Send Orders for Reprints to reprints@benthamscience.ae

406 *The Open Cybernetics & Systemics Journal*, 2015, *9*, 406-414

Open Access

Research on the High Capacity Hiding Information Algorithms in Binary

Li Dongmei[1,*] and Wei Ke[2]

[1]*School of Computer & Information Engineering, Nanyang Institute of Technology, Henan, Nanyang, 473000, P.R. China;* [2]*Zhongzhou University, Henan, Zhengzhou, 450044, P.R. China*

Abstract: High capacity hiding information algorithms in binary host image is developed according to characteristics of binary image. The binary host image is first divided into blocks of size 2×2, and then each sub-block adaptively determines the embedded bits number and their positions according to the difference between its black pixel and white pixel number. Each sub-block in the transitional block between black pixel and white Pixel can hide at least two bits messages. Therefore, its embeddable capacity is two times more than that of other presented blind algorithms.

Keywords: Binary image, embedding strategy, hiding capacity, information hiding.

1. INTRODUCTION

Binary image that is black and white image, is a very broad application of the image, with the accelerating global information digitization process, a large number of important information, such as personal files, medical records, diploma, patent documents, handwritten signature, design drawings, collection of books, confidential documents, scanned documents into digital binary image of the way to keep. Obviously, with the general grayscale, color images or audio, video compared to the value of these binary image to be much more expensive, and therefore, in which the protective hide information (such as a digital watermark) it is very important. Unlike binary image grayscale images with rich gradation level, in which each pixel represents only a "0" represents black, "a" represents white. Because black and white pixel comparison of the two distinct, randomly change the binary image pixels are likely to cause significant changes signs, especially in the flat area binary image by flipping black and white pixels to hide information without introducing visible traces is almost impossible, so the large capacity of information embedded in the binary image is very difficult. In grayscale or color images in an effective information hiding method can't be directly used to contain only two colors black and white binary image, you m must study for this type of image information hiding technique.

2. BINARY IMAGE INFORMATION HIDING FEATURES

Binary image has only two gray levels, *i.e.*, content images "more black and white", each pixel may use a binary code "0" or "1", "0" indicates a black pixel, "1" indicates white pixels. According to Weber's law shows that, when modifying the amount of pixel values of a pixel is below the contrast threshold, the visual system can not feel the presence of the signal, a pixel grayscale image will be in the LSB from "0" to "1" does not cause (or vice versa) to embed a bit of information in any visible distortion, which is the LSB algorithm works, but is different in the binary image, a pixel value from "0" to "1" (or vice versa), the amount of change has gone beyond the point of perception threshold, thus changing the value of a pixel in the binary image, if the pixel considered in isolation, the modification can definitely be perceived. This makes the embedded information in the binary image, like that can't be considered in isolation of one pixel in grayscale or color images at a pixel modify the embedded information, we must consider the condition of the neighborhood pixels. The original binary image is shown in Fig. (**1a**). If the surrounding area is all in the "0" of a pixel from "0" to "1", a white spot appears in a black area, you can easily be detected, as shown in Fig. (**1b**) as shown ; all around, "'the region will be one pixel from" 1 "to" 0 ", that is the emergence of a black dot on a white area, the situation is the same, as shown in Fig. (**1c**) as shown ; if the black and white areas at the junction of a pixel from "0" to "1 " (or vice versa), is generally not noticeable, as shown in Fig. (**1d**). Thus, the binary image is embedded in the information must be in black and white areas junction.

For grayscale images can be hidden information in the transform domain, due to the spatial domain image data is an integer, so when the transform domain back to the spatial domain by inverse transform, generally there will be a rounding error, but the error will not be great, so long as the embedding strength not too low, the hidden information will not be lost this rounding error. However, there are cases of the binary image significantly different, the binary image, the watermark information is embedded in the transform domain, and then returns back to the spatial domain, in order to ensure that the binary image of, the need for the binarization processing, and this binary operation is usually greatly weaken the hidden information, and even remove the watermark. Lu *et al.* [1] Proof: For a binary image, if the watermark is embedded by modifying the DC coefficient of the DCT coefficients (DC), then after binarization operation, the embedded information will cease to exist, can be seen, Gray

*Address correspondence to this author at the School of Computer & Information Engineering, Nanyang Institute of Technology, Henan, 473004, P.R. China; Tel: 13803770071; E-mail: ywind2005@163.com

