Steganography Of Images Using Hilbert Curve

DESAI Hardikkumar V.
Assistant Professor,
Naran Lala College of Professional and
Applied Sciences, Navsari, India.
hardik4dreamz@yahoo.com

DESAI Apurva A.
Department of Computer Science,
Veer Narmad South Gujarat University,
Surat, India.
aadesai@vnsgu.ac.in

Abstract

The development of Information and Communication Technology draws significant attention towards security and integrity of data. ICT leads researchers to develop applications which are used to communicate secretly. Steganography is a very useful information hiding technique. The major advantage of using this technique is that the message which was hidden is not easy to detect. Our research has addressed major aspects of Steganography. We develop new algorithm to achieve steganography using Hilbert Curve. The algorithm is developed by keeping in mind that the message which was hidden is not easy to detect by naked eye. We have generated gray scale and RGB histogram for different images to check the distribution of data. Also we checked standard deviation and mean of the images to check the intensity of pixel. The paper also focuses on various steganographic techniques like HIDE and SEEKS, JSteg, OutGuess 0.1, OutGuess 0.2, F3, F4, and F5.

Keywords: - Steganography, Hide&Seek, JSteg, OutGuess0.1, OutGuess0.2, F3, F4, F5.

1. Introduction

Internet is an open resource for all, so this technology is very much useful to transmit data from one end to other very easily and speedily. Therefore information security draws attention of researchers, government agencies; law makers, military, intelligence agencies as well as criminals who require uninterrupted communications. They are interested in
understanding these technologies and their weaknesses, so as to detect and monitor hidden messages.

1.1 Steganography

The term steganography refers to the art of covert communications. The message is embedded within another object known as a cover work, by tweaking its properties. The resulting output is known as stegogramme.

One of the oldest examples of Steganography dates back to around 440 BC in Greek History. Herodotus, a Greek historian from the 5th century BC, revealed some examples of its use in his work entitled “The Histories of Herodotus”. One elaborate example suggests that Histaeus, ruler of Miletus, tattooed a secret message on the shaven head of one of his most trusted slaves. After the hair had grown back, the slave was sent to Aristagorus where his hair was shaved and the message that commanded a revolt against the Persians was revealed [1].

In modern terms, steganography is usually implemented computationally, where cover work such as text files, images, audio files & video files which are tweaked in such a way that a secret message can be embedded within them.

1.2 Hilbert Curve

A Hilbert curve (also known as a Hilbert space-filling curve) shown in figure 1(a) is a continuous fractal space-filling curve first described by the German mathematician David Hilbert in 591, [2] as a variant of the space-filling curves discovered by Giuseppe Peano in 590. [3] He is recognized as one of the most influential and universal mathematicians of the 19th and early 20th centuries. Hilbert discovered and developed a broad range of fundamental ideas in many areas, including invariant theory of geometry.

![Hilbert Curves](image)

Figure 1 (a): Hilbert Curves with ever increasing density

2. Literature Review

We review different Steganographic techniques like Hide and Seek, JSteg, OutGuess 0.1, Out Guess 0.2, F3, F4 and F5 and other related works regarding steganography are presented. We also presented analysis of various steganography techniques.
2.1 Hide and Seek
Steganography is applicable to all data objects that contain redundancy. People often transmit digital pictures over email and other Internet communication, and JPEG is one of the most common formats for images. Moreover, steganographic systems for the JPEG format seem more interesting because the systems operate in a transform space and are not affected by visual attacks [4]. Visual attacks mean that you can see steganographic messages on the low bit planes of an image because they overwrite visual structures; this usually happens in BMP images. Neil F. Johnson and Sushil Jajodia [5] showed that steganographic systems for palette-based images leave easily detected distortions.

2.2 JSteg
The JSteg algorithm was developed by Derek Upham and is essentially as same as a copy of the Hide & Seek algorithm, because it employs sequential least significant bit embedding. In fact, the JSteg algorithm only differs from the Hide & Seek algorithm because it embeds the message data within the LSBs of the DCT coefficients, rather than its pixel values [6].

