

As discussed, Ni et al. [3] present a RDHPI scheme using histogram modification of the original image. In the approach, they embed data bits in peak point of the histogram that is a pixel with most frequent in the image. Accordingly, they vacate room by histogram modification (VRHM) to provide empty position used for embedding data. In the approach, reversible reconstructing of the original image is possible. For the example, peak point in the histogram of Lena (Fig. 1a) is "155". One to vacate room adds all intensities of the image more than

\[ "155" \] by 1 (Fig. 1b).

In the proposed procedure, we apply the idea of VRHM in the encrypted image.

B. Chess-board predictor

The CB predictor, i.e. non-causal, can provide prediction better than the causal ones, MED and GAP. As shown in Fig. 2 in CB predictor a black/white pixel may be predicted using white/black neighbors. For example, prediction of a white pixel \( \mathcal{P}_q \) employing neighboring black ones \( \{ p_{q-1}, p_{q}, p_{q+1}, p_{q+3} \} \) is done by

\[
\mathcal{P}_q = \left[ \frac{p_{q-1} + p_{q} + p_{q+1} + p_{q+3}}{4} \right]
\]

where \([.]\) is the round function. Using \( \mathcal{P}_q \) and \( \mathcal{P}_q \), the prediction-error \( e_q \) is computed by

\[
e_q = \mathcal{P}_q - \mathcal{P}_q
\]

more specifically let’s denote the prediction of a white pixel by neighboring black ones as white plus prediction (WPP). In addition, a central black pixel \( (P_q) \) can be predicted using neighboring black ones, \( \{ p_{q-2}, p_{q-1}, p_{q+2}, p_{q+3} \} \), by
prediction-errors in an image. Accordingly, we define

$$f_{\text{predictor}} = 1/L$$  \hspace{1cm} (7)

as a probability of failure in reconstruction of a reformed MSB of the pixel that is randomly picked up. Accordingly, we have $f_{\text{WPP}} = 0.0017$, $f_{\text{BCP}} = 0.021$, $f_{\text{GAP}} = 0.014$ and $f_{\text{MED}} = 0.016$ for Baboon image. Although $f_{\text{WPP}}$ significantly is less than others, there is still a probability of failure in using WPP. As a solution diminishing risk of failure, we get together MSBs of the target pixels. Therefore, $N$ ones of the MSBs are modulated and so modulated one is used to carry data bit.

It can be proved that there is always a $N$ can be taken to guarantee LR of the original image. The amount of computed $f_{\text{pre}}$ depends not only on the kind of an employed predictor but also the type of a chosen image. The lower the entropy may be provided by an image, the lower the probability of the failure in reconstructing can be attained. In other words, generally, smoother images have less $f_{\text{pre}}$. For example, unlike Baboon, Lena and F16 have $f_{\text{WPP}} = 0$.

Accordingly, in the proposed method we employ WPP predictor. We may also exploit BCP to achieve more EC.

III. PROPOSED SCHEME

In this section, we introduce our proposed method in details including image encryption, simple scrambling and unscrambling, embedding and extracting data bits, recovering the original image and overall view of the proposed scheme.

A. Image encryption

Encryption of the original image is done using content-owner secret key, $K_C$. However, for encryption, not only any method of encryption but also any standard encryption algorithm can be exploited. The only restriction is no change in positions of the image pixels during encryption. If it occurs, data hider will have knowledge of the new positions to realize proposed procedure. Therefore, in term of security, content-owner can chose a desire encryption algorithm.

Regarding $O$ as the original image, encrypted image is denoted $\hat{O}$. Accordingly, let’s classify pixels of the $\hat{O}$ to three groups, namely encrypted white target pixels (EWTPs), $\{\hat{p}_0, \hat{p}_1, ..., \hat{p}_{q-1}, \hat{p}_q\}$, encrypted black pixels (EBPs), $\{\hat{p}_0, \hat{p}_1, ..., \hat{p}_{q-1}, \hat{p}_q\}$, including encrypted reference black pixels (ERBPs) and encrypted target black pixels (ETBPs).
B. Simple scrambling and unscrambling

Let's assume a set \( X = \{x_0, x_1, ..., x_q, ..., x_0\} \) that is going to be scrambled in a simple way. In the approach, \( X \) is separated into \( N \) smaller sets, including \( \frac{q+1}{N} \) members, i.e. \( N \in \mathbb{N} \). In view on that, putting the corresponding members of each set together, scrambled one \( Y = \{y_0, y_1, ..., y_1, ..., y_0\} \) is brought out. The number of members in \( X \) set, \( (Q+1) \), is selected in such a way to be divisible by \( N \). For example, regarding \( X = \{x_0, x_1, x_2, x_3, x_4, x_5\} \) and \( N = 3 \), by dividing \( X \) into 3 smaller sets, we have \( X_1 = \{x_0, x_1\} \), \( X_2 = \{x_2, x_3\} \) and \( X_3 = \{x_4, x_5\} \). The new positions, \( Y = \{x_0, x_2, x_4, x_1, x_3, x_5\} \), are achieved by putting the corresponding members of each set together.