1874-110X/15 2015 Bentham Open

RETRACTED ARTICLE

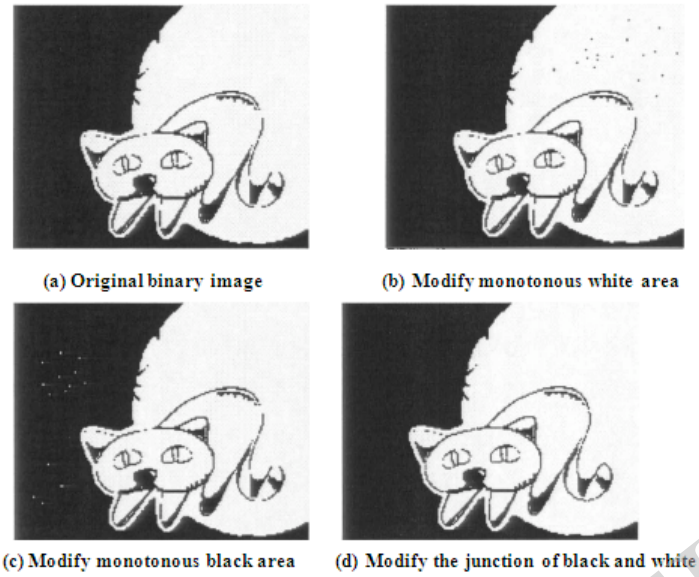

Fig. (1). Perceived modified binary image pixel values in different regions.

of the image applies DCT transform domain information hiding algorithm is not suitable for binary images.

In the binary image information hiding, there are two basic ways [2]: First, modify the value of a single pixel, the second is to modify a set of pixel values. The first method is the modification of a black pixel into a white pixel or a white pixel will be modified to a black pixel, the second method is to modify the binary image of the strokes, lines, etc. The thickness, curvature, relative positions and other features, this method in general more dependent on the type of image, such as text, drawing and the like. At the request of the hidden information extraction and invisible blind premise, except in special binary image types, the amount of information hiding second method is extremely limited.

3. STATUS QUO OF THE BINARY IMAGE INFORMATION HIDING

In recent years, scholars have been at home and abroad in terms of binary image information hiding proposed a series of algorithms for the binary image information currently hiding algorithm can be divided into the following categories: (1) By modifying the binary text document line spacing or word spacing embedded data, (2) The binary image into image blocks of a specific size, the embedded data in the image block by calculating the statistical characteristics of the pixels of certain methods, (3) By modifying the boundary mode to certain regions of the embedded data, (4) By modifying characters some local features embedded data, (5) By modifying the run length to hide secret information to fax a document, (6) In the halftone image of an image based on error diffusion to embed data.

3.1. Modify Binary Text Document Line Spacing or Word Spacing Information Hiding Algorithm

According to human visual characteristics, when the vertical displacement is less than 1/300 inch line of text, the horizontal displacement amount is less than 1/150 inch words, the human eye will not be identifiable. For text documents, in normal circumstances, all the spacing between two adjacent rows is the same, the spacing between adjacent characters is also the same. Modify the line spacing or word spacing information hiding algorithm is between two values by the rows and rows of text or images between words to make minor adjustments to the displacement of the embedded information, even though the line between the word line or word the distance is no longer the same, but if there is no original document as a control, the human eye is unable to identify them, thus completing the hidden information.

From the principle point of view, modify line spacing of information hiding algorithm embedding strategy is very simple, you just need to want to hide secret information, in a word processing program to adjust the line spacing, use of confidential information to the modulation of the text line spacing parameters, you can information embedding process is completed. For example, the character line up move embedding "1" to embed down to "0"; the character to the left embedding "1" to the right to embed a "0." Document hidden information can be delivered in electronic document distribution, or it can be printed as a hard copy print distribution. Line spacing adjustment embedding algorithm has strong robustness, even after repeated copying, or page repeatedly press a telescopic zoom factor, the embedded information can be detected. This is because the distortion caused by the page copy operation in more gradual and primarily in the same direction, it almost does not change the relative distance between the upper and lower reference lines, does not affect the detection performance, these features make the line spacing coding technique can withstand most deformable attack.

When extracting information by analyzing the line spacing or word spacing to determine the embedded content, without the need for the original document as a control, you can achieve blind extraction. Line spacing hiding algorithm based on the specific extraction method is: embedded into rows and rows on the information embedded control, line spacing control lines remain unchanged, only the line spac-

RETRACTED ARTICLE

ing can be embedded in the line for testing. Thus the detection of line spacing and character spacing control line can be used as a reference spacing, line spacing, by detecting and comparing the detection can be completed, if the line shift is coded as "1", if the line down, the coding "0." Testing and comparing the line spacing can be used to detect the centroid method, the centroid is defined as the center of the horizontal axis line. With $\Delta_{R,+}$ represents the distance to be detected on a row, and does not move the line between the centroids, with $\Delta_{R,-}$ row is detected and it does not move to the next row between the centroid representation distance, and with $\Delta_{X,+}$ and $\Delta_{X,-}$ denote two without making changes in the original document, the corresponding centroid distance.

