Method of detecting hidden data transmission via the Koch-Zhao steganographic algorithm

S V Belim, D E Vilkhovskiy
Dostoevsky Omsk State University, Omsk, Russia
Email: sbelim@mail.ru, vilkhovskiy@gmail.com

Abstract. The article proposes an algorithm for analyzing images with embedded message based on the Koch-Zhao steganographic method. The basic idea is to compare the pairs of coefficients of the discrete cosine transform. For this purpose the dependence of the coefficient difference from the block number is constructed. Numerical differentiation of this dependence allows us to identify the boundaries of the embedded message. After that the analysis of the initial dependence on the allocated interval allows to define the used parameters of Koch-Zhao method. The computer experiment was conducted. It is shown that the proposed algorithm allows to determine exactly the location, size and contents of a hidden message if it is embedded in a continuous sequence of blocks.

1. Introduction

For an attack on steganography algorithms in general it is put three tasks. At first, It is necessary to detect the fact of embedding a hidden message. Secondly, It is necessary to determine the size of the embedded message and its location in the container image data array. Thirdly, it is required to restore the built-in message as accurately as possible.

The greatest success was achieved by the steganographic analysis in the solution of the first problem. Most methods for detecting the presence of an embedded message are based on a study of the statistical properties of an image used as a container. Based on the assumption of a random bits distribution of least significant bits in the blue component of the color image, a statistical method of steganographic analysis based on the Chi-square criterion was proposed in [1, 2]. This method allows to detect with a high probability the fact of embedding a message when the container is evenly filled.

The method of visual color layers comparison is proposed in [3]. This approach is effective only if there are enough large areas of uniform fill in the image-container. With a large number of small parts, distortions made by the embedded image are not visually distinguishable. Image modeling with built-in message based on Markov chains used in [4, 5, 6]. This approach is effective if the built-in message is characterized by certain statistical characteristics. In articles [7, 8], the influence of the steganographic insert on the degree of compression of the image-container. It is shown that changes in statistical characteristics due to steganographic embedding of the message reduces the compressibility of the image. And the reduction in compression depends on the size of the message. In articles [9, 10], based on the analysis of the lower layer using the method of analyzing decision hierarchies, an algorithm is proposed that allows to identify with a high enough accuracy the position of LSB inserts. The aim of this work is to develop an algorithm for detecting steganographic insertions in the image embedded with the Koch-Zhao method [11].

2. Formulation of the problem and Algorithm of embedding
As an object of research, we will consider a digital image that is missing information about the presence or absence of an embedded message. It is known only that the method of embedding is Koch-Zhao [11]. We formulate the following three problems. First, it is necessary to determine the presence or absence of a steganographic insert. Secondly, if there is an embedded message, determine its position in the container image and the dimensions. Thirdly, it is necessary to determine as accurately as possible the built-in message, if any, without any a priori information.

The Koch-Zhao steganographic method [11] is based on a two-dimensional discrete cosine transform (DCT). The algorithm for embedding a message consists of the following steps:

1. The original image is divided into blocks of $8 \times 8$ pixels in size.
2. A DCT is applied to each block, the result is a matrix of coefficients $D_i (i = 1, \ldots, N; N$ – is the number of blocks) with a size of $8 \times 8$.
3. Select the sequence of blocks in which embedding will be performed. Each block of information is recorded 1 bit.
4. Two DCT coefficients are selected in each block located in the mid-frequency range of the coefficients, symmetric with respect to the main diagonal ($D_i[3,4]$ and $D_i[4,3]$, $D_i[3,5]$ and $D_i[5,3]$, $D_i[4,5]$ and $D_i[5,4]$).
5. To transfer a bit 0 it is necessary that the difference of modules of a pair of coefficients exceeds some positive value $M_0$. To transfer bit 1, the difference must be less than $-M_0$. Thus, at transmission 0 we increase the module of the first coefficient and reduce the module of the second. When transferring 1 reduce the module of the first coefficient and increase the module of the second.
6. We pass on each block and execute steps 4 and 5.
7. For each block, we perform the reverse DCT.

The choice of the mid-frequency DCT coefficients is due to the need to minimize the effect of embedding on the visual properties of the modified image. The choice of high-frequency or low-frequency coefficients leads to the appearance of effects that are noticeable visually.