Regarding simple scrambling, that scrambles \( X \) to \( Y \), in a reverse procedure, \( Y \) set is divided into \( \frac{q+1}{N} \) smaller ones including \( N \) members and corresponding members of each set get together to reversibly bring out \( X \). As an example, considering \( Y = \{x_0, x_3, x_4, x_1, x_2, x_5\} \) and \( N = 3 \), we divide \( Y \) into 2 sets, \( Y_1 = \{x_0, x_3, x_4\} \) and \( Y_2 = \{x_1, x_2, x_5\} \). Putting together corresponding members of these two sets, \( X = \{x_0, x_1, x_2, x_3, x_4, x_5\} \) is achieved outcome of a simple unscrambling.

C. MSB modulation, demodulation and embedding data bits

In this section, we demonstrate procedure of embedding data bits in the modulated MSBs of the encrypted pixels. MSB modulation gives more intensive value prevents failure of reconstruction. In other words, employing modulation, several MSBs must be changed instead of just one MSB to embed data bit that provides more resistant against failure of reconstruction.

This modulation is done by a new binary representation of \( N \)-bits of MSBs that take different positions. It represents a whole number less than \( 2^N \). Using the histogram modification of the modulated MSBs some rooms are vacated to realize RDHEI.

Let's assume, \( \hat{\mathcal{P}}_q = \hat{\mathcal{P}}_{(q)7} \ldots \hat{\mathcal{P}}_{(q)3} \ldots \hat{\mathcal{P}}_{(q)0} \), as a bitwise demonstration of an encrypted pixel, i.e. from LSB, \( \hat{\mathcal{P}}_{(q)0} \), to MSB, \( \hat{\mathcal{P}}_{(q)7} \). As mentioned when (6) is confirmed for \( \hat{\mathcal{P}}_{(q)} \), \( \hat{\mathcal{P}}_{(q)7} \) may be replaced by a bit of data so that it is perfectly retrievable. Regarding MSBs of \( Q \) ones of the target pixels, \( \mathcal{P}_7 = \{\hat{\mathcal{P}}_{(0)7}, \hat{\mathcal{P}}_{(1)7}, ..., \hat{\mathcal{P}}_{(a)7}, ..., \hat{\mathcal{P}}_{(Q)7}\} \), we describe process of modulation by

\[
m_j = \sum_{k=0}^{N-1} 2^k \times (\hat{\mathcal{P}}_{(q-k)7}, q = N \times (j + 1) - 1, j = 0, 1, ..., J \quad (8)
\]

Algorithm 1: Shrinking the value of \( m_j \) for \((0 \leq j \leq J)\).

For \( j = 0 \) to \( J \) do

\[\delta_j = m_j\]

if \((m_j < 2^{N-1})\) then

\[\delta_j = 2^N - 1 - m_j\]

end if

end for

Algorithm 2: Embedding data bits for \((1 \leq q < m \times l)\).

For \( j = 0 \) to \( J \) do

\[c_j = \delta_j\]

if \((\hat{\Delta}_q = 1)\) then

\[c_j = 2^N - 1 - \delta_j\]

end if

end for

so modulated MSBs are denoted \( M = \{m_0, m_1, ..., m_j, ..., m_J\}, 0 \leq m_j < 2^N, J = \frac{2^N - 1}{N} \).

Regarding the nature of the encrypted image, the histogram of \( M \) demonstrates a uniform distribution. We should vacate rooms employing histogram modification to embed data bits. The approach is a shrinking process that shifts each \( m_j \) less than \( 2^{N-1} \) to a larger one. The shrunken ones, \( S = \{s_0, s_1, ..., s_j, ..., s_J\}, 2^{N-1} \leq s_j < 2^N \), is achieved using Algorithm 1.

In this algorithm, if \( m_j < 2^{N-1} \), it is reformed to \( s_j \) what has most bitwise mutation than \( m_j \). In other words, \( m_j \) is replaced with its 1’s complement.

Data bits may be encrypted using \( K_d \). Vacating rooms by the shrinking, \((J+1)\) bits of the encrypted data, \( \mathcal{D} = \{\hat{d}_0, \hat{d}_1, ..., \hat{d}_j, ..., \hat{d}_J\} \), may be embedded in \( S \) employing Algorithm 2. Consequently, we have carrier ones, \( \mathcal{C} = \{c_0, c_1, ..., c_p, ..., c_q\} \). In this algorithm, to embed a encrypted data bit with value “1” in \( s_j \), it is replaced by its 1’s complement and to embed a one with value “0” it is remained intact.

