ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 9, Issue, 06, pp.51810-51817, June, 2017

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

RESEARCH ARTICLE

SURVEY OF STEGANOGRAPHIC TECHNIQUES FOR COMPRESSED MEDIA

***Deepak Kumar, Aniruddha Singh and Manish Verma**

Department of CSE, GU, Greater Noida, UP, India

ARTICLE INFO

Article History:
Received 28th March, 2017
Received in revised form
08th April, 2017
Accepted 10th May, 2017
Published online 20th June, 2017

Key words:

Steganography,
Information Hiding,
Cover Image, Secret Image,
Huffman, LZW.

ABSTRACT

Steganography is the old craft of concealing messages. The word steganography originates from Greek inception and signifies "secured or shrouded expressing" in which hiding information is implanted into a blameless looking transporter. The transporter might be a content record, advanced picture, sound document or a video document. After installed information is exchanged crosswise over correspondence diverts or posted openly territories. In this way, the transporter ought to appear to be blameless under generally examinations. There are a few systems of steganography; some utilization the minimum critical bits of the picture information to conceal the hiding message, some shroud the data in a particular band of the spatial recurrence segment of the transporter, some of them make utilization of the testing mistake in picture digitization. Every one of these procedures are constrained as far as data concealing ability To build the inserting limit we need a steganography system in view of the compressed lossless steganography. In this paper we survey various techniques in which information hidden using compression over the cover picture.

Copyright©2017, Deepak Kumar et al. *This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.*

Citation: Deepak Kumar, Aniruddha Singh and Manish Verma, 2017. "Survey of steganographic techniques for compressed media", *International Journal of Current Research*, 9, (06), 51810-51817.

INTRODUCTION

The proliferation of network technology and digital devices makes the transport of digital multimedia fast and clean. But, distributing virtual information over public networks together with net is not dependable because of copyright violation, counterfeiting, forgery, and fraud. Therefore, techniques for shielding virtual records, in particular sensitive information, are extraordinarily important (Chang and Kieu, 2010). Despite the fact that the usage of electronic documents is big, only a few humans can recognize that those documents include "hid-den facts". The motive for the usage of the phrase "hidden" is that those facts are normally located within a file, but can't be diagnosed the use of common strategies. Hidden records can be classified into two types. The first is automatically created by the utility, and the second one is created and concealed by way of an individual for precise purposes (Park and Lee, 2009). Secret facts may be included through cryptographic methods, conventionally. But, transmitting the encrypted mystery records by using cryptosystems is unlawful through a few dictatorial governments, or the meaningless form of the encrypted information might also attract the eye of interceptors (e.g., wardens or sensors) which can be designed to prevent any mystery communications (Chang and Kieu, 2010). Instead, exclusive information can be protected by means of employing records hiding strategies.

Corresponding autour:* **Deepak Kumar,
Department of CSE, GU, Greater Noida, UP, India.

Generally, statistics hiding consists of digital watermarking and steganography (Chang and Kieu, 2010). Watermarking isn't the same as steganography in its predominant aim. Watermarking is used for copyright safety, broadcast tracking, transaction tracking, and comparable activities. A watermarking scheme alters a cover object, either imperceptibly or perceptibly, to embed a message approximately the quilt object (e.g., the owner's identifier). it could be determined as steganography that is concentrating on high robustness and really low or nearly no protection (Gutub and Fattani, 2007). In comparison, steganography is used ordinarily for secret communications (Chang and Kieu, 2010). Steganography is the art of writing mystery data in this sort of way that no person except the intended receiver is aware of approximately the lifestyles of secret records. a success steganography relies upon upon the service medium not to elevate interest (Sajedi and Jazzed, 2010). There are three essential troubles to be taken into consideration whilst analyzing steganography structures: potential (or bitrate), security and robustness (Al-Haidari *et al.,* 2009). Capacity refers to the amount of facts bits that can be hidden in the cover medium. Protection relates to the ability of an eavesdropper to determine the hidden information without problems. Robustness is worried approximately the withstand opportunity of editing or destroying the unseen facts (Gutub and Fattani, 2007). In stegno- raphy for virtual structures, the duvet media used to cover the message may be text, picture, video or audio? les (Aabed *et al.,* 2007).

We suggest a compression based totally textual content steganography technique as a way to improve capacity and security. Specifically, the trouble is to attain a widespread increment in the amount of secret data that is aimed to be hidden in cover medium at the same time as we choice to complicate the extraction technique of the secret facts. Within the proposed technique, mystery information has been embedded within the selected text from the previously constructed textual content base. The textual content base includes naturally generated texts like notification texts, abstracts of articles, and so on. Which can be used for a collection speech. While embedding, originality of the selected textual content has been protected through handiest camouflaging the secret information. E-mail has been chosen as verbal exchange channel among the two parties, so the stego cover has been arranged as a ahead mail platform. Even as arranging the stego cowl as a forward mail platform, we use the previously organized e mail address listing for choosing the email addresses. Meanwhile, this e mail address list has been used as a worldwide stego key that is shared both by means of the sender and the recipient beforehand.