When extracting a message, it is considered that pairs of variable DCT coefficients are known.

The extraction algorithm coincides with the embedding algorithm in the first four steps:

5. Calculate the difference between the module values for the coefficient pairs that were embedded.
6. If the difference is greater than $M_0$, then bit 0 was embedded. If the difference of values is less than $-M_0$, then bit 1 was embedded.
7. Sequentially extract bits built into all blocks.

Analysis of the embedding and extraction algorithms shows that in order to make a successful attack on the Koch-Zhao steganographic method it is necessary to determine the blocks in which the message was embedded, the indices of the variable DCT coefficients, and the threshold value $M_0$.

For the correct extraction of the message, the sending and receiving parties must have general secret information about the embedding parameters. We will proceed from the fact that the information about the parameters has a minimum size. In this case, we can formulate three assumptions used in the sequel. First, the embedding is performed in a continuous sequence of blocks. Second, all the blocks use the same pairs of DCT coefficients. Third, all blocks use the same value $M_0$. Any deviation from these assumptions increases the amount of classified information.

3. Algorithm of steganography analysis

To determine the parameters of the steganographic insert, we will use the fact that the parameter $M_0$ must have a large value, allowing the receiving side to extract the hidden message without loss from any image. If $M_0$ is not chosen to be large enough, then the extracted message may cause errors due to the peculiarities of the image-container.
First of all, define the DCT coefficients, which are produced in embedding. For this, as in the embedding algorithm, we split the image into blocks $B_i (i = 1, ..., N)$ with size $8 \times 8$ pixels. For each block $B_i (i = 1, ..., N)$, we perform DCT. As a result, we get a set of coefficient matrices $D_i (i = 1, ..., N)$ with size $8 \times 8$. Let us analyze the mid-frequency elements of the matrix $D_i (i = 1, ..., N)$.

We introduce three sequences of values $(i = 1, ..., N)$:

$C_i^{(1)} = |D_i[3,4]| - |D_i[4,3]|$

$C_i^{(2)} = |D_i[3,5]| - |D_i[5,3]|$

$C_i^{(3)} = |D_i[4,5]| - |D_i[5,4]|$

Embedding a message leads to changes in one of these sequences. We construct the histograms of the dependence $C_i^{(j)} (j = 1,2,3; i = 1, ..., N)$ on the block number $i$. Embedding the message causes one of the sequences to change as a "step" appears with the height $M_0$. An example of such a change is shown in Figure 1.

![Figure 1](image.png)

**Figure 1.** The histogram of the $C_i^{(1)}$ dependence on the block number $i$: a) the image without the built-in message, b) the image with the built-in message.

The problem of detecting the steganographic insert is reduced to the analysis of the dependencies $C_i^{(j)} (j = 1,2,3; i = 1, ..., N)$ on the block number $i$ and the search for step changes. The boundaries of the steps on the histogram can be detected by numerical differentiation based on difference schemes. We numerically differentiate the dependence of $C_i^{(j)} (j = 1,2,3; i = 1, ..., N)$ with respect to $i$.

$$dC_i^{(j)} = C_i^{(j)} - C_{i-1}^{(j)}$$

As a result of this operation, step changes will give high peaks that will allow to define the boundaries of the embedded message. The dependence of $dC_i^{(j)}$ on the block number $i$ for the dependence shown in Figure 1 (b) is shown in Figure 2.
Figure 2. The histogram of the dependence \( dC_i^{(1)} \) on the block number \( i \).

To automatically determine the boundaries of the built-in message for each \( dC^{(j)} \) array, calculate the following values: \( M_j \) - is the maximum value of the array elements \( dC^{(j)} \), \( N_j \) - is the average value of the array elements \( dC^{(j)} \), \( O_j \) - root-mean-square deviation for array elements \( dC^{(j)} \). Calculate \( R_j = N_j + O_j \). We introduce the value \( Y_j \), which will vary in the range of values from \( R_j \) to \( M_j \). Choose the value \( Y_j \) such that there exist exactly two values \( C_{i_1}^{(j)} > Y_j \) and \( C_{i_2}^{(j)} > Y_j \). The values of \( i_1 \) and \( i_2 \) will serve as the boundaries of the built-in message. To determine the value of \( M_0 \) it is necessary to find the minimum value of \( C_i^{(j)} \) in the interval from \( i_1 \) to \( i_2 \).