Let’s demonstrate the process of the embedding more precisely with an example having \( N = 3 \) and modulation of 27 MSBs, \( \mathcal{M} = \{3,0,4,7,2,6,1,5\} \). Employing Algorithm 1, rooms are vacated by shrinking procedure, \( S = \{4,7,4,7,5,5,6,6,5\} \). In other words, using histogram modification of \( \mathcal{M} \), modulated ones between 0 and 3 are shifted to make rooms embed \( \mathcal{D} = \{1,0,1,0,0,1,0,1\} \). Eventually, exploiting Algorithm 2 carrier set, \( \mathcal{C} = \{3,0,4,0,5,5,1,6,2\} \), is achieved. Histogram of \( \mathcal{M} \), \( S \) and \( \mathcal{C} \) sets are depicted in Fig. 4-a, b and c respectively.

In order to create a marked encrypted image, at first, the set \( \mathcal{C} \) is demodulated using

\[
[\hat{\mathcal{P}}_{(N \times j + k-1)7}] = \text{mod}(\frac{c_j}{2^{\frac{N-1}{2}}}, 2), 0 \leq j \leq J, 0 < k \leq N \quad (9)
\]

Consequently, we have marked encrypted MSBs, \( [\hat{\mathcal{P}}_q] = \{[\hat{\mathcal{P}}_{(0)7}], [\hat{\mathcal{P}}_{(1)7}], ..., [\hat{\mathcal{P}}_{(a)7}], ..., [\hat{\mathcal{P}}_{(Q)7}]\} \). A marked pixel, \( \hat{\mathcal{P}}_{(q)} \), \( q = 0, 1, ..., Q \), is created by replacement of a \( \hat{\mathcal{P}}_{(q)7} \) in

![Fig. 4. The histogram of an example of \( \mathcal{M} \), \( S \) and \( \mathcal{C} \) sets.](image-url)
MSB of $\tilde{p}_a$. Therefore, we have marked EWTPs, $[\tilde{P}] = \{[[\tilde{p}_0]], [[\tilde{p}_1]], ..., [[\tilde{p}_a]], ..., [[\tilde{p}_q]]\}$.

D. Extracting data bits

For extracting data bits, we just need data hidder key, $K_d$. Firstly, MSBs of the target pixels in the marked encrypted image are picked up, denoted $[\tilde{P}_7]$. In this paper, reconstructed sets or pixels are in bold, $[\tilde{P}_7]$ is modulated to recover $C$ using (8) where $(\tilde{p}_{(a-k),7})$ and $m_j$ respectively are replaced by $[\tilde{P}_{(a-k),7}]$ and $e_j$, $j = 0,1, ..., J$. Having $C = \{c_0, c_1, ..., c_j, ..., c_q\}$, extraction of the data bits is done employing Algorithm 3. In Algorithm 3, $j$’th bit of data, $\tilde{d}_j$, is extracted using $c_j$. They are decrypted using $k_d$ attaining $D$.

E. Recovering the original image

Reconstructing the original image may be initiated by decryption of the marked encrypted image using $K_d$. Let’s assume $P' = \{p'_0, p'_1, ..., p'_a, ..., p'_q\}$ is the decrypted white target pixels (DWTTPs). The subject is reconstructing the MSB of DWTTPs, $P'_s = \{p'_{0,7}, p'_{1,7}, ..., p'_{a,7}, ..., p'_{q,7}\}$. Other bits remain intact.

The reconstructing process continues by calculating 1’s complement of $P'_s$ that denoted $P''_s = \{p''_{0,7}, p''_{1,7}, ..., p''_{a,7}, ..., p''_{q,7}\}$. By replacement of $P'_s$ and $P''_s$ in MSB of DWTTPs we respectively have $P' = \{p'_0, p'_1, ..., p'_a, ..., p'_q\}$ and $P'' = \{p''_0, p''_1, ..., p''_a, ..., p''_q\}$.

After decryption, black pixels completely remain intact and are employed to predict $p'_a$ using (1) where $p'_a$ is replaced by $p'_b$. Therefore, having predicted one $p'_b$, prediction-error ($e'_b$) may be calculated using (2).

Having $P'_s = \{e'_0, e'_1, ..., e'_a, ..., e'_q\}$, similarly, $P''_s = \{e''_0, e''_1, ..., e''_a, ..., e''_q\}$ is also computed.