For the first cause, capacity increment, we favor to use records compression techniques. In a statistics compression system the goal is to lower the redundancy of a given statistics description (Gal ambos and Bekesi, 2002). Typically, records compression algorithms are labeled as lossless or loss. Lossless statistics compression involves a metamorphosis of illustration of the authentic statistics set such that it's miles possible to breed exactly the original statistics set by way of in step with-forming a decompression transformation and it's miles used when the original and the decompressed? les have to be same (in com- urgent textual content files, executable codes, phrase processing ?les, and so on.). Loss statistics compression includes a change of representation of the unique records set such that it's far impossible to breed precisely the original records set, but an approximate illustration is reproduced by means of performing a decompression transformation. This type is used on the net and specifically in streaming media and telephony packages (Al-Bahadili, 2008). In case of textual facts, at the same time as performing a compression/ decompression procedure we need to recover exactly the authentic statistics. In case of photographs or voices – without stepping into deep hassle – it is allowed to get an approximation of the authentic data (Galambos and Bekesi, 2002). In our problem, we should defend the originality because of dealing with textual statistics. So we ought to use a lossless records compression method. For that reason, we advocate to hire LZW facts compression set of rules due to its true compression ratio and frequent usage within the literature. The LZW algorithm first reads the records and tries to match a chain of records bytes as big as possible with an encoded string from the dictionary. The matched information sequence and its succeeding person are grouped collectively and then brought to the dictionary for encoding later records sequences (Liang *et al.,* 2008). For the second reason, security development, we recommend to appoint stego-keys. We are able to classify the hired stego keys into two instructions consistent with their missions. one among them is the constructed stego keys at some point of embedding segment of the seasoned- posed scheme and the alternative is the previously constructed worldwide stego key which is shared both with the aid of the sender and the receiver ahead. Meanwhile, employing Combinatory-based coding with a purpose to guide the favored randomness (see Jun *et al.,* 2011 for added facts) affords a fantastic contribution to the security. Combinatory-primarily based coding is predictable to the receiver but pretty random to an observer who attempts to investigate the steganography cover, rendering the steganography cowl extra resilient (Desoky, 2009). For this reason, Latin rectangular has been employed (see Easton and Gary Parker, 2001 and Colbourn, 1984 for additional information). By way of basing on Bailey and Curran (2006), we are able to say that LZW coding also provides contribution to safety. Assessment process has been performed through capacity measurements. Ability has been measured in terms of percentage, via calculating the rate of secret records this is embedded inside the stego cowl. Except, a popular evaluation has been executed in phrases of capacity by means of comparing the proposed scheme with the opposite contemporary methods in the literature.

Compression algorithms

Compression is the combination of additives. One is encoding algorithm, another one is deciphering set of rules. In encoding set of rules makes the message as compressed illustration. In deciphering algorithm reconstructs the message from compressed representation to unique message or it reconstructs some approximation. Compression algorithms are categorized into categories lossless algorithms reconstruct unique message from compressed message. Lossless compression is used for textual content; loss compression is used for pictures and sound (Jafari *et al.,* 2013; Bashardoost *et al.,* 2013; Lin *et al.,* 2014). Textual content compression is one technique to growth the performance of textual content compression. Enter textual content may be modified a incredibly redundant text via using pre-described distinctly redundant codes in place of words or terms. This high redundant textual content will increase the overall performance of the text compression set of rules. The already current arithmetic coding,

Huffman coding, LZ set of rules, PPMC, RLE cannot deliver better compression ratios (Satir *et al.,* 2014; Zhang *et al.,* 2015). Better compression ratio is finished via the usage of dictionary based totally compression. Steganography, from the Greek, manner blanketed or mystery writing, and is a long-practiced form of hiding statistics. Despite the fact that related to cryptography, they're now not the equal. Steganography rationale is to hide the lifestyles of the message, while cryptography scrambles a message so that it cannot be understood. greater precisely, “the intention of Steganography is to cover messages inner other innocent messages in a manner that does not allow any enemy to even detect that there's a second secret message present. “Steganography includes a full-size array of techniques for hiding messages in a spread of media. among those techniques are invisible inks, microdots, digital signatures, covert channels and unfold-spectrum communications. these days, thanks to trendy technology, steganography is used on textual content, photos, sound, indicators, and extra. „ cover? is an audio, picture, video soon that's used to hide the authentic message. The duvet signal used in the system of steganography is referred to as the „host alerts?. Records hidden in cover information is referred to as embedded data (Chetan *et al.,* 2015; Garg *et al.,* 2016). There is no essential to encrypt the hidden message .but it depends on the safety of the device, the layout of the complete understanding of it. The benefit of steganography is that it could be used to secretly transmit messages without the reality of the transmission being found. Frequently, the usage of

encryption may perceive the sender or receiver as somebody with something to hide.

