Journal of Engineering and Applied Sciences
MEDWELL PUBLICATIONS

Analysis and Implementation of Steganography on Tiff Image using Bit-Plane Complexity Segmentation (BPCS) and Spread Spectrum Method

Latifah Uswatun Hasanah, Tito Waluyo Purboyo and Randy Erfa Saputra
Department of Computer Engineering, Faculty of Electrical Engineering, Telkom University, Bandung, Indonesia

Key words: Steganography, TIFF, BPCS, Spread Spectrum, MSE

Corresponding Author:
Latifah Uswatun Hasanah
Department of Computer Engineering, Faculty of Electrical Engineering, Telkom University, Bandung, Indonesia

Page No.: 2529-2535
Volume: 15, Issue 12, 2020
ISSN: 1816-949x
Journal of Engineering and Applied Sciences

Abstract: Nowadays, with the development of internet media and applications, that use the Internet is increasing also a crime in information systems. Information security becomes very important in human life as a social being. Much confidential information such as personal data, financial data and even state secrets that need to be protected, so that, it cannot be misused by others who are not authorized to do so. One technique to maintain information security is by using steganography technique. Steganography is a technique where messages can be hidden by certain methods. Steganography on the image is the development of science from steganography. In this final project will explain about the development of steganography application on TIFF (Tagged Image File Format) digital image using Bit-Plane Complexity Segmentation (BPCS) method which utilizes the human vision characteristic that cannot see the change of binary pattern that happened in the picture and spread spectrum in message randomization. The final part of this final project is to calculate the parameters commonly used as an indicator to measure the similarity of two images, namely the Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) value, resulting in the quality of the digital image file that has been inserted the message did not undergo any significant changes and the data residing in the digital image could be re-extracted.

INTRODUCTION

Today, the development of all-digital technology is growing so fast. Communication technology is one of them. Many messages are delivered in digital media. Digital communication technology is a computer-based electrical communication technique using the binary number system. Binary numbers will form the codes that represent a certain information. After going through the process of digitization, the incoming information will turn into a series of binary numbers that form information in the form of digital code. Digital messages can be text, images, audio or video. The security of digital messaging, especially confidential digital messaging, is necessary. With the continuous development of digital image processing, the security techniques for digital messages

can be solved. Digital secret messages can be inserted into a digital image using a technique called steganography[1].

MATERIALS AND METHODS

Substances and methods

Steganography: Steganography is a technique that allows concealment of information or data on various types of digital media. Steganography requires two properties of the container media (images, audio, video and text) and secret messages to be hidden. Secret message can be hidden by the sender unnoticed by anyone other than the recipient by using steganography. Steganography techniques commonly used for the benefit of communication and information on a particular company that is confidential. The goal of steganography is hiding communication for preventing third parties know about the existence of the message (1998). The general steganography of the system as: First, create a message on the cover media. The second, a stego-message is created by hiding a secret message placed on the cover via message using a stego-message and then the receiver gets the stego-message from the channel that not secure. Lastly, a pre-agreed stego-key is used for extracts secret message when receiver already receives the message[2].

Steganography has two main processes, namely embed or insertion and extraction, as shown in Fig. 1. The insertion process is a process of inserting (hidden object) or information or messages to be inserted, into a cover object to produce new files that have been inserted messages in it which called the stego-file. While the extraction process is the process of returning the hidden object as a whole after inserted into the cover object[3].

There are some important things to be kept in mind before applying or performing a procedure of steganographic: Embedding capacity where the data is hidden in large data volumes called cover or carriers. The embedded capacity is the amount of data that can be hidden or embedded on the cover and will be compared to the cover size because if the size of data that will be inserted on the cover is greater than the cover size then steganography can't be done. Undetectability where the data should be hidden in the carrier file in such a way that any confidential information can't be seen accidentally in the original file. If anyone detects the message in the original file, then the steganography is failed. Robustness, the ability of the embedding algorithm to store embedded data even after going through the compression and decompression process. Security, in most cases, security, including the perceptional transparency of the hidden data is considers the most important issue of hiding data in many different formats[3].