In the following, summation of $N$ amounts of errors is achieved by

$$E'_j = \Sigma_{k=0}^{N-1} |e'_{(a-k)}|, q = N \times (j + 1) - 1 \quad (10)$$

Accordingly, $E''_j$ is attained by (10) where $e'_{(a-k)}$ is replaced with $e''_{(a-k)}$. Finally, the original pixels belong to $P'$ or $P''$ are retrieved by comparing $E'_j$ and $E''_j$ employing Algorithm 4.

Accordingly, for each $j$, if $E''_j \leq E'_j$ the recovered pixels belong to $N$-pixels of $P''$ and in vice versa they belong to $P'$. Eventually, the recovered ones are $P = \{p'_0, p'_1, ..., p'_q, ..., p'_q\}$.

Having $[p_{-N+1}, ..., p_a] = [p''_{-N+1}, ..., p_a]$ means $N$-pixels of the original image are losslessly reconstructed.

If $E'_j$ and $E''_j$ are close together it will be a high risk in LR of the original pixels. Therefore, risk of LR may be analyzed by computing difference between $E'_j$ and $E''_j$.

---

Algorithm 3: Extracting data bits.

for $j = 0$ to $J$ do

if ($e'_j < 2^{N-1}$) then

$d_j = 1$

else if ($2^{N-1} \leq e'_j < 2^N$) then

$d_j = 0$

end if

end for

Algorithm 4: Reconstructing the original pixels.

for $j = 0$ to $J$ do

$q = N \times (j + 1) - 1$

if ($E''_j \leq E'_j$) then

$[p_{-N}, ..., p_a] = [p''_{-N}, ..., p_a]$

else

$[p_{-N}, ..., p_a] = [p''_{-N}, ..., p_a]$

end if

end for

Algorithm 5: Risk of LR.

if ($R_j < 16 \times N$) then

High risk (HR)

else if ($R_j < 32 \times N$) then

Median risk (MR)

else if ($R_j < 64 \times N$) then

Low risk (LR)

else

Very low risk (VLR)

end if

$$R_j = |E'_j - E''_j| \quad (11)$$

The larger $R_j$ may be calculated, the lower risk of LR can be achieved and vice versa. Employing $R_j$ and $N$ in Algorithm 5, we can define 4 classes of risk: namely high risk (HR), median risk (MR), low risk (LR) and very low risk (VLR). Therefore, in reconstruction of each $N$-pixels set we have a measurement of risk. The bigger $N$ is chosen, the more accurate evaluation of risk is realized.

F. Overall view of the proposed method

In Fig. 5a, we depict general diagram of the proposed scheme. As shown, at first, embedding is done on EWT as demonstrated in Subsection C. Then, in a similar way it can be continued by EBTP, $P = \{p_0, p_1, ..., p_q\}$ to increase EC. In more explanation, by choosing N-bits, modulation can also be done for $P_7 = \{p_{0,7}, p_{1,7}, ..., p_{q,7}\}$. $N$ generally is considered bigger than $N$ because WPP predictor can predict better than BCP one, as discussed. Thus, modulated ones $M = \{m_0, m_1, ..., m_q\}, \quad 0 \leq m_j < 2^N, \quad j = \frac{q-N+1}{N}$, are
achieved using (8). Accordingly, shrunken ones, \( S = \{ s_0, s_1, \ldots, s_j, \ldots, s_J \} \), \( 2^{N-1} \leq s_j < 2^N \), are attained using Algorithm 1, then carrier ones, \( C = \{ c_0, c_1, \ldots, c_j, \ldots, c_J \} \), are reached exploiting Algorithm2 and eventually marked EBTPs, \( \hat{P} = \{ \hat{p}_0, \hat{p}_1, \ldots, \hat{p}_q, \ldots, \hat{p}_Q \} \), are achieved employing (9). Thus, marked encrypted image, including marked EWTPs and EBTPs, is created. Meanwhile, BRPs just are encrypted without any more modification. They will be employed to reconstruct other pixels at the recipient. They form 25% of the image.

Let’s assume an original image in size \( P \times Q \) so regarding a set of preferred \( \{ N, N \} \) and employing all target pixels including black and white, EC in bits is achieved by

\[
EC = P \times Q \times \left( \frac{1}{2^N} + \frac{1}{4^N} \right) \tag{12}
\]

choosing \( \{ N = 1, N = 1 \} \) most possible EC, \( \frac{3}{4} (P \times Q) \), is achieved.

In Fig. 5b, extracting data bits and reconstruction of the original image are shown that can be done in a separable procedure.

Extracting data from marked EBTP, \( \hat{P} \), similar to marked EWTP may be done (Subsection D).

Reconstructing the image is initiated by decryption of the marked encrypted image using \( K_e \). Therefore, BRPs are completely recovered just by decryption. Using BRPs, BTPs may be recovered in a similar procedure described in Subsection E except using (3) and (4) instead of (1) and (2) respectively to predict errors. Having black pixels, BPs, WTPs are recovered as also discussed in Subsection E. Risk of LR for \( N \)-pixels/\( N \)-pixels can be computed in meanwhile of recovering.