Background of textual content Compression And Steganography

Lossless compression researchers have advanced rather sophisticated techniques, together with Huffman encoding, mathematics encoding, the Lempel-Ziv (LZ) circle of relatives, Dynamic Markov Compression (DMC), Prediction by way of Partial Matching (PPM),Run period coding(RLE) and Burrows-Wheeler remodel (BWT) based algorithms .[5,12,14] However, none of those strategies has been capable of reach the theoretical pleasant-case compression ratio continually. Dictionary based Encoding (DBE) technique for looking to obtain better compression ratios is to develop new compression algorithms. To be able to growth the secrecy of the text message compressed by dictionary based compression, its miles hidden in the audio file. If the text message is hidden using stenographic system, it could be detected through attackers. To keep away from this, the enter message can be transformed into fairly redundant code and then hidden. This method will help hold secrecy.

Dictionary making algorithms

- Calgary corpus files are taken as take a look at text files.
- Phrases are accumulated from all text documents. this is 6, 18,108 variety of phrases.
- In this word, letters in uppercase are transformed into letters in lowercase.
- Four to shape the dictionary, phrases are indexed in descending order after locating how normally each word occurs.
- 8900 words have been indexed inside the trendy dictionary
- For the primary 169 phrases, single ASCII man or woman is assigned as code.
- For the words from one hundred seventy to 4300, single 169ASCII character with every uppercase letter is assigned as double codes.
- For the ultimate words, with each top case letter previous person combination is coded.

Hiding the compressed text in audio will decorate the security. as compared to other text compression algorithm dictionary primarily based audio steganography machine gives better value.

A. LSB Insertion technique

- Audio file is converted into the statistics samples.
- First 40 bytes are allotted for header element.
- The compressed textual content message is transformed as binary.
- The length of textual content message is likewise transformed as binary.
- The identifier is chosen to cover the textual content message.
- An identifier enables in the restoration of textual content.
- If there's no identifier in audio document, audio document no hidden text message.
- The identifier's binary is 10101010.
- Identifier can be hidden in eight data samples.
- The next 10 data samples will function the length of textual content message.
- The subsequent 10 information samples can be as the width of text message.
- The compressed text message in the remaining records samples lsb is to be hidden.

B. Statistics Extraction process

- Text can be recovered in a reverse manner of the way the text is hidden.
- Now check the obtained audio report whether or not identifier present or now not.
- Without identifier, there can be no hidden textual content in records samples.
- Both the duration and width of the text message from the records samples lsb are to be measured.
- The lab little bit of information samples ought to be taken till the period of the message is obtained
- Then the message inside the lsb bit is to be transformed into text. deciphering algorithm

The deciphering is easier than the encoding. Top case letters followed through unmarried ASCII character is identified as a code. If top case letters are observed with the aid of two ASCII characters, the second one ASII man or woman is diagnosed as separate code. Extracted code is as compared with dictionary desk and corresponding phrases are gathered in the output document. This output document after processing seems the same as the preliminary report since the compression and decompression is lossless.

Purpose of Cryptography

Cryptography affords some of safety desires to ensure the privacy of statistics, non-alteration of records and so forth. Because of the excellent security blessings of cryptography it's far broadly used today. Following are the diverse goals of cryptography.

1. **Confidentiality:** Information in pc is transmitted and needs to be accessed most effective with the aid of the legal birthday party and now not by means of absolutely everyone else.
2. **Authentication:** The facts obtained by way of any system has to check the identification of the sender that whether or not the records is getting back from a licensed character or a false identification.
3. **Integrity:** Only the authorized birthday party is authorized to regulate the transmitted records. No person in between the sender and receiver are allowed to alter the given message.
4. **Non Repudiation:** Guarantees that neither the sender, nor the receiver of message ought to be able to deny the transmission.
5. **Get admission to manage:** Only the legal events are able to get entry to the given records.

Asymmetric Encryption

Uneven cryptography or public-key cryptography is cryptography in which a couple of keys is used to encrypt and

decrypt a message so that it arrives securely. To begin with, a community consumer receives a public and personal key pair from certificate authority. Every other person who wants to ship an encrypted message can get the supposed recipient's public key from a public directory. They use this key to encrypt the message, and that they send it to the recipient. When the recipient gets the message, they decrypt it with their non-public key, which no person else have to have get right of entry to too.

Diffusion and Confusion

Shannon, in one of the fundamental papers at the theoretical foundations of cryptography [1, 2], gave residences that a terrific cryptosystem need to have to prevent statistical evaluation: diffusion and confusion. Diffusion method that if we trade a character of the obvious textual content, then numerous characters of the cipher textual content must trade, and similarly, if we exchange a individual of the cipher textual content, then numerous characters of the obvious text have to trade. Which means that frequency information of letters in the apparent textual content are subtle over several characters inside the cipher textual content, this means that that rather more cipher textual content is wanted to do a meaningful statistical attack. Confusion means that the key does now not relate in a simple way to the cipher text. Especially, every character of the cipher text have to depend upon several components of the important thing.