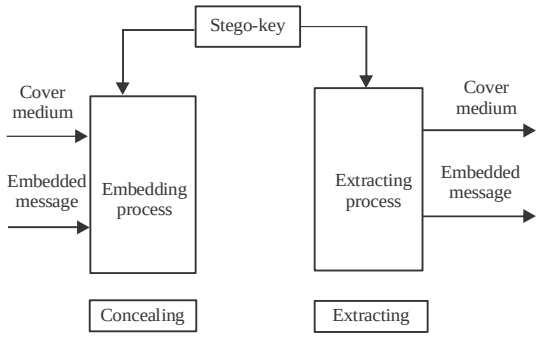

Fig. 1: Steganography process

Steganography in images: Hiding messages in pictures are the most commonly used technique today. The image with a secret message in it can be easily disseminated via the web or forum. The application of steganography techniques has been investigated by Niels Provos, a German Steganographer. The usual methods used to hide information in images are LSB, masking and filtering, and transformation on the cover image. This technique can be used with various degrees of success with different types of image files[4].

TIFF image: The TIFF format is the best image format with its meaning all data and information (RGB data, CMYK data and others) related to the manipulation of images that are not lost. The usual TIFF format is used for printing needs with very high image quality. Font size for format this is usually very large This format is capable of storing images with quality up to 32 bits. The TIFF file format can also be used for inter-platform streaming purposes (PC, Macintosh and Silicon Graphic). In addition, this format is easy to use for transfer between programs almost all programs are able to read bitmap file formats as well capable of reading TIFF format file TIFF format using LZW compression algorithm.

Steganography property: The following properties contained in the steganography are:

- Embedded message (hidden text) is the file to be hidden. (a text, video, audio or image)
- Cover (cover text) object is a file used as containers to hide the embedded message
- The stego-object (stego-text) is a file container that's been given message embedded message
- Key stego is the key that is used to insert the message and extracting a message from the stego-text[1]

Steganography methods: In this section will focus discusses the TIFF digital image, method of BPCS and spread spectrum.

BPCS method: In 1997, Eiji Kawaguchi and R.O Eason introduced the Bit-Plane Complexity Segmentation (BPCS) which is a steganography technique with image media. The BPCS takes an advantage of human vision which that can't see the change of the binary pattern that occur in the image. Implementing BPCS on TIFF-formatted image files will not cause damage to secret messages because TIFF uses lossless compression techniques[4].

In the BPCS method, image documents are divided into multiple segments, measuring 8×8 pixel per segment. pixels on each segment have an 8-bit plane that will represent the value of the pixel. The bit-slicing process is the process by which the 8×8 segment is converted into 8 pieces of bit-plane. The eighth bit-plane representation is a Pure Binary Code (PBC). PBC is the password used to represent each digit in a decimal number with its binary equivalent. The data insertion process is performed on segments that have high complexity. This high complexity segment is called the noise-like region.

In these segments, data will be hidden in the Most Significant Bit (MSB) and Least Significant Bit (LSB). That means it is done on all the bit-plane. The insertion process in bit-plane uses Canonical Gray Coding (CGC) system which produces better results than using a Pure Binary Code (PBC) system in bit-slicing process. CGC is a password with minimum changes which means each number is only 1 bit different from the previous bit[5].

In one pixel is 8-bit. One bit consists of a 1-bit plane. The plane '0' contains the Lowest Sequence Bit (LSB), while plane '7' contains high order bits (MSB). For example, suppose there is a P image with n-bit depth can be shown P = (P1, P2 ... Pn). Pi is the i-th plane bit with i = 1, 2, ..., n. If the image of P consists of 3 colour, red, green, blue, then P = (PR1, PR2, ..., PRn, PG1, PG2, ..., PGn, PB1, PB2, ..., PBn) with PRi is the i-bit of bit-plane for red, PGi is the i-bit of bit-plane for green, and PBi is the i-bit of bit-plane for blue.

Meanwhile, the complexity of binary imagery is a complexity parameter of a binary image. Discoloration black and white in binary images on each row and column horizontally (left to right) and vertical (top to bottom) is a good measure to calculate the value of complexity. If color changes occur many, then the image has a high level of complexity. If otherwise, then the image is a simple picture. The formula of complexity image:

$$\alpha = \frac{k}{2 \times 2^{n}}\left(2^{n-1}\right)$$

Where:
'k' = A B&W color change
'α' = The value of complexity

For a square binary image with a 2n×2n size, the maximum possible color change is 2*2n*(2n-1) and possibly minimum the color change is 0, obtained for all black images or all white.

The complexity in the bit-plane area is the parameter that used to determine the bit-plane is an informative region or noise-like region. This complexity parameter must have a forgotten boundary a second separator called the threshold (α0).