According to changes in the distance to the center of mass can be determined on this line moves up or down the volume. In case:

$$\frac{\Delta_{R,+}-\Delta_{R,-}}{\Delta_{R,+}+\Delta_{R,-}} > \frac{\Delta_{X,+}-\Delta_{X,-}}{\Delta_{X,+}+\Delta_{X,-}} \quad (1)$$

Indicates that it is increased with the distance from the previous line, *i.e.* the line is down, the embedded information is "0." In case:

$$\frac{\Delta_{R,+}-\Delta_{R,-}}{\Delta_{R,+}+\Delta_{R,-}} < \frac{\Delta_{X,+}-\Delta_{X,-}}{\Delta_{X,+}+\Delta_{X,-}} \quad (2)$$

Indicates that it is reduced with the distance on a line, *i.e.*, the line is moved up, the embedded information is "1". Defined functions:

$$f(R)=\left(\operatorname{sgn}\left(\frac{\Delta_{X,+}-\Delta_{X,-}}{\Delta_{X,+}+\Delta_{X,-}}-\frac{\Delta_{R,+}-\Delta_{R,-}}{\Delta_{R,+}+\Delta_{R,-}}\right)+1\right)\Big/2 \quad (3)$$

Word spacing adjustment information hiding algorithm embedding process and adjust the line spacing consistent hiding algorithm, such as the word information to be embedded blocks and increasing the distance between the reference chunks left of it, said embedded information bit is "l", it the distance between the right side of the base chunks embedded information, said the increase is "0." Based on experience, the human eye can not identify changes within 1/150 inch. In fact, when most word processing software, document alignment processing typeset documents often use variations of the word spacing, making the text attractive in appearance, the reader can accept the text on one line word spacing varies widely. And line spacing adjustment hiding algorithm is similar in the embedding process, the need to set up some word spacing nonlinear distortion constant reference chunks of information used as a reference for the detection and compensation for printing, scanning, etc. caused. In principle, the information is embedded by adjusting the spacing between any two characters, the only limit is the sum of all the rows of the coded words of the displacement of the pitch should be equal to 0, in order to maintain the correct ordering of rows is not disrupted, so the application can be embedded in each row a plurality of bits of information. When extracting the embedded information, you need to determine the exact position of the reference word block. Therefore, the information extraction process to adjust line spacing than the method to more complex, the ability to resist attack methods are also weaker than the line spacing adjustments.

If the document after embedding information in the form of electronic document distribution, it can be directly extracted, when text is printed if the image print, scan obtained during the extraction of information contained in confidential documents because of possible contamination through wear and tear, or multiple copies text image scanned inevitably introduces some noise, not suitable for direct extraction of information that is in front of the extraction requires some pretreatment. Major steps have to remove the salt and pepper noise preprocessing, tilt correction, binarization. Text image is usually caused by some equipment regeneration, copiers, scanners, and so can be seen as a noisy channel, the noise generated can be considered primarily salt and pepper noise. An effective way to reduce the salt and pepper noise, median filtering is performed. Copying and scanning process will make the text tilted, before extracting need to tilt correction. Finally, convert the scanned image format is binary image, binarization threshold to determine the available global threshold method, the local threshold method or a combination of algorithms [3].

Based on similar principle, in [4] describes a to embed information, characters or words by changing the size of the character height can be slightly increased a little, or slightly increased width that does not cause significant visual abnormalities. Adjust the text with line spacing or word spacing compared embedding method, this information is embedded method has high embedding capacity. Experimental results show that the change in the size of text characters 1/300 inch can be reliably detected, and can withstand repeated photocopying attacks. When embedding, always modify the current character or word size, the same size is modified characters or words adjacent words or characters remain unchanged, so that when the size of adjacent watermark extraction of words or characters can be used as a reference dimension. This watermark embedding method is very sensitive to the character of the tilt, rotate small text page will seriously affect the watermark extraction, even if the tilt correction does not help. Therefore, how to resist the spin attack is the method further improved content.

3.2. Block Concealment Methods

Information Hiding image block is a more typical concealment methods, such method is to embed the image into a size of M × N sub-blocks, and then divided into sub-blocks are analyzed to determine which blocks can be used to embed information of information can be used to embed sub-blocks according to certain rules embedded computing determine strategies to identify the pixels that can be modified according to the embedding strategy, and then directly to the pixels to be modified. This method is simple algorithm, high hiding capacity, compared to other algorithms in terms of having a more practical. In this way all the binary image are applicable.

Wu [5] proposed a typical block hiding method for binary image is divided into 3 × 3 sub-block size, data is embedded, the embedded information is required by 1 or 0 to adjust the black pixels within the sub-block parity, the total

RETRACTED ARTICLE

number of black pixels to the total number of odd and even represent 1 and 0 is embedded. To reduce the potential distortion modify pixels, first calculation sub-pixels in the block inversion priority, the priority of a pixel can be expressed as flip flip visual distortion estimates, with the lowest priority of the pixels are used to embed data. Visual distortion metric, consider the smoothness and connectivity to the midpoint of the pixels is 3 × 3 window. Smoothness is determined by the pixel changes in horizontal, vertical and diagonal directions, the connectivity within the window by black and white cluster number of clusters evaluation. To equalize the embedded data in the image, increasing the random permutation of all pixels reversible step. This method when extracting data is as simple as can be determined based on the information hidden within the black pixel sub-blocks of the total parity is a 1 or 0, without the original image. Odd number of pixels can also be controlled by another means of: selecting a value, the size of Q, forces the black pixel group is equal to the number of pixels in the embedded watermark 2kQ "0", where k is a positive integer ; or force the pixel group number of black pixels is equal to (2k + 1) Q to embed watermark information "1." Q can be obtained by increasing the value of better robustness, but also increase the number of data processing and image quality.

