LSB Embedding with Hillbert Curve

Ravi Kumar Yadav

Assistant Professor, Department of Computer Science, Keshav Mahavidyalaya, University of Delhi

Abstract

Information security has pulled in open consideration within the past decade, effective strategies for information security assurance are in request. This method inserts the secret information into the pixels and the pixel order for embedding is by Hillbert Curve. With standard petitcola images many experiments are carried out. The prevalence of ciphertext and encrypted images in cyberspace has led to increased interest from academics towards RDH algorithms in encrypted images.

Keywords: Steganography, Hilbert Curve

1. Introduction

The technology and speed of communication are having latest developments and information sharing via public networks. Information systems are now being regularly attacked by the cyber criminals. Therefore, Information security importance has increased tenfold. Military organisations, e-commerce, banking, research and intelligence agencies are using the applications of Information security. There was as well as there is need of Hidden communication for information processing and communication systems. The hidden communication concept is in use through various forms for more than 2000 years.[1][2]

2. History of Data Hiding & Steganography

In Greece, the greek historian Herodotus recorded stories of information hiding. King Darius of susa shaved slave's head and tattoo a secret message. The slaves's hair is allowed to grow before he can be sent to recipient carrying secret message in the form of tattoo on head. The recipient shaved the slave's head and read the secret message on hisarrival.[3][4] Soldier Demeratus to send a message to sparta that "Xerses intended to invade greece"[5][6]

The text was written on wax-covered tablets. The wax is removed from the tablet and secret message is written, waxed the tablet to make it looks as blank tablet and sent. Invisible ink was u seed by romans for information hiding. Fruit juices milk was used as invisible ink. The scret message was readable after heating the surface. The invisible ink is still in limited use for Hidden communication. The Steganographia(Johannes Trithemius1462-1526), Steganographia (Gaspari Schotti 1665), Cryptographic Militaire, Les filigrannes are the text attributed foundation of information hiding.

These text and memoirs indicate, hidden communications concepts are not new idea. Computers and internet are the digital medium used to communicate covertly. Information hiding field became now significantly advanced. The computer systems, communication networks, multimedia objects, programs are the technology used as host cover in information hiding. The two areas of Information Hiding are Digital Watermarking and Steganography Most of the information are available in public domain, more protection from tempering and impersonation of such information is required. Modern Steganography is an art of data hiding in a digital cover so it is not detectable by eavesdropper. The primary goal of steganography is that perceptually and statistically information is undetectable. Digital Watermarking, is to protect digital information from tamper, copyright violation, illegal copying and unauthorised modifications. The carriers of data hiding are image, audio, video, text, tcp/ip header, all these contains high degree of redundancy. The most popular carrier for digital steganography is image. Modern image steganography steps include: selection of cover image, encrypt and/or compress the original message, sender embed secret message in cover image and use a key known to sender/receiver, recipient decode the secret message using the key. The attributes of Digital steganography are Invisibility or imperceptibility, Payload or capacity undetectability [7]

Intuitive Compression Techniques

Data compression is the conversion of input stream data into output stream data that with smaller size. The long inefficient representation of data into short meaningful representation of data. Any non random collection of data will always have some data redundancy which can be exploited to have a smaller representation which means removing the redundancy from the data. In computers data is represented in formats that is longer than necessary. For additional security compression techniques, error correcting codes and cryptography methods are used with steganography. Some of the Compression methods are :

Braille: Well known code that enable blind to read developed by louis Braille. Braille code 3x2 dots cells of code written on thick paper. Each dot may be flat or raised, so the information is 6 bits(64 possible groups). Irreversible Text Compression: sometimes the compression of text is done by throwing away a part of the information called irreversible text compression. For ex. more number of spaces are replaced with single space. Symbols are encoded in 5bits instead of 8 bits. This is a type of special-purpose method. Ad-Hoc Text Compression: Packing where ascii code of 8 bits per character is replaced with the 7bits per character. This makes compression ratio : 7/8 = 0.875. If the text contains only uppercase letters, digits and punctuation marks then 6 bit CDC display code may be used.

Baudot code: 5 bit code by JME Baudot in 1880 for telegraph communication. This is unreliable code with no parity bit. Run Length Encoding: The basic idea of run length encoding is that if a data item A occurs N consecutive times in input stream then n occurrences are replaced with single pair AN. N times consecutive occurrence of a data item is run length of N. This coding is run length encoding.

RLE text Compression: When three or more consecutive bytes are found in an input stream then data compressor writes the byte followed by repetition count. Digram encoding and pattern substitution are the variants of RLE. RLE Image Compression: Digital Image : Image with small dots called pixels, where each pixel is one bit Black or white, or several bits to indicate several colors or shades for gray. Array that store pixels in memory are Bitmap. RLE based image compression is the idea that if a random pixel is selected then there is always a good chance that its neighbors also have same colors. If the bitmap starts with 16 white then two black then 17 white pixels this is encodes as 16 2 17

The size of output compressed stream depends on the image type. For grayscale images, the grayscale value and the repetition value are stored in the compressed stream. RLE image compressor scans the image bitmap by rows, by columns or zigzag all three ways to have best compression results. Bitmap rows are encoded by rows because normally adjacent pixels are identical. The disadvantage of RLE is that the runlength is completely redone for the modified image.