Discrete Cosine Transform (DCT)

The Discrete cosine rework (DCT) is maximum popular due to several motives. One of the purpose is that maximum of the compression techniques developed inside the DCT area & consequently photo processing is extra acquainted with it. DCT is one of the maximum commonplace linear changes in digital image manner technology. Two-dimensional discrete cosine transform (2d-DCT) is defined as

$$F(jk) = a(j)a(k)\sum_{m=0}^{N}\sum_{n=0}^{N-1} f(mn)cos\left[\frac{(2m+1)j\pi}{2N}\right]cos\left[\frac{(2n+1)k\pi}{2N}\right]$$

The corresponding inverse transformation (Whether 2DIDCT) is defined as

$$F(mn) = \sum_{m=0}^{N}\sum_{n=0}^{N-1} a(j)a(k)f(jk)cos\left[\frac{(2m+1)j\pi}{2N}\right]cos\left[\frac{(2n+1)k\pi}{2N}\right]$$

The 2nd-DCT cannot best listen the main statistics of original photo into the smallest low frequency coefficient, however additionally it is able to purpose the image blocking effect being the smallest, that may recognize the best compromise among the records centralizing and the computing hassle. The DCT lets in an image to be damaged up into one of a kind frequency bands, making it an awful lot easier to embed steganography information into the middle frequency bands of an image. So that you can invisibly embed the steganograph that may continue to exist lossy statistics compressions, a reasonable tradeoff is to embed the steganograph into the middle-frequency range of the photo. The middle frequency bands are chosen such that they've minimized that they avoid the maximum visual essential elements of the picture (low frequency) without over-exposing themselves to removal through compression and noise assaults. DCT area steganography can live to tell the tale against the assaults including noising, compression, polishing, and filtering.

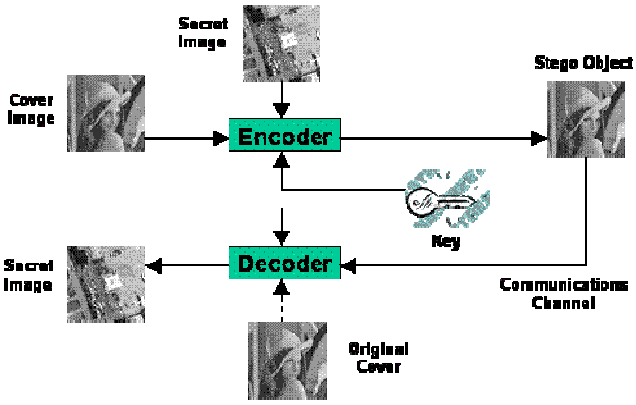

Fig.1. Steganograph Embedding and Detection in DCT

Steganography applications

Despite the fact that the important application of virtual steganography is to defend the copyright, however its applications are not that constrained. It has a huge variety of applications. a few essential applications are:

1. **Broadcast tracking:** we are able to use digital steganography to display that how regularly a specific advertisement has been broadcasted [30,31]. In broadcast tracking the machine gets the broadcast. Then the system searches for the detection of steganographs and identifies whilst, wherein and the way generally this paintings /advertisement is broadcasted. TV is an example in which news includes watermarked videos from broadcasters.
2. **Proprietor identity:** in this software steganography is used to affirm the proprietor. The writer of paintings which include songs, eBook, and portray keep the copyright as quickly as it's miles published /revealed. Textural copyright notices were used however they've some obstacle. it could be without problems eliminated from a file after which be copied even by way of those who don't have any incorrect intentions. Because the steganography may be hid imperceptibly in to the work, it can become aware of the owner of steganography higher than the textual form of proprietor identity.

1. Clinical applications

Steganography is used to perceive the scientific x-ray photographs and different facts of sufferers thereby decreasing the chances of tampering of the scientific data.

Requirements for a Steganography set of rules

the main objectives for any steganography set of rules are potential, undetectibility and robustness. Even though it is tough for a steganography set of rules to have all of the traits at the equal time because there's typically trade-off between those characterstics.

- **Capability:** The amount of facts to be embedded in cover medium and can retrieved later effectively without appreciably changing the duvet medium.

- **Undetectibility:** There have to be no visual difference among cover and stego object i.e. embedded message must now not be visible to human eye.
- **Robustness:** A stego gadget is stated to be robust if it can endure any attack and if it undergoes transformation such as scaling, rotation, filtering and lossy compression and so on. it must remain intact.
- **Protection:** An embedding set of rules is said to be relaxed if the embedded information could not be eliminated after detection by way of the attacker. It relies upon at the knowledge approximately the embedded set of rules and mystery key.