To improve the robustness, [6] in the vector image is divided into sub-blocks of 8 × 8, with the number of black and white pixels in each block to represent the embedded information, the specific embedding strategy: The image is "a "percentage of pixels, to set a threshold, with the block to" 1 "the percentage of pixels exceeding the threshold to indicate" 1 ", is below the threshold, said" 0. "As described above for each sub-block is embedded to make modifications to complete embedding. For a general binary image, the algorithm modified when embedded within the image in black and white change Comparative distinct boundary pixels, for the halftone image is embedded modifying those pixels isolated, or else a large number of pixel changes will cause significant visual anomalies. The method for some lossy compression, low-pass filtering and image format conversion and other operations with a certain robustness, but the image quality of the embedded information block will be significantly reduced, and the higher the threshold, the lower the quality, the more obvious.

3.3. Text Feature Modification Method and Boundary Modification Method

These two methods are of fine-stroke text embedded information processing, mainly for text documents to hide information.Text features modified algorithm based on the characteristics of the text stroke is embedded with a better visual effect.

Amamo and Misaki [7] gives the specific implementation of this algorithm.

Step 1: analysis identified the text area of the image using the connection components. Then follow the enclosed space grouping, each group is divided into sub-blocks of a four-part, four sub-block is divided into two sets. Firstly stroke fonts connection, and then based on the analysis by stroke block, and then share the strokes block into four parts. Step 2: Calculate the average width of each stroke, the mean value, generally used to calculate the run. Step 3: Define two operations: "thicker" and "thinning", which by increasing and decreasing the width of the vertical run to achieve. Step 4: The four strokes block is divided into two groups, one group of stroke by making thicker, thinner strokes another group to complete the information embedded. Embedding strategy is simple: the first group thicken thinning the second group, said embedding "1", and vice versa, said embedding "0."

Information extraction is relatively simple, first by blocking mode when embedding the image block, and then calculate the thickness of strokes, each stroke thickness comparison of the four sub-blocks are the same to complete the extraction, which can resist printing: Scan re-digitize deformation attacks, strokes long digitized image is still located.

The literature [8] propose a method based on hidden boundary, it is first divided into a plurality of basic character boundary, the length of each border fixed boundary is 5 pixels, the fixed length of the boundary information embedded in the boundary mode 5, embedding strategy for the two visually similar boundary as the boundary of the information-carrying (Fig. **2**), wherein a stands for "0", the other representing "1", the boundary of visually similar to carry information interchange. Without reference to the original text extracted by the algorithm, the application does not require any special enhancement techniques to extract information can be carried out. Experimental results show that the method may allow a resolution of embedding articles on 300dPi full digitized documents, each character block of information can be embedded 5.69bits resolution 200dPi each character block can be embedded 5.69bits information, but the resolution 100dPi block for each character can only be embedded 0.17bits information. This method can be used, such as a text document in addition can also be used in engineering drawing document image.

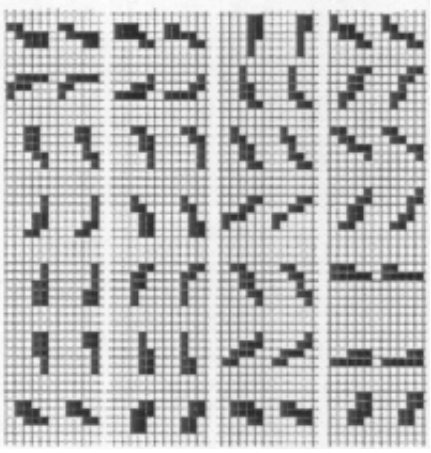

Fig. (2). Fixed length of the boundary of the boundary mode 5.

3.4. Based on a Halftone Image Embedding Algorithm

Halftone technique is to use a single color in a certain range of concentrations of hundreds of kinds of mold continuous gradation, which uses digital spatial resolution in exchange for brightness amplitude resolution, so its essence is a binary image. Due to the limited capacity of the reasons for the human visual resolution, halftone image at a distance

RETRACTED ARTICLE

of visual effects and original images similar to that that it is a continuous-tone images. Halftone image is now widely used in newspapers, magazines, documents, money, checks and confidential documents such as image output. Since the half-tone image with a spatial resolution to represent luminance resolution, and therefore can be used to directly modify the pixel information embedding method. Of course, embedding method for halftone images, there are some special requirements, some algorithms require the original grayscale image.

Matsui and Tanaka [9] proposed a modification of the pixel run length information embedding method carried out by modifying the runlength hide the secret information, the main embodiment for the halftone image fax files. Fax document contains 1728 pixels per line, which is 1728 pixels contains many run, run data recorded data using Huffman coding. Specific embedding strategy is very simple, that modify pixel run edge, so run plus one or minus one, with a run of parity to carry information. Obviously this method can also be applied to all of the binary image, but in the short run when more will affect the visual effect of the image, but when you run too embedding capacity and relatively small.

