Signal Processing: Image Communication 64 (2018) 78–88

Contents lists available at ScienceDirect

Signal Processing: Image Communication

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

Reversible data hiding in encrypted images using adaptive block-level prediction-error expansion

Shuang Yi [a], Yicong Zhou [a,*], Zhongyun Hua [b]

[a] *Department of Computer and Information Science, University of Macau, Macau 999078, China*
[b] *School of Computer Science and Technology, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen 518055, China*

ARTICLE INFO

Keywords:
Reversible data hiding
Encrypted images
Adaptive embedding
Block-level prediction-error expansion

ABSTRACT

As directly reserving room from the encrypted image for data embedding is difficult and inefficient, many encryption domain based reversible data hiding schemes have disadvantages such as small embedding rate and low visual quality of the directly decrypted image. In order to solve these problems, this paper first introduces a reversible data hiding method for natural images using the block-level prediction-error expansion. The method can embed secret data into 2 × 2 image blocks by exploiting the pixel redundancy within each block. Extending this concept to the encrypted domain, we then propose a reversible data hiding method in encrypted images using adaptive block-level prediction-error expansion (ABPEE-RDHEI). ABPEE-RDHEI encrypts the original image by block permutation to preserve spatial redundancy for data embedding, and applies a stream cipher to the block permutated image to further enhance the security level. Due to the adaptive pixel selection and iterative embedding processes, the proposed ABPEE-RDHEI can achieve a high embedding rate and pleasing visual quality of the marked decrypted images. Experimental results and analysis show that ABPEE-RDHEI has a better performance than several state-of-the-art methods.

1. Introduction

Reversible data hiding (RDH) in images aims to embed the secret data into an original image in an imperceptible way. Unlike watermarking that should be robust to against malicious attacks, RDH emphasizes perfect data extraction and image recovery at the receiver side. Many efficient RDH methods have been proposed in recent years. The early RDH is based on image compression, where a number of possible features of the original image are extracted and losslessly compressed. The reserved spare space is utilized for embedding the secret data. Later, more RDH methods have been proposed by exploiting the pixel spatial correlations of the image (e.g., difference expansion (DE) [1] and histogram shifting (HS) [2–4]) and exploiting modification direction (EMD) [5,6]. The DE methods embed secret data by enlarging the difference of two adjacent pixel values, and HS based methods embed secret data into the shifted histograms. Another commonly used RDH method is called prediction-error expansion (PEE) [7,8]. It exploits the prediction-error to embed the secret data using the HS technique.

Nowadays, many researchers show their interest in developing reversible data hiding methods in encrypted images (RDHEI), where the original image is first encrypted by the content provider, and then the data hider embeds secret data into the encrypted image without knowing the original image content. At the receiver side, RDHEI aims to completely recover both the secret data and original image. With the reversibility property, RDHEI has wide applications in many critical scenarios such as medical image sharing, law forensics, Cloud storage and military applications. If users want to store their images into the Cloud but do not want any unauthorized access, they can encrypt the images before sending them to the Cloud. Although the Cloud does not know the image contents, it is able to add some additional information to the encrypted images for the management of the resources (e.g., add notations or location information to the encrypted images). In this scenario, no transmission is involved, and thus no errors or attacks either [9].

Existing RDHEI methods can be divided into two categories: one is called vacating room after encryption (VRAE), the other is reserving room before encryption (RRBE) [10]. In the VRAE methods, the content provider needs do nothing but encrypt the original image. Many VRAE methods have been proposed in recent years [11–19]. They use the stream cipher [12–15,20–22], permutation [23–26] or Paillier cryptosystem [27] to encrypt the original image, and embed secret data by bits flipping [20–22], compression [28,29], HS [23–26,30] et al. However, vacating room for data embedding after image encryption

[*] Corresponding author.
E-mail address: yicongzhou@umac.mo (Y. Zhou).

https://doi.org/10.1016/j.image.2018.03.001
Received 3 November 2017; Received in revised form 26 January 2018; Accepted 2 March 2018
Available online 8 March 2018
0923-5965/

may be difficult and inefficient. In order to increase the embedding rate, Ma et al. [10] proposed a RRBE method to reserve room from the original image before image encryption. Secret data then can be embedded into the reserved spare space directly. Later, some RRBE methods have been proposed by reserving the spare space using different techniques such as traditional RDH methods [10,31–33] and sparse representation technique [34].

Although RRBE can achieve a relatively high embedding rate, the content owner is required to perform an extra operation of reserving a spare space before encrypting the original image. The content owner may have difficulty to accomplish this extra operation to vacate rooms and/or may have no idea about the forthcoming secret data to be embedded. Therefore, many researchers show interest in developing VRAE methods. Existing VRAE methods for RDHEI can be divided into two categories, namely separable VRAE (S-VRAE) and joint (J-VRAE) methods. The former one can perform data extraction and image recovery separately, while the latter one cannot. Due to the difficulty of vacating room from the encrypted image, the VRAE methods are limited in the embedding rate. In addition, most VRAE methods extract the secret data and recover the original image by utilizing the fluctuation measure function to evaluate the smoothness of the decrypted image. They may fail to obtain the precise measure results when the images contain many textures. This leads to incorrect data extraction and/or partial image recovery. To address these problems, this paper proposes an adaptive block-level prediction-error expansion (ABPEE) to perform RDHEI. The main contributions of this work are summarized as follows:

(1) We propose a new block-level predictor (BLP) to predict the pixel values within a 2 × 2 image block. It can obtain more precise prediction results than the commonly used predictor like median edge detector (MED) [35].