Formally, the algorithm can be written in the form of a sequence of the following steps:

1. Split the image into \( B_i \) blocks of size \( 8 \times 8 \) pixels.
2. Apply to each block \( B_i \) a discrete cosine transformation. As a result, we obtain matrices of DCT coefficients \( D_i \) with size \( 8 \times 8 \).
3. Calculate a sequence of three values (\( i = 1, \ldots, N \)):
   \[
   C_i^{(1)} = ||D_i[3,4]|| - ||D_i[4,3]||
   \]
   \[
   C_i^{(2)} = ||D_i[3,5]|| - ||D_i[5,3]||
   \]
   \[
   C_i^{(3)} = ||D_i[4,5]|| - ||D_i[5,4]||
   \]
4. Perform numerical differentiation of \( C_i^{(j)} \) (\( j = 1,2,3; i = 1, \ldots, N \)) with respect to \( i \):
   \[
   dC_i^{(j)} = C_i^{(j)} - C_{i-1}^{(j)}
   \]
5. Calculate: \( M_j \) - the maximum value in array \( dC^{(j)} \), \( N_j \) - average value in array \( dC^{(j)} \), \( O_j \) - root-mean-square deviation in array \( dC^{(j)} \). Calculate \( R_j = N_j + O_j \).
6. To enumerate the value of $Y_j$ in the interval from $R_j$ to $M_j$ with step $dY$. Determine the value $Y_j$ such that there are exactly two values $C_{i_1}^{(j)} > Y_j$ and $C_{i_2}^{(j)} > Y_j$. If such a value can not be determined, then reduce the step $dY$. Define $i_1$ and $i_2$.

7. Find the minimum value of $C_i^{(j)}$ on the interval from $i_1$ to $i_2$. Assign $M_0$ to the value found.

8. Extract the message using the found parameters.

The computer experiment showed that this algorithm allows to accurately find and extract the built-in message for the values $M_0 > 54$.

**Conclusion**

Thus, the Koch-Zhao steganographic algorithm is not resistant to the attack of analysis of the DCT coefficients. The algorithm proposed in this article allows you to accurately extract the built-in message, provided that it is the only one and is embedded in the continuous area. Deviation from these assumptions increases the stability of the steganographic algorithm.

**References**

[1] Provos N and Honeyman P 2001 Detecting steganographic content on the internet *CITI Technical Report 01–11* (Ann Arbor: University of Michigan)

[2] Westfeld A and Pfitzmann A 2000 Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and STools-and Some Lessons Learned *International Workshop on Information Hiding* pp 61-76

[3] Aliev A T 2004 On application LSB steganographics method to digital images with the greater monochrome areas *Vestnik DGTU* 4 pp 454-460

[4] Barsukov V S 2000 Ocenka urovnya skrytnosti multimedijnych steganograficheskikh kanalov hraneniya i peredachi informacii *Specialnaya Tekhnika* 1

[5] Kustov V N and Paraskevopulo A Ju 2005 Prostye tajny stegoanaliza *Zashhita informacii INSIDE* 4 pp 72-78

[6] Golub V A and Dryuchenko M A 2009 Steganographic information detection in JPEG files with the help of complex usage of several stego-attackes *Infocommunication Technologies* 7 pp 44-50

[7] Zhilkin M Yu 2008 Stegoanaliz graficheskikh dannykh v razlichnykh formatakh *Doklady TUSURA* 2 pp 63-64

[8] Monarev V A 2012 Sdvigovyy metod obnaruzheniya skrytoy informatsii *Vestnik SibGUTI* 4 pp 62–68

[9] Vilkhoskiy D E and Belim S V 2017 Algorithm for detection of steganographic insertions of LSB-substitution type based on the least significant bits analysis *Informatika i sistemy upravleniya* 4 pp 3-11

[10] Vilkhoskiy D E and Belim S V 2017 Algorithm for detection of steganographic inserts type LSB-substitution on the basis of an analysis of the zero layer *J. Phys.: Conf. Ser.* 944 010212

[11] Koch E and Zhao J 1995 Towards robust and hidden image copyright labeling *IEEE Workshop on Nonlinear Signal and Image Processing* pp 452-455