Related Work

Jafari, Reza *et al.* (2013) In this paper, the goal of image compression is to remove the redundancies for minimizing the number of bits required to represent an image while steganography works by embedding the secret data in redundancies of the image in invisibility manner. Our focus in this paper is the improvement of image compression through steganography. Even if the purposes of digital steganography and data compression are by definition contradictory, we use these techniques jointly to compress an image. Hence, two schemes exploring this idea are suggested. The first scheme combines a steganography algorithm with the baseline DCT-based JPEG, while the second one uses this steganography algorithm with the DWT-based JPEG. In this study data compression is performed twice. First, we take advantage of energy compaction using JPEG to reduce redundant data. Second, we embed some bit blocks within its subsequent blocks of the same image with steganography. The embedded bits not only increase file size of the compressed image, but also decrease the file size further more. Experimental results show for this promising technique to have wide potential in image coding.

Bashardoost, Morteza *et al.* (2013) In this paper, the challenge of steganography methods is to create a rational balance between the quality of the file and the size of data that can be transferred. In addition, the robustness of the technique and security of the obscure data are the facts that cannot be dissembled. The Least Significant Bit (LSB) insertion approach provides a high degree of visual quality and a large amount of capacity for the concealed data, but the covert message is not well protected in this method. In the proposed method, the secret data is firstly encoded by using the Vigenere encryption method to guarantee the protection of the hidden message. Afterward, the Lempel Ziv Welch (LZW) technique compresses the data to reduce the occupational capacity of the confidential data. Then, by utilizing the extended knight tour algorithm, each bitstream of the data is spread out on the image to increase the robustness of the method. The results show that the proposed method not only improves the security and payload capacity problems of the simple LSB method, but also increases the visual quality of the stego image.

Lin, Chi-Yuan *et al.* (2014) In this paper, a steganography system for VQ codebooks using section-based informed embedding (SBIE) is presented. The goal is to offer a stego-image for secure communication. In the proposed scheme, the secret image is first compressed by a unsupervised fuzzy competitive learning network clustering technology (named FCLN) based on VQ, then a section-based informed embedding algorithms are presented to provide the mechanism of the hiding system for adjustable robustness and fidelity performance. The FCLN generate optimal codebook for VQ. Then we embedded VQ codebook of secret image information into the cover image by a section-based informed embedding scheme. Finally, the experimental results demonstrate two objectives: (1) the promising codebook can be obtained using our proposed image compression scheme FCLN method, and (2) the high fidelity performance offered by the presented SBIE algorithm.

Satir, Esra *et al.* (2014) In this paper, capacity and security issues of text steganography have been considered by proposing a compression based approach. Because of using textual data in steganography, firstly, the employed data compression algorithm has to be lossless. Accordingly, Huffmann coding has been chosen due to its frequent use in the literature and significant compression ratio. Besides, the proposed method constructs and uses stego-keys in order to increase security. Secret information has been hidden in the chosen text from the previously constructed text base that consists of naturally generated texts. Email has been chosen as communication channel between the two parties, so the stego cover has been arranged as a forward mail platform. As the result of performed experiments, average capacity has been computed as 7.962 % for the secret message with 300 characters (or 300·8 bits). Finally, comparison of the proposed method with the other contemporary methods in the literature has been carried out

Zhang, Yi *et al.* (2015) In this paper, current typical adaptive Steganography algorithms cannot extract the embedded secret messages correctly after compression. In order to solve this problem, a JPEG-compression resistant adaptive steganography algorithm is proposed. Utilizing the relationship between DCT coefficients, the domain of messages embedding is determined. The modifying magnitude of different DCT coefficients can be determined according to the quality factors of JPEG compression. To ensure the completely correct extraction of embedded messages after JPEG compression, the RS codes is used to encode the messages to be embedded. Besides, based on the current energy function in the PQe steganography and the distortion function in J-UNIWARD Steganography, the corresponding distortion value of DCT coefficients is calculated. With the help of that, STCs is used to embed the encoded messages into the DCT coefficients, which have a smaller distortion value. The experimental results under different quality factors of JPEG compression and different payloads demonstrate that the proposed algorithm not only has a high correct rate of extracted messages after JPEG compression, which increases from about 60% to nearly 100% comparing with J-UNIWARD steganography under quality factor 75 of JPEG compression, but also has a strong detection resistant performance.

Chetan, Er *et al.* (2015) In this paper, image compression plays an important role in our day to day activities. Image compression is the process of reducing the amount of data required to represent given quantity of information in image to reduce storage requirements and many other reasons. In a computer image is represented as an array of numbers, integers and is known as digital image. Image array is mainly of two dimensional and three dimensional, if the image is two dimensional than the image is black and white and if image is color than it is three dimensional. Fractal Image Compression

is an approach for better image compression. The main objective of this method is to provide simple and better compression results, which is based on proposed Quad tree Decomposition and Discrete Wavelet Transform method for a color image. Fractal image compression can be obtained by dividing the cyber image into overlapped blocks depending on a quantization value and the well-known techniques of Quad tree decomposition. The compression ratio (CR) and Peak Signal to Noise Ratio (PSNR) values are determined for three types of images namely standard Lena image, Baboon image and Pepper image.

