Signal Processing 104 (2014) 387–400

Contents lists available at ScienceDirect

Signal Processing

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

High-capacity reversible data hiding in encrypted images by prediction error

Xiaotian Wu [a], Wei Sun [b,c,*]

[a] School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510006, China
[b] School of Software, Sun Yat-sen University, Guangzhou 510006, China
[c] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China

ARTICLE INFO

Article history:
Received 30 December 2013
Received in revised form
25 April 2014
Accepted 29 April 2014
Available online 9 May 2014

Keywords:
Privacy preserving
Reversible data hiding
Image encryption
Prediction error

ABSTRACT

In recent years, signal processing in encrypted images received much attention from academia due to the privacy preserving property. Reversible data hiding in encrypted images is a technique that embedded additional data into an encrypted image without accessing the content of the original image, the embedded data can be extracted and the encrypted image can be recovered to the original one. In this paper, two reversible data hiding methods in encrypted images, namely a joint method and a separable method, are introduced by adopting prediction error. In the joint method, data extraction and image reconstruction are performed at the same time. The reversibility, number of incorrect extracted bits are significantly improved while maintaining good visual quality of recovered image, especially when embedding rate is high. In the separable method, data extraction and image recovery are separated. The separable method also provides improved reversibility and good visual quality of recovered image for high payload embedding.

1. Introduction

Reversible data hiding in images is a methodology that embeds additional messages into some distortion-unacceptable covers, such as medical, military or law forensic images, with a reversible manner that the original covers can be losslessly recovered after the embedded messages are extracted. In recent years, many reversible data hiding approaches have been proposed. Tian [1] introduced a difference expansion method, where two pixels are used as a group and a bit is embedded into each group by expanding the pixel difference. Ni et al. [2] exploited the image histogram and concealed the secret data by shifting the histogram. Tai et al. [3] proposed an efficient reversible data hiding method by shifting the histogram with the assistance of a binary tree. More investigations on reversible data hiding can be found in [4–11].

Signal processing in encrypted domain has attracted considerable research interest. With regard to providing confidentiality for images, encryption is an effective and popular means for a content owner to convert the original and meaningful content to incomprehensible one. However, in some scenarios a content owner does not trust the processing service provider, and does not want the service provider to access the content of the original image. The content owner may encrypt the image before transmission. The service provider would embed some additional messages within the encrypted image for other purposes such as image notation or authentication.

Recently, some methods on reversible data hiding in encrypted images have been proposed. Those proposed methods can be classified into two categories: joint methods

* Corresponding author at: School of Software, Sun Yat-sen University, Guangzhou 510006, China.
E-mail addresses: wxt.sysu@gmail.com (X. Wu), sunwei@sysu.edu.cn (W. Sun).

http://dx.doi.org/10.1016/j.sigpro.2014.04.032
0165-1684/

and separable methods. For those joint methods, data extraction and image recovery are performed jointly. In [12], one additional bit is embedded into an associated block of cipher-text image encrypted by AES algorithm. Data extraction and image recovery are performed based on the analysis of local standard deviation of the marked encrypted image. But the embedding payload is low, and the image decrypted directly from the marked encrypted image is seriously degraded. Later, Zhang [13] introduced a method of reversible data hiding in encrypted images by modifying the least significant bits (LSBs) of the encrypted image. More exactly, a content owner encrypts the original image using XOR operation, and then a service provider partitions the encrypted image into blocks of the same size. Each block is separated into two disjoint sets. According to the embedded data, the 3 LSBs of one set are flipped. A receiver decrypts the image by using XOR operation, and uses the block smoothness to extract the embedded bits and to recover the original block. The embedding payload increases and the high fidelity of directly decrypted image is preserved. However, the probability of correctly retrieving the embedded bits and recovering the image significantly decreases when high embedding payload is adopted. Hong et al. [14] improved Zhang's method by using side match technique and a better metric for measuring the block smoothness. The capability of extracting correct embedded data and reconstructing the image is further enhanced. Another improved joint method [15] is proposed by adopting a pseudo-random sequence modulation mechanism. The additional bits are embedded by modifying the LSBs of some encrypted pixels which are determined by the pseudo-random sequences. Main advantage of this method is that the reversibility for obtaining correct extracted bits and recovering the original image is improved, when a small number of bits are embedded. In all, those reported joint methods are not capable of obtaining error-free extracted bits when high payload embedding is used.

For the second type of methods, data extraction and image decryption are separable so that perfectly extracting the embedded bits is guaranteed. Zhang [16] proposed a separable method, where some encrypted data are firstly compressed, and space for data embedding is emptied out. A receiver having the data hiding key can extract the additional data with any error, while a receiver having the encryption key can decrypt received data to obtain an image similar to the original one. If both the data hiding and encryption keys are available, the receiver can retrieve the additional data and recover the original image. Zhang's method [16] guarantees an error-free data extraction, but it is not suitable for high payload embedding. For providing better rate-distortion performance, an efficient method using low-density parity-check codes and side information is given in [17]. To obtain an error-free recovered image, Ma et al. [18] introduced a reversible data hiding methodology for encrypted images by reserving room before encryption with a conventional reversible data hiding algorithm, where the reserved room is used to accommodate the additional data. In their method, data extraction and image reconstruction are free of any error. Zhang et al. [19] proposed a reversibility improved method. Prior to encrypting the image, room for data embedding is vacated by shifting the histogram of estimating errors of some pixels. Data retrieving and image recovery in their method are error-free. Although these two methods improve the embedding capacity and the reversibility significantly, empty out room for data embedding by the content owner might be impossible, because reversible data hiding in encrypted image always requires the content owner to do nothing except image encryption, and data embedding is supposed to be accomplished by the service provider. Qian et al. [20] introduced a separable reversible data hiding approach for encrypted images using a histogram modification and n-nary data hiding method. Reserving room is no longer required at the content owner's side, and an error-free recovered image is obtained by their method.

For the mentioned methods such as [13,14,16], the reversibility, number of incorrect extracted bits and visual quality of lossy recovered image are not satisfactory when high payload embedding is carried out. In this paper, two reversible data hiding methods for encrypted images based on prediction error are introduced to improve the mentioned problems. Both the methods can provide improved reversibility and better visual quality of lossy recovered image. Further, the number of incorrect extracted bits is significantly reduced by the proposed joint method.

The remaining part of this paper is organized as follows. The joint method of reversible data hiding in encrypted images is described in Section 2. Section 3 introduces the separable method. Experimental results and discussions are provided in Section 4. Section 5 gives some concluding remarks and future work.

2. The joint method

The first method consists of three phases: image encryption phase, data hiding phase and joint data extraction and image reconstruction phase, as depicted in Fig. 1. In image encryption phase, a content owner encrypts an original uncompressed image by using an encryption key, and produces an encrypted version of the original image. In data hiding phase, a service provider (also called data-hider) embeds some additional data within the encrypted image by utilizing a data hiding key. Note that, the service provider does not know any information about the original image. In joint data extraction and image reconstruction phase, a receiver can decrypt the marked encrypted image by the encryption key, and obtain a directly decrypted image which is similar to the original one. Further, the receiver can convert the directly decrypted image to the original version and extract the embedded data with the aid of data hiding key.

2.1. Image encryption phase

The original uncompressed image C with $M \times N$ pixels is assumed to be gray level, and each gray level is denoted by 8 bits. Let (i,j) be the pixel location, and let $C(i,j)$ be the associated gray value, where $C(i,j) \in [0,255], 1 \le i \le M$, $1 \le j \le N$. The original image C is decomposed into 8 bit