Displays 39 (2015) 109–116

Contents lists available at ScienceDirect

Displays

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

A reversible data hiding scheme based on the Sudoku technique

Thai-Son Nguyen [a,b], Chin-Chen Chang [a,*]

[a] Department of Information Engineering and Computer Science, Feng Chia University, Taichung 40724, Taiwan, ROC
[b] Department of Information Technology, Tra Vinh University, Tra Vinh Province, Viet Nam

ARTICLE INFO

Article history:
Received 14 July 2015
Received in revised form 5 October 2015
Accepted 6 October 2015
Available online 9 October 2015

Keywords:
Reversible data hiding
Steganography
Sudoku
High visual quality

ABSTRACT

Data hiding, also known as information hiding, plays an important role in information security for various purposes. Reversible data hiding is a technique that allows distortion-free recovery of both the cover image and the secret information. In this paper, we propose a new, reversible data hiding scheme that is based on the Sudoku technique and can achieve higher embedding capacity. The proposed scheme allows embedding more secret bits into a pair of pixels while guaranteeing the good quality of the stego-image. The experimental results showed that the proposed scheme obtained higher embedding capacity than some other previous schemes. In addition, our proposed scheme maintained the good visual quality of the stego-image (i.e., PSNR > 46 dB), which outperforms some existing schemes.

1. Introduction

Data hiding, also known as information hiding, is important in information security for various purposes, such as copyright protection and content authentication. The main idea of data hiding is to embed the secret data into the cover media, i.e., images, audio, video or text, to avoid attracting the attention of malicious attackers in the Internet, which is a public channel. In recent years, many data hiding schemes have been proposed, and most of them have been irreversible data hiding schemes [1–3]. This means that after the secret data have been extracted from the stego-image, the cover image is distorted permanently and cannot be restored correctly. However, some special fields, such as the military and medical fields, require that the cover image be restored to its original condition after the secret data have been extracted. As a result, various reversible data hiding schemes [4–7,9–13,16,20] have been proposed to solve this issue. These schemes allow the correct extraction of the secret information and the reconstruction of the cover image without any distortion.

Most reversible data hiding schemes are based on difference expansion (DE) [4,5,9,12,13,18] and histogram shifting [6,7,10,11,19]. In 2003, Tian introduced a reversible DE-based data hiding scheme [4]. In this scheme, the difference between two neighboring pixels is calculated. Then, the difference value is doubled to embed one secret bit. In 2007, Thodi and Rodriguez [5] proposed an expansion scheme based on error prediction (PE) to hide the secret data. In this scheme, a prediction technique is used to predict the pixel value. Then, the difference between the pixel value and its predicted value is calculated to embed the secret data. The above two schemes [4,5] are based on the DE technique to obtain high embedding capacity and good image quality. However, in DE-based schemes, the pixels may have an overflow or underflow problem. To achieve better visual quality, many researchers have proposed histogram-shifting schemes. In 2006, Ni et al. proposed the first histogram-shifting scheme [6]. In Ni et al.'s scheme, most of the pixels are modified by one grayscale value to embed secret data. Their scheme achieved high visual quality of the stego-image, but the embedding capacity is limited. In 2009, Tai et al. [11] proposed a new reversible data hiding scheme based on histogram modification. In Tai et al.'s scheme, a binary tree structure is designed to determine the peak point, which is used for embedding data. To extract the secret data, the level of the binary tree must be sent to the receiver. Also in 2009, Kim et al. [7] introduced a reversible data hiding scheme based on the different histogram shifting to obtain high capacity and imperceptible embedding by dividing the cover image into several sub-images. The difference values between the sub-sampled images are calculated. Then, the difference values are shifted to embed more secret data. To improve Kim et al.'s scheme, Luo et al. [9] proposed a reversible data hiding scheme based on selecting the median pixel of each block to structure the reference sub-image. They separated the image blocks into four categories, and, for each category, one corresponding embedding strategy is used to hide the secret message. However, the embedding capacity of their scheme also is limited because the reference pixels are not used for embedding secret data. To obtain better embedding

* Corresponding author.
E-mail addresses: thaison@tvu.edu.vn (T.-S. Nguyen), alan3c@gmail.com (C.-C. Chang).

http://dx.doi.org/10.1016/j.displa.2015.10.003
0141-9382/

capacity, Lou et al. [13] designed a novel technique by using a multiple-layer embedding technique and a logarithm transformation function to hide secret data into the reduced difference expansion. In 2010, Li et al. [10] proposed a reversible data hiding scheme based on the difference between adjacent pixels to embed the secret data. In Li et al.'s scheme, the difference sequence of pixels is explored for concealing the secret data. In order to increase the hiding capacity and maintain good image quality, Qin et al. introduced a new scheme in 2012 [12] that exploited the relationship between the prediction error and the threshold to decide whether the current pixel is embeddable or not. To provide a general framework for histogram shifting-based reversible data hiding, Li et al. [19] designed two simple functions, i.e., shifting and embedding functions, for reversible data hiding. Then, based on these two functions, two different reversible data hiding schemes are introduced. These two schemes are expected to further improve the performance of the previous works. Their scheme I obtained high embedding capacity. However, their scheme II only worked well for low embedding capacity. Instead of using histogram shifting for embedding data as was done in [19], Gui et al. [18] used prediction-error expansion for reversible data hiding. In Gui et al.'s scheme, the complexity measurement is divided into various levels, and the size of the embedded data is calculated. Then, the smoother regions are used to embed more secret bits. However, the quality of the stego images in this scheme is unsatisfactory. The average PSNR of their scheme is always less than 40 dB for the embedding rate of 1 bit per pixel (bpp).