A bit-plane belongs to an informative region if has a smaller complexity value than the threshold value ($\alpha<\alpha 0$) and if it has a greater complexity value than the threshold value ($\alpha \geq \alpha 0$) would be regarded as a noise like region.

Conjugation of binary image (P) is a binary image that has a complexity value of one minus the value of complexity (P). The conjugation of image P is denoted by P*. To build a conjugation of P* of an image P can be found by using this following formula[5]:

$$P^{*} = P \oplus Wc$$
$$(P^{*})^{*} = P, P^{*} \neq P$$

If 'α (P)' is the complexity of P, then: 'α (P *) = 1- α (P)'. Steps taken on the BPCS algorithm when inserting data are as follows:

Cover-object with the PBC system converted into CGC system, then an image is a slice into bit-plane in the form of the binary image. Every bit-plane represents the bits of each pixel in the image.

Segmentation of each bit-plane on the cover-object becomes informative and noise-like region using the limit value/threshold (α0). The common value of threshold = 0.3.

The group of secret message bytes is converted into a series of secret message blocks. If the block (S) is less complex than the threshold value, then conjugate the block value (S) to obtain a more complex conjugate block (S*) value. the conjugate block (S*) obtained must be more complex than the threshold value.

Each block of secret messages will be inserted into a noise-like region (or replace all the bits in the noise-like region) in bit-plane. Also, paste the conjugation folder as it does on the secret message block. Change a stego-image of the CGC system into a PBC system[4].

Spread spectrum method: Spread spectrum is part of the realm transforms. A transmission technique with using the pseudo-noise code which is independent of data information as a form modulator wave to propagate the inner signal energy a larger communication path

(bandwidth) rather than an information, communication path signal. By the receiver, the signal is reassembled using replica pseudo-noise code synchronized. By definition, it can be said that steganography using a spread spectrum method creates the cover-object either as noise (noise) or as an attempt to add pseudo noise (pseudo-noise) into the cover-object. In the spread spectrum method, insertion message or information there is a key or key used to encrypt messages. The key is obtained from the pseudo number results of the calculation[6].

Parameters

Linear Congruential Generator (LCG): LCG is a pseudo number algorithm to get the key used as message encryption:

$$X_{n+1} = (aX_{n+b}) \bmod m$$

Where:

X_n = n random numbers from his series
X_{n+1} = Previous random numbers
a = Multiplier
b = The value increments
m = Modulus (a, b and m are constants)

Mean Square Error (MSE): Is an amount of the number of errors between the results of image processing with the original image. From the above formula has the following information. MSE values can be found from this equation[7]:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} f(i,j) - g(i,j)^2 \quad (4)$$

Where:

m = Number of lines on the cover image
n = Number of columns on the cover image
f(i, j) = Intensity of the cover image
g(i, j) = Intensity of the stego-image

Peak Signal to Noise Ratio (PSNR): The ratio between the maximum value measured by the magnitude of the error affecting the signal. PSNR is usually measured in decibels[7]:

$$PSNR = 10 \log 10\left(\frac{C^2MAX}{MSE}\right) \quad (5)$$

Where:

CMAX = The largest pixel value on the entire image

RESULTS AND DISCUSSION

Steganography using BPCS method:

171	133	172	127
82	0	91	73
71	129	47	61
97	143	110	152

In Table 1, shows the results of manual calculations as it is known the image size has been listed above and the number of characters that can be inserted messages.

In Fig. 2, the number of characters in messages that can be inserted on each pixel is shown on the y-axis while on the x-axis represents the resolution of an image When the pixels become larger, the message can be inserted more than usual. The CMYK (cyan, magenta, yellow, key) is a part of the coloring model often used in color printing. CMYK has 4 a color, that is Cyan (greenish blue), Magenta (mixed red and blue), Yellow and Key (black).

Steganography using spread spectrum method: This part will discuss about the results of the experiment using the spread spectrum method. The example uses 4×4 pixel image with a TIFF image format.

In Fig. 3, the picture above is a TIFF image format with 4×4 pixel size that has convert, for spread spectrum method experiment. A value of CMYK (in decimal form) convert to be a binary number. then afterwards the message can be inserted.