Garg, Nancy *et al.* (2016) In this paper, the secure data storage on cloud environments is the primary requirement of such applications, where data are being transferred or transmitted between the servers and their users. The data security is quite important because they belongs the users. With an internet based development and use of computer technology several trends are opening up in the era of cloud computing .Moving data into the cloud offers much ease to users since they don't have to care about the complexities of managing hardware directly. Steganography is the practice of hiding a content of a file, a message, image, or video within other file, message, image, or video. Generally, the hidden messages appear to be part of other: shopping lists, articles images. When steganography is combined with encryption it provides security. In this paper two approaches which include steganography along with encryption is presented for security of data storage on cloud.

Kumar, Rajeev *et al.* (2016) In this paper, we propose a high capacity text steganography method using Huffman compression. The forward email platform is used to hide the secret data. We make use of the number of characters used in email id to indicate the hidden secret data bits. So, to make optimal utilization of number of characters in email ids, the characters added to the email id to indicate the secret data bits are taken from the processed secret data. Hence, the hiding capacity is further increased. The new characters are appended just before the `@' symbol of email ids. Experimental results show that our method performs better than the some important existing methods in terms of hiding capacity.

Malik, Aruna *et al.* (2017) In this paper, capacity and security issues of text steganography have been considered by employing LZW compression technique and color coding based approach. The proposed technique uses the forward mail platform to hide the secret data. This algorithm first compresses secret data and then hides the compressed secret data into the email addresses and also in the cover message of the email. The secret data bits are embedded in the message (or cover text) by making it colored using a color coding table. Experimental results show that the proposed method not only produces a high embedding capacity but also reduces computational complexity. Moreover, the security of the proposed method is significantly improved by employing stego keys. The superiority of the proposed method has been experimentally verified by comparing with recently developed existing techniques.

Zhang, Yi *et al.* (2017) In this paper, in order to improve the JPEG compression resistant performance of the current steganography algorithms resisting statistic detection, an adaptive steganography algorithm resisting JPEG compression and detection based on dither modulation is proposed. Utilizing the adaptive dither modulation algorithm based on the quantization tables, the embedding domains resisting JPEG compression for spatial images and JPEG images are determined separately. Then the embedding cost function is constructed by the embedding costs calculation algorithm based on side information. Finally, the RS coding is combined with the STCs to realize the minimum costs messages embedding while improving the correct rates of the extracted messages after JPEG compression. The experimental results demonstrate that the algorithm can be applied to both spatial images and JPEG images. Compared with the current S-UNIWARD steganography, the message extraction error rates of the proposed algorithm after JPEG compression decrease from about 50 % to nearly 0; compared with the current JPEG compression and detection resistant steganography algorithms, the proposed algorithm not only possesses the comparable JPEG compression resistant ability, but also has a stronger detection resistant performance and a higher operation efficiency.

Challenges and issues with Steganography

After analyzing and studying to be had literature and existing strategies, it changed into discovered that steganography algorithms are going through numerous challenges and issues that call for further exploring and investigations. A number of the prominent problems and regions are as follows:

- Data hiding in still photograph poses diverse challenges as this offer less redundancy and imperceptibility compared to audio and video files.
- It is also a undertaking to embed message into organization photographs, which are noticeably inter correlated and regularly manipulated in compressed form.
- Steganography algorithms commonly battle for presenting excessive information price and imperceptibility. If a technique provides excessive payload potential then it could become less robust and vice versa. Necessities for higher ability and comfortable conversation are often contradictory. depending upon the precise application this alternate off wishes to are searching for out and at the identical time there is also want to produce high pleasant stego algorithm by means of accomplishing excessive fee of PSNR (peak signal to Noise Ratio).
- Steganography strategies are very sensitive to various modifications in cowl medium like photograph processing operations (smoothing, filtering, photograph transformations and so forth.) compression strategies, removing and filtering virtual noise techniques because those techniques cause elimination or adjustments of secret embedded information too. There may be also need to design steganography algorithms able to bearing photo processing operations.
- Hidden message need to be at ease each from perceptual and statistical assaults. There's requirement to design greater sturdy steganography algorithms and there's want to pay special attention for the presence of active and malicious attacks.
- Steganography has various useful applications however like different technologies, criminals and terrorists also can misuse it for unwell functions. there is want to recognize all steganography as well as steganalysis are

concepts, practices and its programs for social purposes as a substitute then unwell functions.

Statistical Compression strategies

Lossless statistical facts compression algorithms have generated a number of interest during the last two decades or so, typically outperforming older algorithms consisting of LZ77 with the aid of a big margin. those algorithms are often sequential: they technique a facts circulate from beginning to cease, incrementally building and redefining a model based totally on statistics that has been processed, without the want to get entry to symbols in addition upstream than the primary un-encoded one. Building a compressor on this way has some of critical benefits. first off, all records that the compressor makes use of is available to the decoder as properly, so there may be no need to output greater records for the decoder a good way to discover how the facts are encoded. Secondly, the size of the document that wishes to be compressed does no longer need to be acknowledged in advance and may be arbitrarily huge, as one may think to be the case for a facts channel among computer systems on a network (the compression serving to enhance bandwidth).