(2) Using BLP, we introduce a block-level prediction-error expansion (BPEE) method to embed secret data into an image. It embeds secret data into image block by block rather than the conventional PEE that embeds data pixel by pixel in the raster-scan order.

(3) We further propose an Adaptive BPEE based RDHEI (ABPEE-RDHEI) scheme.

(a) It inherits the merits of VRAE that reserving room process is not required at the content-owner side. In addition, data extraction and image recovery can be performed separately and independently.

(b) It is fully reversible. This means that the secret data and original image can be recovered without any error.

(c) It can achieve a high embedding rate, as iterative embedding is used.

(d) It can obtain a larger embedding rate and generate marked decrypted images with higher visual quality than several state-of-the-art methods.

The rest of this paper is organized as follows: Section 2 briefly reviews some related works. Section 3 introduces BPEE. Section 4 presents the proposed ABPEE-RDHEI. Section 5 discusses several characteristics of ABPEE-RDHEI. Section 6 provides simulation results and comparisons with the state-of-the-art methods. Finally, Section 7 concludes this paper.

2. Related works

In this section, we review some commonly used predictors in image processing and some existing VRAE methods for RDHEI.

2.1. Existing predictors

Recently, many predictors have been proposed to estimate pixel values in spatial domain, such as MED [35], gradient adjusted predictor (GAP) [36], simplified gradient adjusted predictor (SGAP) [37], partial differential equations (PDE) predictor [38] and checkerboard based prediction (CBP) [39]. GAP and SGAP use the pixels in a half-surrounded structure with an irregular shape to predict the target pixel. They are suitable for predicting pixels in a raster-scan order. PDE completes the prediction process using the predefined reference pixels. CBP uses 25% pixels in a host image to predict the remaining 75% pixels, and each target pixel is predicted by pixels with a 3 × 3 region. MED uses pixels within a 2 × 2 image block to predict the target pixel, and it is described as follows:

$$\hat{x} = \begin{cases} \min(x_r, x_c), & \text{if } x_d \geqslant \max(x_r, x_c) \\ \max(x_r, x_c), & \text{if } x_d \leqslant \min(x_r, x_c) \\ x_r + x_c - x_d, & \text{otherwise} \end{cases} \tag{1}$$

where (x, x_r, x_c, x_d) denotes the pixels within a 2 × 2 image block, x is the target pixel, x_r, x_c and x_d are the three remaining pixels located in the same row, column, and diagonal directions with x, respectively.

2.2. VRAE methods

Here, we review two types of VRAE methods, J-VRAE and S-VRAE, separately.

2.2.1. J-VRAE methods

The J-VRAE methods usually first encrypt the original image using a stream cipher, and then embed the secret data using different techniques such as least significant bits (LSBs) flipping [12–15,20–22] and public key modulation [16]. Zhang's method [20] divides the encrypted image into a number of non-overlapped blocks, separates pixels in each block into two groups, namely S_0 and S_1, and embeds one bit of secret data into one image block by flipping the 3 LSBs of S_0 (if secret data bit is 0) or S_1 (if secret data bit is 1). Data extraction and image recovery are accomplished by comparing the smoothness of the decrypted image blocks. Yu et al. [12] proposed an improved version of Zhang's method [20] by randomly selecting p% ($p \in (0, 100]$) of pixels in the encrypted image as the active pixels, and flipping the 3 LSBs of all active pixels in S_0 or S_1 of an image block to embed different values of secret data bits. Thus, Zhang's method [20] is a special case of Yu et al.'s method [12] with $p = 100$. Hong et al. [21] and Liao et al.'s [22] methods use a more precise fluctuation measure function to calculate the complexity of the image block, and apply the side match technique to reduce the rate of incorrect data extraction. Wu et al. [13] embed one bit of the secret data into a group of pixels by flipping their ith$(1 \leqslant i \leqslant 6)$ least significant bits (LSBs). In Li et al.'s method [14], one bit of the secret data is embedded into a pre-selected pixel by flipping its 3 LSBs. In addition, it duplicates the secret data before embedding to reduce the rate of incorrect data extraction. Method in [15] flips only the LSBs of fewer pixels by the elaborate selection, so that visual quality of the marked decrypted image can be improved. Meanwhile, the adaptive judging function based on the distribution characteristics of image local contents effectively decreases the error rate of extracted secret data bits. Zhou et al. [16] use a public key modulation method to embed n $(n \geqslant 1)$ bits of secret data into an image block. The support vector machine technique is utilized for data extraction and image recovery.

2.2.2. S-VRAE methods

In order to perform data extraction and image recovery separately, a number of S-VRAE methods have been proposed. Wu et al. [13] embed the secret data by replacing the ith$(i \geqslant 7)$ bit of a stream-cipher-encrypted pixel. Methods in [28,29] and [18] encrypt the original image using the stream cipher. In [29] and [28], a number of LSB planes and the 4th bit plane of a stream-cipher-encrypted image are compressed to accommodate secret data, respectively. Zhang et al. [18] use the pseudo random sequence modulation technique to embed secret data into 3 LSB planes of the encrypted image. Previous S-VRAE methods [11,13,18,28,29] can perfectly extract the secret data without any error but the recovered image may have data loss. This is because image recovery is accomplished by analyzing the local standard