The literature [10] proposed a method based on hidden DCT domain embedding strategy is to adjust the three intermediate frequency DCT coefficients of the maximum absolute value of a position to realize that the embedded information, but strictly speaking, the algorithm does not suitable for halftone images, because of the embedding process is completely against the gray image, but by the print / scan after the embedded information can be extracted. In order to ensure accurate extraction of information by repeatedly embedding method of embedding capacity of the process is relatively small, in 256x256 gray scale image, only embedded in 56 bits, the method of superposition of noise resistance, attacks such as cropping the ability is relatively poor.

The literature [11, 12] proposed three kinds of halftone images directly embedded in hiding algorithm was DHST, DHPT, DHSPT, these three kinds of hiding algorithm does not require the original grayscale. DHST method for random modification of an isolated halftone image so that the pixel values of the pixels with the same data bit to be embedded. DHPT method for modifying randomly selected pair of black and white pixels, so as to carry the embedding information. DHSPT also modify a pair of black and white pixels, while taking into account the changes will not modify significantly reduce the visual approach.

The literature [13] is modified to the above 3 kind of embedding method. This method is no longer embedded in randomly selected position, but the size is selected according to the brightness of the embedded position. The method generally select high brightness or low brightness region information is embedded in the intermediate luminance region information embedding is not performed. The method of embedding capacity without reducing the visual effect of improving the premise after embedding, but performing the algorithm needs to be embedded to make the luminance reference picture. The literature [14] proposed a high-capacity halftone watermarking algorithm based on error diffusion, the embedding capacity than the literature [11, 12] have significantly improved.

In summary, because of the special nature of the binary image, the current applied to the binary image information hiding algorithm is far from hiding algorithm based on grayscale image of a mature, there are still insufficient in many ways, and some can not be easily applied to binary text other than the binary image of the image, some small amount of data can be embedded, and some can not modify the image will maximize the spread to the entire image, and some can not be blind to extract information, but also in the robustness, hidden capacity , imperceptibility, security, and so far from perfect, the most prominent of which is hidden a small capacity, so there is still much work to be done.

4. LARGE-CAPACITY INFORMATION HIDING ALGORITHM BASED ON PARTIAL BLOCK

4.1. Block and Embedding Strategy

In the binary image hiding algorithm, a method block relative to the other parties who have a simple embedding method, embedding large capacity, good concealment features, and modify binary image strokes, lines, etc. The thickness, curvature, the relative position of other characteristics of the amount of information hiding algorithm is extremely limited, so applications need to embed more information block algorithms are generally used. The principle is to block algorithm image into a size of M × N sub-blocks of the partition of the sub-blocks according to certain rules and calculations to determine the embedding strategy, and then find you can modify pixels based on the embedding strategy, and then directly to the modify pixels (black pixels will be modified into a white pixel or a white pixel to modify the black pixels), enabling embed hidden information. Literature [4, 5] have proposed an algorithm based on block, [4] in the image is divided into small pieces 3 × 3, the literature [5] in the image into small blocks of 8 × 8, the literature [2] in the image is divided into small blocks of 2 × 2, they are each sub-block is embedded a bit of information, clearly the smaller block, the more the number of available sub-block, the greater the capacity can be embedded, literature [1] is an algorithm embedded in a larger capacity. Therefore, this algorithm uses 2 × 2 block method.

Because the binary image "black and white" features, embedding hidden information in the sub-block after block, in order not to have a significant visual distortion, so can not be considered in isolation as a pixel grayscale image embedded case, which must be the same of the pixels adjacent to the pixel considered in conjunction with the other according to human visual characteristics. According to the human visual characteristics, in a white area in a black spot, because of contrasting visually is very obvious, also in a black area in a white point is very clear, and the other eye on the horizontal and vertical directions the sensitivity is greater than two diagonal directions. Therefore, this algorithm is not all white or all black sub-blocks embedded information, while at the same time there is only black and white pixels within the sub-blocks of embedded information, embedding modify pixel color information as possible be carried out in a diagonal direction. Meanwhile, in order to realize blind extraction of information embedded in the process of embedding information needed to avoid the sub-block after the changes to all white or all black.

RETRACTED ARTICLE

Because the binary image "black and white" features, embedding hidden information in the sub-block after block, in order not to have a significant visual distortion, so can not be considered in isolation as a pixel grayscale image embedded case, which must be the same of the pixels adjacent to the pixel considered in conjunction with the other according to human visual characteristics. According to the human visual characteristics, in a white area in a black spot, because of contrasting visually is very obvious, also in a black area in a white point is very clear, and the other eye on the horizontal and vertical directions the sensitivity is greater than two diagonal directions. Therefore, this algorithm is not all white or all black sub-blocks embedded information, while at the same time there is only black and white pixels within the sub-blocks of embedded information, embedding modify pixel color information as possible be carried out in a diagonal direction. Meanwhile, in order to realize blind extraction of information embedded in the process of embedding information needed to avoid the sub-block after the changes to all white or all black.