1. Run length Encoding techniques (RLE)

One of the best compression techniques referred to as the Run-duration Encoding (RLE)[2] is created especially for facts with strings of repeated symbols (the period of the string is referred to as a run). the main concept behind RLE is to encode repeated symbols as a pair: the duration of the string and the image. For example, the string 'abbaaaaabaabbbaa' of length 16 bytes (characters) is represented as 7 integers plus 7 characters, which may be without problems encoded on 14 bytes (as for instance '1a2b5a1b2a3b2a'). The biggest problem with RLE method is that within the worst case the dimensions of output records can be instances greater than the scale of input records. To do away with this hassle, every pair (the lengths and the strings one by one) can be later encoded with an set of rules like Huffman coding.

2. Huffman coding

The Huffman coding algorithm [31] is called after its inventor, David Huffman, who evolved this set of rules as a scholar in a category on statistics idea at MIT in 1950. It's far a greater a success technique used for textual content compression. The Huffman's idea is to replace fixed-duration codes (together with ASCII) by means of variable-length codes, assigning shorter code words to the extra regularly occurring symbols and for this reason reducing the overall period of the statistics. When using variable-length code phrases, its miles proper to create a (uniquely decipherable) prefix-code, keeping off the need for a separator to decide code phrase obstacles. Huffman coding creates the sort of code. The Huffman algorithm is straightforward and may be defined in terms of making a Huffman code tree.

The procedure for constructing this tree is:

- Start with a listing of unfastened nodes, in which each node corresponds to a symbol in the alphabet.
- Pick two free nodes with the lowest weight from the list.
- Create a parent node for those nodes selected and the weight is equal to the load of the sum of two infant nodes.
- Eliminate the two baby nodes from the listing and the figure node is delivered to the list of unfastened nodes.
- Repeat the manner starting from step-2 until handiest a single tree stays.

After building the Huffman tree, the set of rules creates a prefix code for every symbol from the alphabet absolutely by using traversing the binary tree from the root to the node, which corresponds to the image. It assigns 0 for a left department and 1 for a proper department. The set of rules supplied above is called as a semi adaptive or semi-static Huffman coding as it calls for expertise of frequencies for every image from alphabet [6]. Together with the compressed output, the Huffman tree with the Huffman codes for symbols or just the frequencies of symbols which can be used to create the Huffman tree ought to be saved. This statistics is needed at some point of the deciphering method and it's miles positioned in the header of the compressed document.

3. Adaptive Huffman coding

The basic Huffman algorithm suffers from the downside that to generate Huffman codes it calls for the opportunity distribution of the enter set that's often now not available. Moreover it isn't suitable to cases while chances of the input symbols are converting. The Adaptive Huffman coding method changed into developed primarily based on Huffman coding first by using Newton Faller [30] and via Robert G. Gallager [26] and then advanced by Donald Knuth and Jeffery S. Vitter [20,19]. in this technique, a exclusive method called sibling assets is followed to construct a Huffman tree.

In this approach, both sender and receiver hold dynamically changing Huffman code timber whose leaves constitute characters visible to this point. Initially the tree includes simplest the 0-node, a special node representing messages that have not begun to be seen. Right here, the Huffman tree includes a counter for every image and the counter is up to date each time whilst a corresponding input image is coded. Huffman tree below creation continues to be a Huffman tree if it's miles ensured by means of checking whether or not the sibling assets is retained. If the sibling assets is violated, the tree must be restructured to make certain this assets. Generally this set of rules generates codes which might be extra powerful than static Huffman coding. The adaptive Huffman coding is advanced to Static Huffman coding in elements: firstly It requires best one bypass thru the input and it provides little or no overhead to the output. However this set of rules has to rebuild the whole Huffman tree after encoding each symbol which becomes slower than the static Huffman coding.

LZ77 set of rules

Jacob Ziv and Abraham Lempel have supplied their dictionary-primarily based set of rules in 1977 for lossless records compression [27]. LZ77 exploits the truth that phrases and phrases inside a textual content file are in all likelihood to be repeated. When there is repetition, they may be encoded as a pointer to an earlier occurrence, with the pointer accompanied by way of the number of characters to be matched.

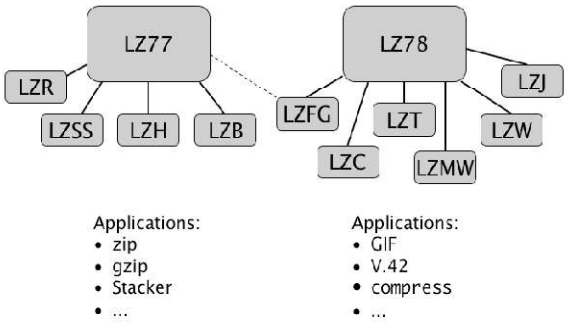

Fig.2. LZ Family

LZ77 is a totally easy adaptive scheme that calls for no previous information of the source data and appears to require no assumptions approximately the characteristics of the supply. In this method, the dictionary is truely a portion of the formerly encoded sequence. The encoder examines the enter collection thru a sliding window which consists of two parts: a search buffer that consists of a part of the lately encoded series and a glance beforehand buffer that contains the following portion of the collection to be encoded. The set of rules searches the sliding window for the longest fit with the start of the appearance-in advance buffer and outputs a reference (a pointer) to that suit.