The closer the pixels are the more correlation they have. By breaking the correlation, we can diminish risk of LR at the recipient. The approach may be realized by scrambling of MSB of target pixels before modulation. In other words, the scrambling prevents modulation of MSBs belong to near rough regions to keep together. Although, the scrambling can randomly be performed using \( K_d \), for less complexity, in this scheme, we scramble the encrypted MSBs, \( \frac{P_2}{P_1} \), in a simple way (Subsection B). Thus, presenting more efficient procedure, the scrambled MSBs are modulated, marked and finally unscrambled to form the corresponding marked encrypted pixels.

If MSBs of target pixels are scrambled by data hider to embed data bits, MSBs of encrypted target pixels will scramble as well in the same approach for extracting data. Accordingly, at the recipient, for original image reconstruction, we must scramble prediction-errors before computing \( N/N \) amounts of the prediction-errors. Therefore, in Algorithm 4, we should consider corresponding pixels of the scrambled prediction-errors for reconstructing original ones.

**IV. EXPERIMENTAL RESULT**

The performance of the proposed separable VRAE algorithm is confirmed designing several experiments. Twelve grayscale images F16, Lena, Splash, Splash, House, Boat, Elaine, Lake, Peppers, Baboon, Stream, Aerial and APC from the USC-SIPI database are used as test images. Also, BOWS2 original database, including 10000 greyscale images, are employed to verify that the proposed algorithm can achieve LR of the original image and error-free extraction of the data bits. All test images are \( 512 \times 512 \) in size. In experiments, the first and the last two rows and columns of the image are ignored in data embedding process. Peak signal-to-noise ratio (PSNR) is used to estimate the quality of a recovered image. PSNR=∞ means LR of the original image.
Choosing a set of \( \{ N, N \} \) may be tied to entropy of an image. The greater the entropy of the image is estimated, the greater the value of \( \{ N, N \} \) must be selected for LR. However, the more \( \{ N, N \} \) is preferred, the less EC is attained. Therefore, there is a tradeoff between EC and PSNR of the reconstructed image. In Table I, the precise \( N/N \) for LR of the test images is inserted. By taking \( \{ N = 1, N = 1 \} \) for Elaine the most possible EC in the proposed method, 193548 bits, is achieved. In the other side, Stream image provides lowest EC. In this image, the lowest possible \( N/N \) for LR is \( \{ N = 2, N = 4 \} \) that definitely may be employed for LR of the other test images. Also, in the table, risk of LR is evaluated by computing the number of \( N \)-pixels that take HR or MR. Although F16, Peppers and Baboon are reconstructed perfectly, they respectively include 4, 8 and 5 high risk sets of \( N \)-pixels. About reconstruction of \( N \)-pixels, Baboon and Elaine take 13 and 1 sets of high risk, respectively. Considering the number of HR and MR sets, generally, Baboon have most risk of LR. However, by choosing \( \{ N = 3, N = 5 \} \) for embedding data in Baboon, no set of HR is left. It is attained by paying the cost of diminishing EC to 55913 bits.

On the other hand, we assume that the proposed scheme is completely blind to original-content so we can not find out the least possible precise \( \{ N, N \} \) for LR. However, in the next subsection we confirm that a set of \( \{ N, N \} \) may be always found out for LR.

In the proposed scheme, error-free extraction of data bits is realized for all test images under any circumstances.

### A. Lossless retrieval

In Fig. 6 we demonstrate performance of the proposed scheme in LR and error-free extraction of data bits using 10000 test images of BOWS2 original database. As shown, for various \( \{ N, N \} \) we accomplished the proposed algorithm to bring out the number of failure in LR. For example, in \( \{ N = 2, N = 6 \} \), from 10000 test images there is 11 images that are not perfectly reconstructed so the failure rate is \( \mathcal{F}_r = 0.0011 \). As shown, the more \( \{ N, N \} \) is preferred, the less EC and the less failure are provided. In \( \{ N = 4, N = 6 \} \), \( \mathcal{F}_r \) is zero. Thus, there is permanently a set of \( \{ N, N \} \) to achieve LR.

As demonstrated in the figure, modification of \( N \) affects EC and \( \mathcal{F}_r \) more tangible than \( N \). As instance, in \( N = 6 \), increasing \( N \) from 2 to 3 diminishes EC and \( \mathcal{F}_r \), as much as 21506 bits and 0.0249, respectively while in \( N = 3 \), increment of \( N \) has no noticeable impact on EC and \( \mathcal{F}_r \).