- Cyan value in decimals

187	107	181	125
60	36	94	79
46	125	40	79
87	153	75	135

- Magenta value in decimals; Yellow value in decimals

146	109	154	108
33	44	97	76
8	124	26	44
69	128	64	124

- Key (black) value in decimals

255	255	255	255
255	255	255	255
255	255	255	255
255	255	255	255

- All CMYK values are converted to binary

Table 1: Experimental results using the method of BPCS

Image size	The number of characters
4× 4 pixel	8
6×6 pixel	18
8×8 pixel	32

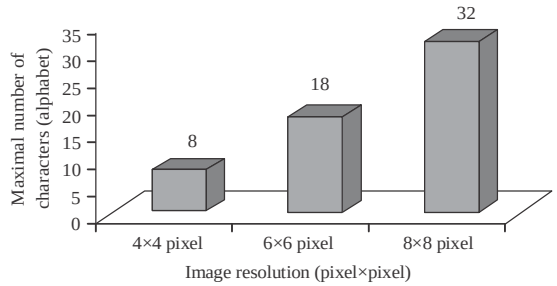

Fig. 2: Maximum characters that can be inserted into each pixel in CMYK

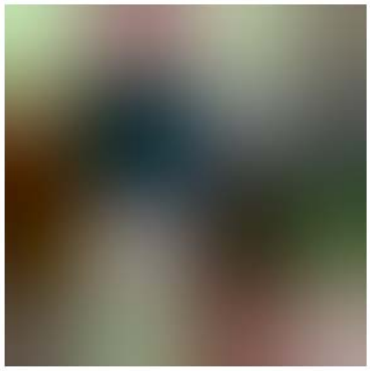

Fig. 3: TIFF image format

- The cyan value in binary

10101011	10000101	10101100	01111111
01010010	00000000	01011011	01001001
01000111	10000001	00101111	00111101
01100001	10001111	01101110	10011000

- The yellow value in binary

10101011	10000101	10101100	01111111
01010010	00000000	01011011	01001001
01000111	10000001	00101111	00111101
10010010	01101101	10011010	01101100

- The yellow value in binary

11111111	11111111	11111111	11111111
11111111	11111111	11111111	11111111
11111111	11111111	11111111	11111111
11111111	11111111	11111111	11111111

- The key (black) value in binary

10111011	01101011	10110101	01111101
00111100	00100100	01011110	01001111
00101110	01111101	00101000	01001111
01010111	10011001	01001011	10000111

Using the spread spectrum method requires 3 input data to start the process, that is a digital image file, the secret messages (text) and keys. The first step that must be done is to do the spreading process with the scalar quantity multiplier that is determined, the scale of the scalar multiplier is 4. For example, in this study has been experimented with segment of secret messages “TASK” (in the form of characters), the shape of the character we can see in the ASCII table, for each character in ASCII form (in decimal) is “84 65 83 75”. To start the insertion, the ASCII form must also be changed first into binary form, which results in "01010100 01000001 01010011 01001011". After spreading with scalar multiplier 4, the binary form generates a new message segment due to the doubling of bits 4-fold, the size of the message segment becomes 4-fold larger in size. Before doing the next step, we must specify the key first. The key to the experiment this time is “LUH”, just as with the secret message segment, the shape of the character must be converted to binary form, as follows “01001100 01010101 01001000”. The binary is of course, we get from each character in ASCII that is worth 81 (in decimal). The second step is to do the modulation process (the process of randomizing the message with a pseudo number taken from the key variable). This modulation process uses LCG algorithm as a calculation to get a pseudo number. Be discovered a = 2, b = 7 and mod = 15:

X1 = (2 x 81 + 7) mod 15 = 4 00000100
X2 = (2 x 4 + 7) mod 15 = 0 00000000
X3 = (2 x 0 + 7) mod 15 = 7 00000111
X4 = (2 x 7 + 7) mod 15 = 6 00000110

After getting the LCG value, then got also pseudo number “00000100 00000000 00000111 00000110”. The pseudo number length must be adjusted to the length of the message. If the message size is shorter than pseudo number then pseudo number size will be cut according to message size. If the message size is longer than the pseudo number then the pseudo number will be repeated until the length is equal to the message length. And then the message segment is modulated with a pseudo number using an XOR function. Message segment: message segment is spreading with the scalar of multiplier 4:

00001111 00001111 00001111 00000000
00001111 00000000 00000000 00001111
00001111 00001111 00000000 11111111
00001111 00000000 11110000 11111111 XOR

- Pseudo noise number: taken from the calculation of LCG

```
00000100  00000000  00000111  00000110
```

- XOR (modulation) result

```
00001111  00001111  00001000  00000110
00001111  00001111  00000111  00001001
00001111  00001111  00000111  11110110
00001111  00000000  11110111  11110110
```

The result of the modulation process will then be inserted into the digital image cover file.In this experiment, the message is inserted in the last digit of the bit and images that have been inserted are called stego-images.