2.3 OutGuess 0.1
In much the same way that embedding the message data sequentially using the Hide & Seek method was not considered very secure, neither was the fact that the JSteg algorithm embedded in the same fashion. The first version of Outguess, designed by Neils Provos [7], improved the JSteg algorithm by scattering the embedding locations over the entire image according to a PRNG on image.

2.4 OutGuess 0.2
Neils Provos [8] created a revised version of the Outguess 0.1 algorithm, called OutGuess 0.2. It was ensure that the statistical properties of the cover image were maintained after embedding, such that stegograms looks statistically similar to a clean image. This would make it harder for steganalysts to calculate the likelihood that their suspect image is a stegogramme. The algorithm is exactly the same for OutGuess 0.2 as it was for OutGuess 0.1. The difference lies in what happens after the information has been embedded. In OutGuess 0.2, corrections are made to the coefficients such that they appear similar to that of a clean image in terms of frequencies of the values. This is known as statistics aware steganography.

2.5 F3
As an alternative to the OutGuess 0.2 algorithm, AndreasWestfeld designed an algorithm called F3 [9]. It was considered even more secure. The reason for this is that it did not instantiate the same embedding process as the JSteg and OutGuess algorithms. Instead of avoiding embedding in DCT coefficients equal to 1, the F3 algorithm permitted embedding in these regions, at the same time as it would still avoid embedding in zeros and the DCT coefficients. The algorithm still embedded the message data sequentially. Another change with this algorithm was that it did not embed directly in the least significant bits of the DCT coefficients, but instead took the absolute value of the coefficients first, before comparing
them to the message bits. If both the absolute value of the coefficient, and the message bit were the same, then no changes are made. If they are different, then the absolute value of the DCT coefficient is reduced by 1. An implication of this however, is that zero values are often created which the decoding algorithm will not be programmed to extract data from. The F5 algorithm worked around this by reembedding $m_i$ when the result is that a zero DCT coefficient is created.

2.6 F4

The main drawback with F3 was the reality that it effectively embedded more zeros than ones as a result of the shrinkage mechanism. This meant that when the statistical properties of the stegograms are examined through its some object of embedding became visible. This is much the same as what happened in the JSteg implementation except a slightly different pattern is derived. In addition to this, steganalysts also found that more odd coefficients existed in F3 stegograms than even coefficients. This now meant that there were two weaknesses that could be examined when viewing the histogram of a suspect image. F4 was developed to remove these properties such that the histogram would appear similar to that of a clean image.

2.7 F5

The F5 algorithm [10] is predominantly the same as the F4 algorithm, at least in terms of its strategy for encoding the message data. However, the F5 algorithm was designed in an attempt to improve on the F4 algorithm by minimizing the disturbance caused when embedding the message data.

The earliest work was done by Davern and Scott (1996) [11], who divide the domain blocks of the image into two sets. They then use standard fractal image compression techniques to select a domain block that matches the range block, however they choose a block from one of the two domain sets depending on whether the data bit they are embedding is a one or a zero.

A. Jacquin (1995) [12], most fractal coding schemes are based on representation by Partition Iterated Function System (PIFS), a solution to inverse problem which was first published by Jacquin. A PIFS differs from IFS in the Individual mappings operating on a subset of image, rather than the entire image.

Pauate et al, Saupe et al, Wotilberg et al, (1997, 1994, 1999) [13, 14, 15], Theoretical background referred to in fractals to generate fractals considers an IFS consisting of a collection of contractive affine transformations. The iteration procedure of applying a number of transforms is terminated when convergence is met i.e. an attractor.

Bas et al, Li C Wang ,Liao P et al (1998,2000,2006) suggested that a great variety of steganographic methods on fractal compression principles [16,17,18] are good, but greatest robustness is ensured by means of the methods [17,18] since they directly manipulate the code of compressed image. Building in secrecy increase the given approaches will provide high level of protection.
Lokesh Kumar (2012) [19], in his proposed system cryptographic and steganographic security is combined to give two tier securities to secret data. In proposed scheme secret message is encrypted before hiding it into the cover image which gives high security to secret data. Advanced encryption standard (AES) is used to encrypt secret message and alteration component technique is used to hide encrypted secret message into cover image. Since the resulting perceptual quality of the mixed images is good, it is hardly attracted from eavesdropper by naked eye. Finally we can conclude that the proposed technique is effective for secret data communication.