Implementing the proposed algorithm by \( \{ N = 3, N = 6 \} \), we descendingly sort all 10000 reconstructed images by the number of their \( N \)-pixels that are HR. The first six sorted images are listed in Table II. As discussed, in this implementation, there are 11 failures. In the table, PSNR demonstrates four images that is not reconstructed losslessly and so 4 out of all 11 failures include in first six sorted images. It proves risk analysis is a proper assessment for evaluation LR, i.e. all 11 failures are included in 50 first sorted images.

In any failure, there is some deformed MSBs that can not be recovered correctly. The number of deformed MSBs are tabulated for images. The maximum one is just 12 bits for 6501.pgm image so there is not much error rate of bits when LR does not realized.

### B. A visual demonstration of the proposed scheme

Fig. 7 is a visual demonstration of the proposed procedure including original, encrypted, marked encrypted and reconstructed images with their depicted histograms. We optionally employ AES in counter mode, i.e. a stream cipher procedure, to encrypt the test image, Lena. As shown in Fig. 7b, after encryption there is no knowledge of the original image that remains discovered. As a proof, its histogram is uniformly distributed. We guarantee the security of the proposed algorithm because of having two important property in the represented algorithm. 1- There is no feature or knowledge of the original-content that remains disclosed. 2- Choosing a desire encryption algorithm is possible regarding its pros and cons including the security. Since our method is completely blind to the original-content, it absolutely preserves content-owner privacy. Moreover, content owner can encrypt original-
TABLE II  SIX IMAGES OF BOWS2 ORIGIANL DATABASE THAT HAVE MOST THE NUMBER OF HR SETS OF N-Pixels.

| Items          | Images (.pgm) |       |       |       |       |
|----------------|---------------|-------|-------|-------|-------|
|                |   6502        | 6501  | 6537  | 6498  | 6514  | 9373  |
| N-pixels sets  | HR           | 15    | 11    | 10    | 6     | 4     | 4     |
|                | MR           | 947   | 923   | 778   | 943   | 415   | 460   |
| N-pixels sets  | HR           | 1     | 1     | 2     | 1     | 0     | 2     |
|                | MR           | 285   | 259   | 413   | 415   | 60    | 170   |
| PSNR           | 52.39        | 49.38 | ∞     | 52.39 | 55.4  | ∞     |        |

The number of deformed-MSBs

Table: Efficiency comparison between the proposed method and other separable VRAE ones for 9 test images.

| Schemes        | Items  | Images          |
|----------------|--------|-----------------|
|                |        | F16  Lena  Splash  House  Boat  Elaine  Lake  Peppers  Baboon |
| Zhang2012 [27] | LR     | 1920  1920  1920  1920  1920  1920  1920  Pass  Pass  Pass |
| Wu [32]        | EC     | 130050  130050  130050  130050  130050  130050  Fail  Fail  Fail |
|                | LR     | Pass  Pass  Pass  Fail  Fail  Fail  Fail  Pass  Pass  Pass |
| Qian [29]      | EC     | 77376  77376  77376  77376  77376  77376  77376  Fail  Fail  Fail |
|                | LR     | Pass  Pass  Pass  Fail  Fail  Fail  Fail  Pass  Pass  Pass |
| Proposed       | EC     | 86021  86021  86021  86021  86021  86021  86021  Fail  Fail  Fail |
|                | LR     | Pass  Pass  Pass  Pass  Pass  Pass  Pass  Pass  Pass  Pass |
|                | LR just by K_e | Pass  Pass  Pass  Pass  Pass  Pass  Pass  Pass  Pass  Pass |

Fig. 7. A visual demonstration of the proposed scheme.

Fig. 8. PSNR comparison between our proposed scheme and other schemes Cao et al. [16], Zhang2011 [30], Zhang2012 [27], Qian [29] and Wu [32] for four test images.
content using any key-based encryption algorithm including symmetric or asymmetric and in the symmetric one stream cipher or block cipher.

Fig. 7c is a description of the marked encrypted image. The histogram still has a uniform distribution. Original image is losslessly reconstructed as shown in Fig. 7d.

C. Comparison with other schemes

Separable VRAE methods are more functional than others. The proposed scheme is an only one of separable VRAE methods that makes possible LR of the original image just by having secret key. There is some experiments to confirm it.

In Table III, we compare proposed scheme with other separable VRAE schemes. Schemes [27] and [29] compress some bits of encrypted pixels to vacate room in order to embed data bits. Meanwhile, they employ some parameters can be used to form a tradeoff between EC and lossless recovery. Their functionality is similar to $N/N$. For fairness comparison between our method and [27] and [29] we employ these parameters so that LR of the all test images is accomplished. In the approach, we apply parameters ($M = 4$, $S = 2$, $L = 271$), $q = 0.1$ and $N = 2$, $N = 3$ to [27], [29], and the proposed one, respectively. It should be noted that by choosing less $L$, $q$ and $N$, both EC and risk of LR are enlarged respectively in [27], [29] and ours. It generally leads more probability of failure in LR.