LZW-BPCS STEGANOGRAPHY

Within the preceding if the records capacity of the image is improved, the photo first-class deteriorates. The hiding ability of those strategies may be very low nearly 15% simplest. Our Steganography gives a good deal greater capacity near 50% compared to those strategies however when statistics capacity will increase it's going to have an effect on the photo satisfactory. So, we can't embed sufficiently massive statistics into the duvet photo. In our proposed approach we conquer this problem via preprocessing of information. On this approach, mystery statistics is first compressed after which the compressed mystery facts is embedded into the bits of the quilt picture depending upon the complexity picture. For compression of information a lossless records compression method LZW is used. Steganography method used is bit aircraft complexity segmentation technique. The name of the game message is embedded inside the noise like regions. So, after LZW compression compressed mystery information is embedded into the noise of the cover image. As compression reduces the size of mystery message, this can increases the hiding capability without any effect on their excellent.

Conclusion and future Work

On this segment first off, we purpose to provide an explanation for the blessings and disadvantages of the proposed approach. An advantage of the proposed method is not being language unique. The technique may be implemented to any language by using reconstituting the textual content database and adapting the Latin rectangular to the regarding language, if essential (for e.g. Chinese and Arabic languages). Any other advantage of the proposed technique is shielding the originality of the cover media at the same time as speaking. The technique does now not produce noise a good way to hide mystery records. It modifications neither which means nor layout of the cover text. in the proposed method, the stego cowl is a ahead mail platform that incorporates two cover medium. One in every of them is the clearly generated cowl text. So the textual content is meaningful, syntactically and grammatically correct and valid. Another is the selected e mail addresses so as to reveal the mail as a forward mail platform. There is not any format or constraint on producing e-mail addresses (numbers, repeating characters can be used) and it isn't vital for them to be significant. in order that they do no longer increase suspicion. Because of those specifications, the proposed technique is robust towards OCR applications and retyping. Moreover, security of the proposed method has been supported with the aid of the employed stego keys. Besides, Combinatory-primarily based coding and LZW compression have also been hired for this motive. As future work, we goal to analyze the outcomes of different lossless data compression algorithms like Huffman Coding and mathematics Coding, first of all on capability. For a more enormous ability increment, we purpose to apply shorter naturally generated texts in textual content base. Ultimately via increasing the sort of text base with those shorter texts, we goal to attain the preferred randomness in case of hiding similar patterns.

REFERENCES

Bashardoost, Morteza, Ghazali Bin Sulong, and Parisa Gerami, 2013. "Enhanced LSB image Steganography method by using knight Tour algorithm, Vigenere Encryption and LZW compression." *IJCSI International Journal of Computer Science Issues,* 10, no. 2, 221-227.

Chetan, Er, and Er Deepak Sharma. "Fractal Image Compression Using Quad Tree Decomposition & DWT." *International Journal of Advanced Research in Computer Scienceand Software Engineering,* 5, no. 6 (2015).

Garg, Nancy, and Kamalinder Kaur, 2016. "Hybrid information security model for cloud storage systems using hybrid data security scheme".

Jafari, Reza, Djemel Ziou, and Mohammad Mehdi Rashidi, 2013. "Increasing image compression rate using steganography." *Expert Systems with Applications,* 40, no. 17, 6918-6927.

Kumar, Rajeev, Aruna Malik, Samayveer Singh, and Satish Chand, 2016. "A high capacity email based text steganography scheme using Huffman compression." In *Signal Processing and Integrated Networks (SPIN), 2016 3rd International Conference on*, pp. 53-56. IEEE.

Lin, Chi-Yuan, Shu-Cing Wu, and Jyun-Jie Wang, 2014. "VQ Image Compression Steganography Based on Section-Based Informed Embedding." In *Computer, Consumer and Control (IS3C), 2014 International Symposium on*, pp. 111-114. IEEE.

Malik, Aruna, Geeta Sikka, and Harsh K. Verma, 2017. "A high capacity text steganography scheme based on LZW compression and color coding."*Engineering Science and Technology, an International Journal* 20, no. 1, 72-79.

Satir, Esra, and Hakan Isik, 2014. "A Huffman compression based text steganography method." *Multimedia tools and applications,* 70, no. 3, 2085-2110.

Zhang, Yi, Xiangyang Luo, Chunfang Yang, Dengpan Ye, and Fenlin Liu, 2015. "A jpeg-compression resistant adaptive steganography based on relative relationship between dct coefficients." In *Availability, Reliability and Security (ARES), 2015 10th International Conference on*, pp. 461-466. IEEE.

****