Rosziati Ibrahim et al (2010) [20], in his proposed a new steganography algorithm with 2 layers of security. A system named SIS (Steganography Imaging System) has been developed using the proposed algorithm. They tested few images with various sizes of data to be hidden. With the proposed algorithm, they found that the stego image does not have a noticeable distortion on it (as seen by the naked eyes). SIS can be used by various users who want to hide the data inside the image without revealing the data to other parties. SIS maintains privacy, confidentiality and accuracy of the data.

Wojciech Frączek et al (2010) [21] , suggested that Stream Control Transmission Protocol (SCTP) is a new transport layer protocol that is due to replace TCP (Transmission Control Protocol) and UDP (User Datagram Protocol) protocols in future IP networks. Currently, it is implemented in such operating systems like BSD, Linux, HPUX or Sun Solaris. It is also supported in Cisco network devices operating system (Cisco IOS) and may be used in Windows. This paper describes potential steganographic methods that may be applied to SCTP and may pose a threat to network security. Proposed methods utilize new, characteristic SCTP features like multi-homing and multistreaming. Identified new threats and suggested countermeasures may be used as a supplement to RFC 5062, which describes security attacks in SCTP protocol and can induce further standard modifications.

Hemalatha S (2013) [22], Observes that two secret images can be hidden in one color image and they can be regenerated without actually storing the image. This approach results in high quality of the stegoimage having high PSNR values compared to other methods. However the disadvantage of the approach is that it is susceptible to noise if spatial domain techniques are used to hide the key. This can be improved if transform domain techniques are used to hide the key. The approach is very simple and the security level can be increased by using standard encryption techniques to encrypt the keys.

Sharon Rose Govada et al (2012) [23], he suggested a method is a combination of Word shifting, Text Steganography and Synonym Text Steganography. So we called this as “Three Phase Shielding Text Steganography” This method overcomes various limitations faced by the existing Steganographic algorithms.

Gowtham dhanarasi (2012) [24], suggested a novel method for image steganography using block complexity analysis in wavelet domain. Many researchers have been reported different techniques but all the methods suffer with image quality problem. So in order to achieve good quality,

3. Proposed Algorithm & Flow Chart
We present algorithm to hide the image within image, we develop a new algorithm with the use of Hilbert Curve. The algorithm is developed by keeping in the mind that no one can easily detect the hidden image.
3.1 Embedding algorithm

In this approach, we take three images for the purpose of steganography of images using Hilbert Curve. Manaal image 2(a) is an original image which refers to as cover or carrier image. Tree image 2 (b) is an image to embed within cover image and image 2(c) is an image of Hilbert Curve which is used to identify location on Manaal image 2(a).

![Manaal Image (Cover/Carrier) (b): Tree Image (Image to Hide) (c): Hilbert Curve Image](image)

Hilbert Curve image 2(c) is generated on Manaal image 2(a) to identify locations in Manaal image 2(a) after that, all pixel values of Manaal image 2(a) are changed to 8 bit value 11111111 using bit manipulation technique and last four bit values are changed to 0 using bitor and bitand operations and make the value 11110000.

Now, image 2(b) bit value is changed to 8 bit value 11111111 using bit manipulation technique. Then embed first four bits of image 2(b) to last four bits of Manaal image 2(a) to generate new image 3(a) containing stegogramme which is looks like Manaal image 2(a).