In this paper, to achieve high embedding capacity and good quality of the stego-image, we proposed a new, reversible data hiding scheme based on a reference matrix, which is constructed by using the Sudoku technique. According to Sudoku's properties, the proposed scheme can embed the secret digit in the base-9 numerical system into a pair of pixels. Our experimental results indicated that the proposed scheme successfully achieved the goals of improving embedding capacity and maintaining the good quality of the stego-image.

The rest of the paper is organized as follows. Section 2 introduces the Sudoku technique and the reference matrix that provide the basis for our scheme. Section 3 provides a description of the proposed data hiding scheme. Section 4 provides the experimental results, and our conclusions are presented in Section 5.

2. Related work

In 2008, Chang et al. [15] proposed an information hiding scheme based on the properties of Sudoku solutions. Sudoku is a logic-based, combinatorial, number placement puzzle [14]. Sudoku is a 9×9 matrix that contains 3×3 sub-matrices. Each sub-matrix contains a different value from 1 to 9. In addition, the same integer digit, from 1 to 9, cannot appear twice in a row or a column of the Sudoku grid. Fig. 1 shows an example of a Sudoku solution.

5	3	4	6	7	8	9	1	2
6	7	2	1	9	5	3	4	8
1	9	8	3	4	2	5	6	7
8	5	9	7	6	1	4	2	3
4	2	6	8	5	3	7	9	1
7	1	3	9	2	4	8	5	6
9	6	1	5	3	7	2	8	4
2	8	7	4	1	9	6	3	5
3	4	5	2	8	6	1	7	9

Fig. 1. Example Sudoku solution.

According to the properties of Sudoku solution, Chang et al. specially designed a reference matrix *RM* to embed secret digits in the base-9 numeral system into each pixel pair of the grayscale cover image. Hence, every digit in the Sudoku grid is reduced by one to generate the sub-matrix *S*, which contains digits from 0 to 8, as shown in Fig. 2. Then, the sub-matrix *S* is used to construct the reference matrix *RM*. Fig. 3 shows an example of the reference matrix *RM*, which is used for embedding and extracting secret data.

3. The proposed scheme

After carefully observing Chang et al.'s scheme [15], we found that it was based on Sudoku's properties to obtain high embedding capacity and good image quality. The properties of the Sudoku technique are the key features that allow the achievement of high embedding capacity and good image quality at the same time. However, the disadvantage of Chang et al.'s scheme is that the cover image is damaged permanently and cannot be recovered after the secret information has been extracted from it. In order to solve this problem, in this section, we present our design of a new, reversible data embedding scheme based on the properties of Sudoku technique to obtain the goals of improving security, increasing embedding capacity, and, especially, obtaining reversibility while guaranteeing good visual quality of the stego-image. Our reversible data hiding scheme is divided into two phases, i.e., the data embedding phase and the data extracting phase, which are described in Subsections 3.1 and 3.2, respectively.

3.1. Data embedding phase

In this section, the pixels of the cover image are first paired by using a pairing technique, and each pair of pixels is used to embed a secret digit. To extract the embedded secret digits and restore the original cover image, the decoder must determine which pair of pixels was selected for embedding the secret digit. To facilitate this process, location information must be embedded so the decoder can use it for extracting the secret data. To accomplish this, we generated a 2-D binary location map *LM*, which contains the location information of all embedded pairs. The location map *LM* is initialized with "0."

The cover image *I* is divided into two areas, i.e., the embeddable area and the non-embeddable area, which are shown in Fig. 4. The least significant bits (LSBs) of the pixels in the non-embeddable area are used to record the information of the location map *LM*. Therefore, they must be extracted and concatenated into the secret data *B* to generate the embedded data, *S*.

For use in our proposed scheme, the reference matrix *RM* must be prepared as shown in Fig. 3, and the embedded data *S* must be converted into based-9 numeral system digits. Let us denote the converted secret digits $D = \{d_1, d_2, \ldots, d_N\}$, where N is the total

4	2	3	5	6	7	8	0	1
5	6	1	0	8	4	2	3	7
0	8	7	2	3	1	4	5	6
7	4	8	6	5	0	3	1	2
3	1	5	7	4	2	6	8	0
6	0	2	8	1	3	7	4	5
8	5	0	4	2	6	1	7	3
1	7	6	3	0	8	5	2	4
2	3	4	1	7	5	0	6	8

Fig. 2. Sub-matrix *S*.