4.2. Secret Information Embedding Algorithm

The binary image is divided into a plurality of sub-blocks of size 2 × 2, and each individual intake embedding and extracting information. Because all "1" (white) or all "0" (black) image block is all white or all black image smoothing block, modify such an image block of pixels is easily perceived by the human eye, it is not used to hide information, other blocks of information can be embedded, in the information extraction also considered all "1" or all "0" of the block without the hidden information. Algorithm is described as follows:

Set of vector images for H={$x_{i,j}$,i=1,2,…2M, j=1,2,…,2N}, Namely: the size of vector image of 2M × 2N, where (i, j) representative the pixel coordinates of the image, x_ (i, j) is the pixel value of the corresponding positions, $x_{i,j} \in \{0,1\}$ to be referred to as the hidden information W={w,t=0,1,…,L-1},W_j={0,1}. Embed hidden information in the following steps:

Step 1: m for hidden information need be carried out under the control of the key k1 scrambling to give m *, so you can improve the security of embedded systems, embedded images for class when the hidden information can also improve its robustness;

Step 2: The image block is divided into a sequence of binary image H MN disjoint 2 × 2 image block $B_{k,l}$.

$$B_{k,l} = \begin{bmatrix} b_{(k,1)(1,1)} & b_{(k,1)(1,2)} \\ b_{(k,1)(2,1)} & b_{(k,1)(2,2)} \end{bmatrix}$$

$$k = 1,2,...,M; l = 1,2,...,N \tag{4}$$

where: $b_{(k,1)(u,v)} = x_{2(k-1)+u,1(l-1)+v} \quad u = 1,2; v = 1,2$

Step 3: Under the control of the key k2, a randomly selected sub-block $B_{k.l} : k = 1,2,...,M; l = 1,2,...,N$ embedding processing information. Because it is embedded randomly selected blocks, knowing concealment methods but do not know the attacker can't extract key k2 is hidden information. Embed policy information as follows:

l) Computing sub-block in the number of white pixels

$$y_{k,l} = \sum_{u=1}^{2}\sum_{v=1}^{2} b_{(k,1)(u,v)} \tag{5}$$

2) If $y_{k,l} = 0$ or $y_{k,l} = 4$, then block $B_{k.l}$ is not embedded information without block $B_{k.l}$ any modification;

3) If $y_{k,l} = 2$, then block $B_{k.l}$ 3-bit information is embedded, embedding principle:

$$b'_{(k,1)(1,1)} = m^*_{z+1} \tag{6}$$

$$b'_{(k,1)(1,2)} = m^*_{z+2} \tag{7}$$

$$b'_{(k,1)(2,1)} = m^*_{z+3} \tag{8}$$

$$b'_{(k,1)(2,2)} = \left(m^*_{z+1} + m^*_{z+2} + m^*_{z+3} + 1\right) \bmod 2 \tag{9}$$

Wherein after z is the number of information bits $B_{k,l}$ has already been embedded, $b'_{(k,1)(u,v)}$ for pixel $b_{(k,1)(u,v)}$ embedding information value.

4) If $y_{k,l} = 1$ or $y_{k,l} = 3$, then block $B_{k,l}$ 2-bit information is embedded, embedding principle:

$$b'_{(k,1)(1,2)} = m^*_{z+1} \tag{10}$$

$$b'_{(k,1)(2,1)} = m^*_{z+2} \tag{11}$$

The other two pixel values within the sub-block modified three cases:

If $m^*_{z+1} + m^*_{z+2} = 0$, then

$$b'_{(k,1)(1,1)} = b'_{(k,1)(2,2)} = 1 \tag{12}$$

If $m^*_{z+1} + m^*_{z+2} = 2$, then

$$b'_{(k,1)(1,1)} = b'_{(k,1)(2,2)} = 0 \tag{13}$$

If $m^*_{z+1} + m^*_{z+2} = 1$, calculate:

$$r_1 = b_{(k-1,l-1)(2,2)} + b_{(k-1,l)(2,2)} + b_{(k-1,l)(1,2)} + b_{(k,l-1)(2,1)} + b_{(k,l-1)(2,2)} \tag{14}$$

$$r_2 = b_{(k+1,l+1)(1,1)} + b_{(k+1,l)(1,1)} + b_{(k+1,l)(2,1)} + b_{(k,l+1)(1,2)} + b_{(k,l+1)(1,1)} \tag{15}$$

If $r_1 > r_2$, then

$$b'_{(k,1)(1,1)} = 1, b'_{(k,1)(2,2)} = 0 \tag{16}$$

Otherwise

$$b'_{(k,1)(1,1)} = 0, b'_{(k,1)(2,2)} = 1 \tag{17}$$

RETRACTED ARTICLE

Step 4: Under the control of the key k2, and then the remaining blocks in the next randomly selected sub- $B_{k,l}$, using the method of embedding other step3 to hide information. This process is repeated until all required information is embedded hidden completed.