![Manaal Image (extracted)](image)

[Algorithm 1 – Embedding algorithm to hide image using Hilbert Curve]

Step 1: Get the image “A”.
Step 2: Generate Hilbert Curve to identify targeted region in an Image
Step 3: Change the Pixel value of Targeted Region in an Image.
Step 4: Get the image ‘B’ to embed.
Step 5: Change the pixel value of image ‘B’ using bit manipulation.
Step 6: Pad the pixel of image ‘B’ to image ‘A’ till all the pixels are embedded to generate stegogramme ‘A1’
Step 7: End.
3.3 Extracting Algorithm

In the extraction process, we take the embedded image 4(a) and recognize targeted region by generating Hilbert Curve image 4(b) and extract last four bits 1 1 1 1 (5678) of on image.

After getting all the values from embedded image 4 (a), join them together so values are changed to 11111111 to extract the hidden tree image 5(a) from main image 4(a) and compare original, embedded and extracted image.
Step 1: Get the embedded image ‘A1’ containing Stegogramme.
Step 2: Generate Hilbert Curve to define the locations of Stegogramme.
Step 3: Extract the embedded Pixels by changing the value using bit manipulation.
Step 4: Extract the actual Image and cover image A2.
Step 5: End

3.4 Flowchart for Extraction
Flowchart 2: Extraction flowchart to hide image using Hilbert Curve
4. Experimental Results

The results are based on comparison of images, histogram of images in gray scale and in RGB. We examine different images images for steganography. Over here, we analyze the mean and standard deviation of images to check the intensity of pixel.

Firstly, we hide image (b) within image (a) and new image (c) is generated containing stegogramme. Then we compare cover (carrier/cover) image (a) with image (c) to check whether it is embedded successfully or not. We compare histogram in grey scale and in RGB to check the data distribution and analyze mean and standard deviation of cover image, embedded image and post extracted image to check the intensity of pixels. Results are as follows.

![Figure 6](image)

Figure 6 (a): Tree House Image (Carrier/Cover) (b): Sunset (Image to Hide) (c): Tree House-em Image (With Stegogramme)

![Histogram 1](image)

Histogram 1 (a): Tree House Image gray scale (Carrier/Cover) (b): Tree House-em Image gray scale (With Stegogramme) (c): Tree House-ex Image gray scale (Cover Image Post-Extraction)

![Histogram 2](image)

Histogram 2 (a): Tree House Image RGB (Carrier/Cover) (b): Tree House-em Image RGB (With Stegogramme) (c): Tree House-ex image RGB (Cover Image Post-Extraction)
Table 1: Mean and standard deviation of Tree House image in gray scale and in RGB

<table>
<thead>
<tr>
<th></th>
<th>Mean $\bar{X}$</th>
<th>Standard Deviation $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cover</td>
<td>Embedded</td>
</tr>
<tr>
<td>Grey Scale</td>
<td>0.8778</td>
<td>0.8747</td>
</tr>
<tr>
<td>RGB</td>
<td>241.74</td>
<td>243.84</td>
</tr>
</tbody>
</table>

Moreover, we find mean and standard deviation of twenty five images to check the intensity of pixels for both gray scale and in RGB as shown in table 2.

Table 2: Analysis of images for steganography using Hilbert Curve (grey scale)

