Expert Systems with Applications 41 (2014) 6123–6130

Contents lists available at ScienceDirect

Expert Systems with Applications

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

A novel image steganography scheme with high embedding capacity and tunable visual image quality based on a genetic algorithm

Hamidreza Rashidy Kanan *, Bahram Nazeri

Department of Electrical, Computer and IT Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran

ARTICLE INFO

Keywords:
Steganography
Tunable image steganography
Spatial domain
Genetic algorithm

ABSTRACT

Steganography is knowledge and art of hiding secret data into information which is largely used in information security systems. Various methods have been proposed in the literature which most of them are not capable of both preventing visual degradation and providing a large embedding capacity. In this paper, we propose a tunable visual image quality and data lossless method in spatial domain based on a genetic algorithm (GA). The main idea of the proposed technique is modeling the steganography problem as a search and optimization problem. Experimental results, in comparison with other currently popular steganography techniques, demonstrate that the proposed algorithm not only achieves high embedding capacity but also enhances the PSNR of the stego image.

1. Introduction

Steganography is the science and art of hiding secret data in a host media (e.g. text, image, audio, video, etc.) (Cheddad et al., 2010). The purpose of a steganography algorithm is hiding a large amount of secret data into a meaningful host media such that the embedded secret data are concealed to prevent the attack of unauthorized persons. In steganography, there are three major goals including increasing hiding capacity, robustness to certain attacks and increasing security level (Cheddad et al., 2010). Image steganography (Morkel et al., 2005), where secret data is embedded within an image has been widely studied during the last decade due to the cost decreasing of image storage and communication and also the weaknesses of the human visual system (HVS). It should be mentioned that the cover or host image is referred as the original image without the embedded secret message, while the image that is obtained by *embedding secret* message into cover *image without* destroying the *cover image* is termed as stego image. The term payload is utilized to describe the size of the secret message that can be embedded in a specific image.

There are different steganography approaches including spatial domain (direct manipulation of image intensities) (Carvajal-Gamez, Gallegos-Funes, & Rosales-Silva, 2013; Chan & Cheng, 2004; Chen, Chang, & Le, 2010; Ioannidou, Halkidis, & Stephanides, 2012; Naor & Shamir, 1995; Sajedi & Jamzad, 2010; Shamir, 1979; Wu & Tsai, 2003; Yang, Cheng-Hsing et al., 2008), frequency domain (manipulates the image indirectly through various transforms like DFT, DCT, DWT, etc.) (Barni & Mauro, 1999; Chen, 2008; Chu et al., 2004; Jafari, Ziou, & Rashidi, 2013; Liu & Qiu, 2002; Noda, Niimi, & Kawaguchi, 2006) and the compression (substitution) domain (Chang, 2007; Chang, Nguyen, & Lin, 2011; Chang, Tai, & Lin, 2006; Chang, Wu, & Hu, 2007; Chung, Shen, & Chang, 2001; Yang, 2011). Each approach has different specifications. For instance, the spatial domain algorithms usually offer large hiding capacity for secret data and good visual quality for stego-images, but may not pass statistical steganalysis (Duric, Jacobs, & Jajodia, 2005; Luo et al., 2008; Nissar & Mir, 2010; Ziou & Jafari, 2012). On the other hand, the compression domain techniques perform better in statistical steganalysis, however may create less embedding capacity for secret data and lower visual quality for stego-images (Chang et al., 2006; Yang, 2011). The frequency domain methods are usually utilized in watermarking applications due to their good robustness against image distortion attacks.

Though numerous methods have been proposed for image steganography, limited studies have been done on metaheuristic-based image steganography and these papers could not present logical reasons for advantage of their methods. For example, (Fard et al., 2006) proposed secure jpeg steganography method based on a genetic algorithm (Goldberg, 1989). This method is based on OutGuess which is proved to be the least vulnerable steganographic system. Tseng et al. (2008) proposed a steganography method based on Optimal Pixel Adjustment Process (OPAP) and

* Corresponding author. Tel.: +98 9122454561.
E-mail addresses: h.rashidykanan@qiau.ac.ir (H.R. Kanan), b.nazeri@qiau.ac.ir (B. Nazeri).

http://dx.doi.org/10.1016/j.eswa.2014.04.022
0957-4174/© 2014 Elsevier Ltd. All rights reserved.