4.3. Hidden Information Extraction Method

The received image is embedded with the same sub-block processing, and then in the key k: under control, in order to identify possible hidden secret information of sub-blocks, and then extract the hidden information in these sub-blocks. Extracting embedded information is very simple, if the block is "0" the number is 2, the pixel value of the lower left and upper right in the block to extract the information, if the block is "0" for the number 1 or 3, block pixel values to extract the information within the 3 other than the lower right pixel, if the block is "0" for the number 0 or 4, it indicates that there is no hidden information within the block.

Sub-block hidden information extraction algorithm is:

Step l: Counting the number of sub-blocks and other information extracted in the white pixel.

$$y'_{k,l} = \sum_{u=1}^{2}\sum_{v=1}^{2} b'_{(k,l)(u,v)} \tag{18}$$

Step 2: Depending $y'_{k,l}$ values, and at the same time, like embedding in three cases to extract different amounts of hidden information:

l) If $y'_{k,l} = 0$ or $y'_{k,l} = 4$, the sub-block without hidden information extraction process is not performed;

2) If $y'_{k,l} = 2$, then block $B'_{k,l}$, containing two bits of information, $b'_{(k,l)(1,2)}$ and $b'_{(k,l)(2,1)}$ is to extract information;

3) If $y'_{k,l} = 1$ or $y'_{k,l} = 3$, then block $B'_{k,l}$ contains 3 bits of information, $b'_{(k,l)(1,1)}$, $b'_{(k,l)(1,2)}$ and $b'_{(k,l)(2,1)}$ is to extract information;

Step 3: All the extracted hidden information are combined, and dried prior to embedding the pre-processing performed corresponding scrambling algorithm and the corresponding inverse scrambling key, to obtain the secret information.

4.4. Experimental Results and Analysis

After the peak signal to noise ratio due to the often used to evaluate gray image distortion (PSNR) is not suitable for binary images, for objective evaluation, respectively [1] and objective metrics VDSF second chapter proposed to measure the embedded information distortion introduced, when the change in [1] occurs in the method of measurement values of the pixels through which the Effect of eight surrounding pixels to measure the distortion introduced after embedding information, is defined as the coordinates (i, j) of the pixel modified distortion D_ (i, j) is introduced:

$$D_{i,j} = \sum_{u=-1}^{1}\sum_{v=-1}^{1} abs\left(X_{(x-u)(j-v)} - \left(1 - x_{i,j}\right)\right) \times k_{(i+2)(j+2)} b'_{(k,l)(u,v)} \tag{19}$$

Formula (19), $k = \begin{bmatrix} 1/12 & 1/6 & 1/12 \\ 1/6 & 0 & 1/6 \\ 1/12 & 1/6 & 1/12 \end{bmatrix}$ and $1 - x_{i,j}$ is $x_{i,j}$ color value after the change. The maximum degree of distortion of a modified pixel is 1, the minimum distortion is 0.

If the image has an S pixels changed, the distortion of the entire image is written as:

$$D = \sum_{r=1}^{S} D^r_{i,j} \Big/ S \tag{20}$$

Figs. (**3** and **4**) are respectively a non-binary image and the binary image of text type experimental results random information embedded information can be embedded in all sub-blocks after embedding capacity and the degree of distortion shown in Table **1**, showing the algorithm embedding capacity than the literature [2] to be more than twice as large, it is much larger than other algorithms, but this time the distortion D with the literature [2] are basically the same, though not imperceptible index [2], but still imperceptible to meet the visual requirements.

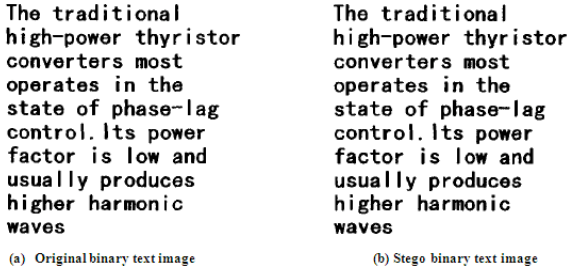

Fig. (3). Data hiding results of english text type binary image.

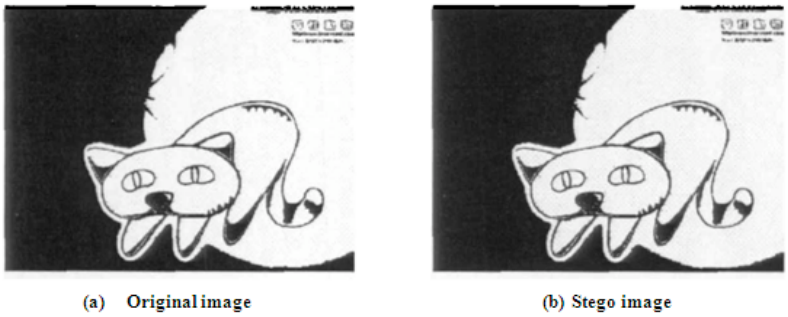

Fig. (4). Non-text-type binary image data hiding results.

RETRACTED ARTICLE

Table 1. Embedding capacity data tables.