<table>
<thead>
<tr>
<th>No</th>
<th>Mean $\bar{X}$</th>
<th>Standard Deviation $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cover</td>
<td>Embedded</td>
</tr>
<tr>
<td>1</td>
<td>0.8874</td>
<td>0.8836</td>
</tr>
<tr>
<td>2</td>
<td>0.8563</td>
<td>0.8569</td>
</tr>
<tr>
<td>3</td>
<td>0.8579</td>
<td>0.9039</td>
</tr>
<tr>
<td>4</td>
<td>0.8812</td>
<td>0.9156</td>
</tr>
<tr>
<td>5</td>
<td>0.9237</td>
<td>0.8905</td>
</tr>
<tr>
<td>6</td>
<td>0.8563</td>
<td>0.9408</td>
</tr>
<tr>
<td>7</td>
<td>0.8556</td>
<td>0.8452</td>
</tr>
<tr>
<td>8</td>
<td>0.8547</td>
<td>0.933</td>
</tr>
<tr>
<td>9</td>
<td>0.8596</td>
<td>0.8963</td>
</tr>
<tr>
<td>10</td>
<td>0.9389</td>
<td>0.9256</td>
</tr>
<tr>
<td>11</td>
<td>0.9458</td>
<td>0.8578</td>
</tr>
<tr>
<td>12</td>
<td>0.8612</td>
<td>0.9641</td>
</tr>
<tr>
<td>13</td>
<td>0.8878</td>
<td>0.9785</td>
</tr>
<tr>
<td>14</td>
<td>0.9128</td>
<td>0.9382</td>
</tr>
<tr>
<td>15</td>
<td>0.9378</td>
<td>0.9621</td>
</tr>
<tr>
<td>16</td>
<td>0.9378</td>
<td>0.8426</td>
</tr>
<tr>
<td>17</td>
<td>0.8659</td>
<td>0.9299</td>
</tr>
<tr>
<td>18</td>
<td>0.8759</td>
<td>0.9065</td>
</tr>
<tr>
<td>19</td>
<td>0.9378</td>
<td>0.9746</td>
</tr>
<tr>
<td>20</td>
<td>0.9426</td>
<td>0.9049</td>
</tr>
<tr>
<td>21</td>
<td>0.9412</td>
<td>0.9146</td>
</tr>
<tr>
<td>22</td>
<td>0.8569</td>
<td>0.8565</td>
</tr>
<tr>
<td>23</td>
<td>0.8579</td>
<td>0.8591</td>
</tr>
</tbody>
</table>
We analyze the mean and standard deviation of images with cover image, embedded image and cover image (post extraction) to find the variance between them. After analysis we find that average mean of cover images is 0.8914, average mean of embedded images is 0.9080 and average mean of cover image (post extraction) is 0.8995. Average standard deviation of cover images is 0.2752, average standard deviation of embedded images is 0.2689 and average standard deviation of cover image (post extraction) is 0.2600. It is calculated that, the mean difference between original image and embedded image is 0.0166, mean difference between embedded and cover image (post extraction) is 0.0085 and mean difference between original image and cover image (post extraction) is 0.0081. Further standard deviation difference between original image and embedded image is 0.0063, standard deviation difference between embedded and cover image (post extraction) is 0.0089 and standard deviation difference between original image and cover image (post extraction) is 0.0152.

Table 3: Analysis of images for steganography using Hilbert Curve (RGB)