In [29], for LR at the recipient side information needed that can be available employing $K_d$. Therefore, as presented in Table III, for this scheme just a high quality version of the original image can be retrieved without having $K_d$ while in the proposed scheme LR of the original image can be accomplished independent of data extraction. Besides, our algorithm can achieve more EC than [29].

As described in Table III, in comparison with [27] we significantly improve EC. Similar to [29] their algorithm is dependent to having $K_d$ for LR of the original image.

In [32] scheme, using data hider key, they prefer some target pixels in the encrypted image to embed data bits and employ a predictor may be denote a modified WPP predictor to reconstruct original image at the recipient. Their methods do not guarantee LR, although it may be possible by diminishing the number of selected target pixels. In this experiment, we consider the most number of target pixels may be chosen in this scheme. In the approach, their algorithm fails in perfect reconstruction of the original image for five test images. As described in Table III, we improve their algorithm designing a procedure to guarantee LR realized by employing histogram modification and MSBs modulation. However, we pay the cost by less EC. All discussed schemes and ours provide an error-free data bits extraction.

In Fig. 8 we compare our scheme with [16], [30], [27], [29] and [32] in quality of the reconstructed original image. In this experiment for reconstructing we suppose [30] and [32] have both encryption and data hider keys while others just have encryption key. In [30] method, original image reconstruction and data bits extraction are joint while our scheme is a separable one. In comparison, we improve not only EC but also image quality. To embed data bits, [16] reserves rooms before encryption aim of patch-level sparse representation. Generally, achieved EC in the proposed scheme is comparable with this scheme. In their scheme, a preprocessing is allowed before encryption while we absolutely are blind to original-content.

As shown, [29] outperforms [16] and [27] in term of PSNR. Achieving LR, we outperform [29].

In Peppers and Lena we improve EC in comparison with [32]. In Lake, Peppers and Baboon they achieve a PSNR less than 55 dB even by using both keys while we realize LR.

V. CONCLUSION

In this paper, by comparing different predictors, we show that WPP is the better one used to diminish probability of failure in reconstruction of the original image. Moreover, BCP predictor is employed to increase embedding capacity. By prediction-error analysis, we just choose MSBs of encrypted target pixels to embed data bits. These MSBs are modulated to be more resistant against failure of reconstruction when they are modified to embed data bits. Employing histogram modification of the modulated ones, we vacate rooms to embed data bits. At the recipient, in a separable way data bits are extracted and original image is reconstructed. We employ the risk measurement to have a probability of LR.

The proposed method improves other separable VRAE schemes by aim of MSBs modulation and histogram modification. We are the only one that realizes LR without having $K_d$.

REFERENCES

[1] Y.-Q. Shi, X. Li, X. Zhang, H.-T. Wu, and B. Ma, “Reversible data hiding: advances in the past two decades,” IEEE Access, vol. 4, pp. 3210-3237, 2016.
[2] J. Tian, “Reversible data embedding using a difference expansion,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 8, pp. 890-896, Aug. 2003.
[3] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, “Reversible data hiding,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, no. 3, pp. 354-362, Mar. 2006.
[4] T. Kalker, and F. M. Willems, “Capacity bounds and constructions for reversible data-hiding,” in Proc. International Conference on Digital Signal Processing, Santorini, Greece, Greece, 2002, pp. 71-76.
[5] X. Li, W. Zhang, X. Gui, and B. Yang, “A novel reversible data hiding scheme based on two-dimensional difference-histogram modification,” IEEE Transactions on Information Forensics and Security, vol. 8, no. 7, pp. 1091-1100, Jul. 2013.
[6] D. M. Thodi, and J. J. Rodriguez, “Expansion embedding techniques for reversible watermarking,” IEEE Transactions on Image Processing, vol. 16, no. 3, pp. 721-730, Mar. 2007.
[7] X. Wu, and N. Memon, “Context-based, adaptive, lossless image coding,” IEEE Transactions on Communications, vol. 45, no. 4, pp. 437-444, Apr. 1997.
[8] M. J. Weinberger, G. Serroussi, and G. Sapiro, “The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS,” IEEE Transactions on Image Processing, vol. 9, no. 8, pp. 1309-1324, Aug. 2000.
[9] P. Tsai, Y.-C. Hu, and H.-L. Yeh, “Reversible image hiding scheme using predictive coding and histogram shifting,” Signal Processing, vol. 89, no. 6, pp. 1129-1143, Jun. 2009.
[10] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y. Q. Shi, “Reversible watermarking algorithm using sorting and prediction,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, no. 7, pp. 989-999, Jul. 2009.
[11] A. Mohammadi, and M. Nakhkash, “Sorting methods and adaptive thresholding for histogram based reversible data hiding,” arXiv preprint arXiv:1907.05123, 2019.