- The cyan value after insertion

10101010	10000101	10101100	01111111
01010010	00000001	01011010	00100100
01000110	10000001	00101110	00111101
01100000	10001110	01101110	10011001

- The magenta value after insertion

10111010	01101011	10110100	01111101
00111100	00100100	01011110	01001111
00101110	01111101	00101000	01001111
01010110	10011001	01001010	10000110

- The yellow value after insertion

11111110	11111111	11111110	11111111
11111110	11111110	11111110	11111110
11111110	11111111	11111110	11111111
11111110	11111111	11111110	11111111

- The key (black) after insertion

10010010	01101101	10011010	01101101
00100000	00101100	01100000	01001101
00001000	01111101	00011010	00101101
01000100	10000001	01000000	01111100

So, that for the message to be read again, it must be extracted. The first step is to do the demodulation process.

- The process of demodulation becomes:

```
00001111  00001111  00001111  00000000
00001111  00000000  00000000  00001111
00001111  00001111  00000000  11111111
00001111  00000000  11110000  11111111
```

Table 2: PSNR and MSE values

Image size	MSE	PSNR
4×4 pixel	0,5625	53,1284
6× 6 pixel	0,16667	63,6938
8×8 pixel	0,09375	68,6913

And then, after process of demodulation,the next process is to divide the four results of demodulation, which is useful for shrinking the demodulation results into the original message. The result of process de-spreading the segment becomes: “01010100 01000001 01010011 01001011”.

Final result will be the same result after de-spreading “01010100 01000001 01010011 01001011” is the same message segment when hidden in the insertion process. The result is then changed to the character form will be “TASK”.

PSNR and MSE stego-image: In Table 2, overall PSNR and MSE data obtained from several dimensions of the image.

Contribution: In this experiment may contribute to the steganography technique in the drawing. In general, many steganography techniques are applied to image formats such as BMP, JPEG or PNG. While the application of steganography on the image with TIFF format quite rarely discussed and the discussion of experiment is explained in detail. Thus, this study may provide references or suggestions for better steganographic methods and in accordance with the TIFF image format to be applied. In this experiment, using the method of BPCS and spread spectrum on the application of TIFF image.

CONCLUSION

Some things that can be concluded from the experimental results in this paper are as follows: The threshold value for message insertion using the BPCS method can be set by the user for different storage capacities.

The larger the value will be little capacity provided to store the secret messages that will be hidden. the more capacity it has the more capacity it provides for storing secret messages.

TIFF digital imagery has a larger insertion message insertion capacity. The result of experiment and analysis using a spread spectrum method before and after dis. Messages are not insignificant significantly also undesirable by the invisible eye of others.

REFERENCES

01. Johnson, N.F. and S. Jajodia, 1998. Exploring steganography: Seeing the unseen. IEEE. Comput., 31: 26-34.

02. Provos, N. and P. Honeyman, 2003. Hide and seek: An introduction to steganography. IEEE Secur. Privacy, 1: 32-44.
03. Kaur, S., S. Bansal and R.K. Bansal, 2014. Steganography and classification of image steganography techniques. Proceedings of the International IEEE Conference on Computing for Sustainable Global Development (INDIACom), March 5-7, 2014, IEEE, New Delhi, India, ISBN: 978-93-80544-10-6, pp: 870-875.
04. Irawan, P.L.T., D.J. Djoko, H. Santjojo and M. Sarosa, 2014. Implementation of cripto-steganography salsa20 and BPCS for digital image data security. J. EECCIS., 8: 175-180.
05. Gkizeli, M., D.A. Pados and M.J. Medley, 2004. SINR, bit error rate and Shannon capacity optimized spread-spectrum steganography. Proceedings of the 2004 International Conference on Image Processing (ICIP'04) Vol. 3, October 24-27, 2004, IEEE, Singapore, pp: 1561-1564.
06. Satish, K., T. Jayakar, C. Tobin, K. Madhavi and K. Murali, 2004. Chaos based spread spectrum image steganography. IEEE Trans. Consumer Electron., 50: 587-590.
07. Joshi, K., R. Yadav and S. Allwadhi, 2016. PSNR and MSE based investigation of LSB. Proceedings of the 2016 International Conference onComputational Techniques in Information and Communication Technologies (ICCTICT), March 11-13, 2016, IEEE, New Delhi, India, ISBN:978-1-5090-0082-1, pp: 280-285.