<table>
<thead>
<tr>
<th>No</th>
<th>Cover</th>
<th>Embedded</th>
<th>Cover Image Post-Extraction</th>
<th>Cover</th>
<th>Embedded</th>
<th>Cover Image Post-Extraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>247.88</td>
<td>247.52</td>
<td>248.12</td>
<td>36.78</td>
<td>36.72</td>
<td>36.61</td>
</tr>
<tr>
<td>2</td>
<td>239.72</td>
<td>240.62</td>
<td>247.66</td>
<td>38.68</td>
<td>38.18</td>
<td>37.82</td>
</tr>
<tr>
<td>3</td>
<td>248.86</td>
<td>246.88</td>
<td>247.93</td>
<td>34.22</td>
<td>39.88</td>
<td>35.42</td>
</tr>
<tr>
<td>4</td>
<td>247.69</td>
<td>247.19</td>
<td>247.22</td>
<td>39.41</td>
<td>36.43</td>
<td>36.32</td>
</tr>
<tr>
<td>5</td>
<td>241.04</td>
<td>247.66</td>
<td>247.67</td>
<td>51.42</td>
<td>39.88</td>
<td>37.88</td>
</tr>
<tr>
<td>6</td>
<td>248.73</td>
<td>247.69</td>
<td>247.87</td>
<td>34.51</td>
<td>36.39</td>
<td>36.48</td>
</tr>
<tr>
<td>7</td>
<td>242.87</td>
<td>245.41</td>
<td>247.51</td>
<td>36.78</td>
<td>36.29</td>
<td>37.12</td>
</tr>
<tr>
<td>8</td>
<td>245.46</td>
<td>245.56</td>
<td>249.32</td>
<td>36.21</td>
<td>36.72</td>
<td>38.73</td>
</tr>
<tr>
<td>9</td>
<td>249.56</td>
<td>249.78</td>
<td>249.68</td>
<td>44.98</td>
<td>44.81</td>
<td>44.31</td>
</tr>
<tr>
<td>10</td>
<td>242.78</td>
<td>242.62</td>
<td>242.65</td>
<td>36.95</td>
<td>36.72</td>
<td>36.18</td>
</tr>
<tr>
<td>11</td>
<td>242.32</td>
<td>242.18</td>
<td>246.37</td>
<td>38.15</td>
<td>46.68</td>
<td>46.43</td>
</tr>
<tr>
<td>12</td>
<td>246.78</td>
<td>246.68</td>
<td>246.65</td>
<td>38.82</td>
<td>35.19</td>
<td>35.46</td>
</tr>
<tr>
<td>13</td>
<td>245.78</td>
<td>245.68</td>
<td>247.88</td>
<td>34.18</td>
<td>34.78</td>
<td>34.83</td>
</tr>
<tr>
<td>14</td>
<td>248.78</td>
<td>247.89</td>
<td>247.78</td>
<td>47.12</td>
<td>48.96</td>
<td>48.12</td>
</tr>
<tr>
<td>15</td>
<td>241.56</td>
<td>242.86</td>
<td>242.96</td>
<td>42.78</td>
<td>34.46</td>
<td>34.12</td>
</tr>
<tr>
<td>16</td>
<td>250.78</td>
<td>249.88</td>
<td>248.76</td>
<td>42.51</td>
<td>34.09</td>
<td>33.46</td>
</tr>
<tr>
<td>17</td>
<td>244.81</td>
<td>245.71</td>
<td>245.81</td>
<td>41.88</td>
<td>49.66</td>
<td>48.53</td>
</tr>
<tr>
<td>18</td>
<td>249.52</td>
<td>249.69</td>
<td>249.86</td>
<td>37.89</td>
<td>37.83</td>
<td>35.96</td>
</tr>
</tbody>
</table>
We analyze the mean, standard deviation of images between cover image, embedded image and cover image (post extraction) to find the variance between them. After analysis of twenty five different images we find that average mean of cover images is 245.76, average mean of embedded images is 246.06 and average mean of cover image (post extraction) images is 246.86. Average standard deviation of cover images is 40.66, average standard deviation of embedded images is 40.19 and average standard deviation of cover image (post extraction) is 39.61. It is calculated that, the mean difference between original image and embedded image is 0.2964, mean difference between embedded and cover image (post extraction) is 0.7988 and mean difference between original image and cover image (post extraction) is 1.0952. Further standard deviation difference between original image and embedded image is 0.468, standard deviation difference between embedded and cover image (post extraction) is 0.5732 and standard deviation difference between original image and cover image (post extraction) is 1.0412.

5. Conclusions

It was concluded that the image containing stegogramme is not recognized by the naked eye detection, there is not much difference between cover, embedded and cover (post-extraction) image. In the analysis we find mean and standard deviation and concluded that there is negligible difference amongst them and not easy to detect the stegogram.

6. Limitations and Future Work

The research provides an opportunity to go ahead in the field of Information Hiding. The research can be extended by using other fractals for the purpose of steganography.

The limitation of our technique is that the size of image to hide should be equal to the size of the fractal used. It may be attempted to overcome.

Moreover, the research may be extended by using different carriers (covers) like audio and video. One of the methods for audio steganography is to substitute the least significant bit of each sample of the cover speech signal with the secret data. While this method is easy to implement and can be used to hide larger secret messages, it cannot protect the hidden message from small modifications that can happen as a result of format conversion or compression, so further advancement is needed. Moreover, in the video steganography generally video is separated into audio and images. Therefore any of them is selected as cover /carrier and data are embedded within them using LSB or any other technique. The research can be furthered by using cryptography with steganography to enhance the security.
because by using cryptography, we can convert our plain text message to cipher text by using the secret key and embed the cipher text into any cover medium using steganographic techniques.

References