GA. This method alters secret bits for achieving more compatibility with host image. Steganography using session based stego-key and GA has been presented in Bhattacharya et al. (2008). This technique is similar to other methods that use encrypted secret data. In fact, this method encrypts secret data in two steps. First step is similar to other methods, but in second step encryption is optimized by GA. In Wang, Yang, and Niu (2010), a steganography method based on GA is presented which is secure against RS attacks. This method in first step embeds secret bits into host image just like simple LSB and in second step alters pixel values to make stego image RS parameters sit in secure area. Ghasemi and Shanbehzadeh (2010) proposed a steganography method based on GA which divides host image into sub blocks and find best pixel sequence in blocks for embedding.

In this paper, we try to find best place for embedding modified secret data in host image to achieve high level of security. The process of embedding is accomplished in two main steps, first to modify secret bits and second to embed it into host image. Different places in host image defined by order of scanning host pixels and starting point of scanning and best LSBs of each pixel. Other options of host bits are defined too. The genetic algorithm which was developed by Goldberg (1989) is utilized to find the best starting point, scanning order and other options such that the PSNR of the stego-image maximized. A feasibility and effectiveness investigation for the proposed method is conducted using some benchmark images. The system performance is compared with the performances of some previously popular existing approaches. Obtained results indicate that our proposed steganography algorithm is superior and reliable.

The rest of the paper is organized as follows: Section 2 presents the main idea of the proposed approach and its capabilities in details. The experimental results and discussion are presented in Section 3. Finally, the paper concludes in Section 4.

2. The proposed steganography method

For introducing the main idea, it is necessary to explain some preliminary definitions. In this section raster order will be explained then the proposed method will be presented in two phases.

2.1. Raster order

In LSB substitution method, host pixels are scanned row by row and trough first row to last one while in each row, pixels are scanned from left to right. This order of pixels is known as raster order. For example if image dimension is 5 × 5, raster order refers to Fig. 1.

2.2. The main idea of the proposed method

The main idea of the proposed scheme is modeling the steganography problem as a search and optimization problem. In other words, for a secret-host image pair, order of Fig. 2 maybe better than raster order or for another one, order of Fig. 3 maybe better. Accordingly, there are different orders and different places in host image for hiding secret image which lead to different PSNRs. The mentioned problem which is the direction of pixel scanning has 16 possible solutions. Therefore, if a mechanism is designed for testing all possible orders to find the best order for host and secret images, the result can be improved from simple LSB.

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Fig. 1. Raster order.

25	24	23	22	21
20	19	18	17	16
15	14	13	12	11
10	9	8	7	6
5	4	3	2	1

Fig. 2. Pixel scanning order 1.

21	16	11	6	1
22	17	12	7	2
23	18	13	8	3
24	19	14	9	4
25	20	15	10	5

Fig. 3. Pixel scanning order 2.

2.3. Development of the main idea

Another option which can be considered through the development of the main idea is the starting point. In other words, in raster order, if we choose another starting point instead of first column and first row for a secret-host image pair, we may be able to obtain better results. For example, Fig. 4 shows raster order with different starting points. For a secret-host image pair, different starting points results different PSNRs and there is no guarantee that the default starting point be the best.

In this paper, steganography is modeled as a search problem in which the purpose is finding the best direction and the best starting point in host image for hiding secret data such that the PSNR of the stego-image maximized. In order to search this space, a genetic algorithm is utilized and the PSNR of the stego-image is considered as a fitness function. In the following the details of the proposed algorithm will be presented.

2.3.1. Chromosome representation

In the utilized genetic algorithm, the proposed chromosome contains 7 genes which are indicated in Table 1.

In the defined chromosome, since the direction of pixel scanning has 16 possible states, so we represented it as a gene with 4 bits length. Starting point is represented as two genes including *X*-offset and *Y*-offset with 8 bits length for each of them. Bit-Planes utilized for determining LSB planes in host pixels which are used for embedding secret data in host pixels. Possible values for Bit-Planes are shown in Table 2.

SB-Pole used to determine secret Bits-Pole, SB-Dire used to determine direction of secret bits and the last gene is BP-Dire