Binary Image	English Text Binary Image	Non-Text-Based Binary Image
Image size	256×256	400×300
The number of available sub-blocks	863	1122
Embedding capacity (bit)	863	1122
This algorithm embedding capacity (bit)	2289	2774
Distortion D	0.3033	0.3128
The algorithm distortion D	0.2988	0.3017
Imperceptible indicators VDSF	45.5088	47.5877
Imperceptible indicators VDS	40.7563	42.9925

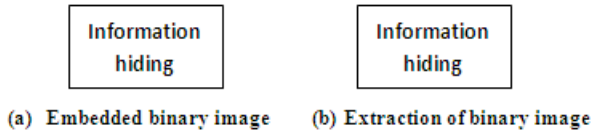

Fig. (5). Contrast of embedded image and extraction image.

As shown in Fig. (**5**) after embedding the binary image shown in Fig. (**4a**) of the binary image of 52 × 52, and the extracted image in Fig. (**5b**) as shown, which in Fig. (**5a**) is identical , illustrated by the algorithm embedded information can be extracted 100% correct.

In this algorithm, information is embedded into the sub-blocks which are controlled by the key, while the sub-block information is not embedded with the sub information is embedded in the characteristics after the block without distinction, so that the information hiding algorithm is fully disclosed, attacker did not extract the embedded key information.

CONCLUSION

On the basis of the block algorithm presents a lot of secret information can be hidden in the binary image algorithm, the binary image into image blocks of size 2x2, and then each sub-block according to the number of black and white pixels different, adaptively determined embedding hidden information can be embedded in the position and number of bits of information. Experiments show that a large amount of information hiding algorithm, the computational complexity is low, stego image without significant distortion. When extracting the hidden information, each block is determined according to only "0" and "1" bits in the sub-block is embedded hidden information, and the number of bits of the secret sub-block and location information, and therefore hidden information extraction without the original carrier image and other auxiliary information, is a completely blind extraction. Meanwhile, at the time the information is embedded, the embedding position is controlled by the key, the key security resides, hence the information hiding algorithm can be public, in line with the principles Kerokhoffs. The algorithm focuses on information hiding method, the influence of noise and long-distance transmission of hidden information for further study.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] H. Lu, X. Shi, and Y. Q. Shi, "Watermark embedding in DC components of DCT for binary images", In: *IEEE International Workshop on Multimedia Signal Processing*, US, Virgin Islands, 2002, pp. 300-303.

[2] L. Zhou, Y. Yang, and Y. Guo, "Summary binary image based on information hiding," *Journal of Sun Yat-Sen University (Natural Science)*, vol. 43, no. 52, pp. 71-75, 2004.

[3] Y. Xu, *Text Carrier Information Hiding Technology Research*, Information Engineering University, Beijing, 2006.

[4] J. Brassil, L. and O. Gorman, "Watermarking document images with bounding box expansion," In: *Proceedings of the Ist International Workshop on information Hiding*, Cambridge, UK, pp. 227-235, 1996.

[5] M. Wu, E. Tang, and B. Liu, "Data Hiding in Digital Binary Image", In: *IEEE International Conference on Multimedia and Exposition*, New York, Aug, 2000, pp. 393-396.

[6] J. Zhao, and E. Koeh, "Embedding robust labels into images for copyright protection", In: *Proceedings of International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technologies*, Vienna, August 21-25, 1995, pp. 242-251.

[7] T. Amamo, and D. Misaki, "Feature calibration method for Watermarking of document images", In: *Proceedings of 5th International Conference on Document Analysis and Recognition*, Bangalore, India, 1999, pp. 91-94.

[8] Q. Mei, E. K. Wong, N. Memon, "Data hiding in binary text documents", In: *SPIE Proc. Security and Watermarking of Multimedia Contents*, San Jose, CA., Jan. 2001, pp. 369-375.

[9] K. Matsui, and K. Tanaka, "Video-steganography: how to secretly embed a signature in a picture", *IMA Intellectual Property Project*, vol. l, no. l, pp. 187-206, 1994.

[10] S. Niu, X. Niu, and Y. Yang, "Halftone image data hiding algorithm," *Journal of Electronic*, vol. 32, no. 7, pp. 1180-1183, 2004.

[11] M. S. Fu, and O. C. Au, "Data hiding for halftone images", In: *Proceedings of SPIE Conference On Security and Watermarking of Multimedia Contents*, Singapore, 2000, pp. 228-236.

[12] M. S. Fu, and O. C. Au, "Data hiding by smart pair toggling for halftone images", In: *IEEE International Conference Acousties, Speech, and Signal Proeessing*, Vienna, 2000, pp. 2318-2321.

RETRACTED ARTICLE

[13] M. S. Fu, and O. C. Au, "Improved halftone image data hiding with intensity selection", *IEEE International Symposium on Circuits and Systems*, vol. 10, pp. 243-246, 2001.

[14] S. C. Pei, and J. M. Guo, "High-capacity data hiding in halftone images using minimal error bit searching", In: *IEEE International Conference on Image Processing*, Singapore, Oct. 2004, pp. 3463-3466.

Received: September 16, 2014 Revised: December 23, 2014 Accepted: December 31, 2014

© Dongmei and Ke; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/-licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.

RETRACTED ARTICLE