[12] S. Xiang, and X. Luo, “Reversible data hiding in homomorphic encrypted domain by mirroring ciphertext group,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 11, pp. 3099-3110, Nov. 2018.

[13] K. Ma, W. Zhang, X. Zhao, N. Yu, and F. Li, “Reversible data hiding in encrypted images by reserving room before encryption,” IEEE Transactions on Information Forensics and Security, vol. 8, no. 3, pp. 553-562, Mar. 2013.

[14] Z. Yin, Y. Xiang, and X. Zhang, “Reversible data hiding in encrypted images based on multi-MSB prediction and huffman coding,” IEEE Transactions on Multimedia, Aug. 2019.

[15] Y.-C. Chen, T.-H. Hung, S.-H. Hsieh, and C.-W. Shiu, “A new reversible data hiding in encrypted image based on multi-secret sharing and lightweight cryptographic algorithms,” IEEE Transactions on Information Forensics and Security, vol. 14, no. 12, pp. 3332-3343, Dec. 2019.

[16] X. Cao, L. Du, X. Wei, D. Meng, and X. Guo, “High capacity reversible data hiding in encrypted images by patch-level sparse representation,” IEEE Transactions on Cybernetics, vol. 46, no. 5, pp. 1132-1143, May, May 2016.

[17] C.-W. Shiu, Y.-C. Chen, and W. Hong, “Encrypted image-based reversible data hiding with public key cryptography from difference expansion,” Signal Processing: Image Communication, vol. 39, pp. 226-233, Nov. 2015.

[18] P. Pateaux, and W. Puech, “An efficient MSB prediction-based method for high-capacity reversible data hiding in encrypted images,” IEEE Transactions on Information Forensics and Security, vol. 13, no. 7, pp. 1670-1681, Jul. 2018.

[19] A. Mohammadi, and M. Nakhkash, “Reversible data hiding in encrypted images using local difference of neighboring pixels,” arXiv preprint arXiv:1907.05123, 2019.

[20] S. Yi, and Y. Zhou, “Separable and reversible data hiding in encrypted images using parametric binary tree labeling,” IEEE Transactions on Multimedia, vol. 21, no. 1, pp. 51-64, Jan. 2019.

[21] D. Xu, and R. Wang, “Separable and error-free reversible data hiding in encrypted images,” Signal Processing, vol. 123, pp. 9-21, Jun. 2016.

[22] Z. Yin, B. Luo, and W. Hong, “Separable and error-free reversible data hiding in encrypted image with high payload,” The Scientific World Journal, vol. 2014, Apr. 2014.

[23] W. Zhang, K. Ma, and N. Yu, “Reversibility improved data hiding in encrypted images,” Signal Processing, vol. 94, no. 1, pp. 118-127, Jan. 2014.

[24] F. Huang, J. Huang, and Y.-Q. Shi, “New framework for reversible data hiding in encrypted domain,” IEEE Transactions on Information Forensics and Security, vol. 11, no. 12, pp. 2777-2789, Dec. 2016.

[25] H. Ge, Y. Chen, Z. Qian, and J. Wang, “A high capacity multi-level approach for reversible data hiding in encrypted images,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 8, pp. 2285 - 2295, Aug. 2019.

[26] X. Zhang, “Commutative reversible data hiding and encryption,” Security and Communication Networks, vol. 6, no. 11, pp. 1396-1403, 2013.

[27] X. Zhang, “Separable reversible data hiding in encrypted image,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 7, no. 2, pp. 826-832, Apr. 2012.

[28] J. Zhou, W. Sun, L. Dong, X. Liu, O. C. Au, and Y. Y. Tang, “Secure reversible image data hiding over encrypted domain via key modulation,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 3, pp. 441-452, Mar. 2016.

[29] Z. Qian, and X. Zhang, “Reversible data hiding in encrypted images with distributed source encoding,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 4, pp. 636-646, Apr. 2016.

[30] X. Zhang, “Reversible data hiding in encrypted image,” IEEE Signal Processing Letters, vol. 18, no. 4, pp. 255-258, Apr. 2011.

[31] W. Hong, T.-S. Chen, and H.-Y. Wu, “An improved reversible data hiding in encrypted images using side match,” IEEE Signal Processing Letters, vol. 19, no. 4, pp. 199-202, Apr. 2012.

[32] X. Wu, and W. Sun, “High-capacity reversible data hiding in encrypted images by prediction error,” Signal Processing, vol. 104, pp. 387-400, Nov. 2014.