Steganography

Image Steganographic schemes Signals are representing information. Signals are fluctuating quantities that are conveying information. TV and Radio signals are an example of signals. Signals convey some types of information to recipients. The analysis of signals and extracting meaningful or useful information with signals is signal processing. The entity or system that completes the task of processing signals called a signal processing system. For example, a Radio set. Signals processing systems are natural processing systems for example yes and man-made processing systems for example TV or Radio. Man made processing systems are Analog processing systems which process continuous signals and digital signal processing systems which process discrete signals.

Image is also an example of signal. Image processing system is a system which process images. Image processing systems are the natural systems like eye and brain pair and manmade that are analog image processing systems like film camera, motion picture, film projector. Image which has two pixel element as 0 for black and 1 for white is a type of Binary image. Grayscale image is a 8bit color image format in which 0 is for black and 255 for white. This type of image can have 127 strands of gray. Color image format having 16 bit information(65536 different colors) for a pixel element is also called High Color Format.

Digital image processing systems are the discrete signal processing systems that captures and process images in digital format. The most important applications where image processing systems needed are astronomy, medical image processing, remote sensing, machine and robot visions, spectroscopy.

Lsb steganography is simple and most popular scheme with more imperceptibility but poor embedding capacity and very low robustness.[11] Image steganography are possible with two domains i.e. spatial domain and frequency domain. Pixels position in image and pixel intensity values are the basis of a steganography technique.

LSB Matching : The pioneering technique in steganography is LSB(Least Significant bit) substitution technique. LSB value in cover binary sequence is replaced with secret data bit. One byte of secret information is divided into eight bits and each bit is embedded in LSB byte of cover binary sequence. This process may be repeated for all the bytes of secret information such that all secret information is hidden into cover binary sequence. And the cover binary sequence can be any type of files.

Image quality metrics : Acquisition and processing of image scan degrade the quality of images and introduce distortion. Quality metrices correlates with subjective perception for the quality of images by human observer. Original image and the image in which secret information is embedded are compared and checked for similarity. If the original image not available for for reference then no-reference image quality metrices can be used. To measure the quality of images with embedding PSNR or WPSNR are used. The structural similarity of images are measured with SSIM.[8] MSE-Mean Square Error : The average squared difference of actual and ideal pixel intensities. This metric may not be align with human perception for the quality.

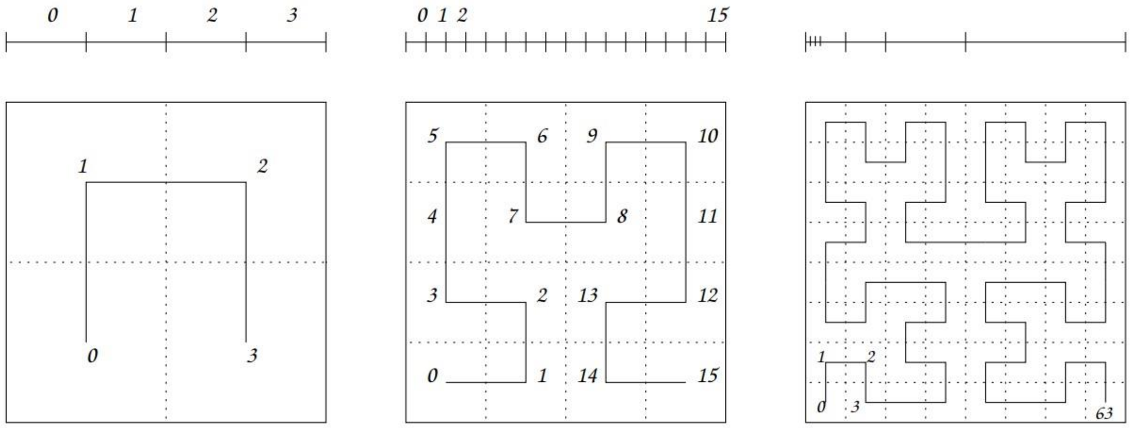
3. Embedding with Hillbert Curve

David Hillbert described space filling curve called Hillbert Curve which is a variant of a continuous sfc(space filling curve). Hillbert curve preserves the locality of reference and provide a mapping from 1D to 2D. Generalisation of gray codes are an instance of Hillbertcurve in Higher dimension. Geometry modelling includes surface and volume oriented models. 2D surface are having resolution and that cane be square based grid which can be recursive to obtain a tree, with each node having four children cells. Quadtree are used for recursively structured grids. Quadtree based grids can be generated and extended to 3D also.

Quadtree cells may have a sequential order. And the data stored using quadtree can be processed in a specific way. The traversal of quadtree cells i.e. all cells are visited at least once. DFS traversal generates a order of cells that leads to pretty large jumps, in sequential order.

It is possible to generate an order where predecessors and successors are direct neighbours of a cell. Large jumps are avoided in this hillbert curve order. Two neighbouring cells according to sequential order will also be neighbours in geometry. The property of sequential order generates a unique index for each data item. This property ensures that the data items are stored and retrieved with this index and items are processed exactly once, no data item skipped and processed twice. Mathematically this mapping indices and data items should be bijective. If two data items are close then their index should be close together. These cells can be pixels for images. For images 2D array of pixels, the order of accessing the pixels is with Hillbert Order. 2D array of pixels are now in 1D array of pixels using this.

Figure 1.



Sometimes data warehouses are compressed using Hillbert curve.[11]

LSB embedded message is "geeksfor" in every 8x8 pixel block for an image. A color image is first converted into grayscale image and lsb bits are changed as per the secret message. For a 8x8block provides 64 lsb bits and the secret message “geeksfor” also converted into 64 bits. Each 64 bits of the block are first converted into hillbert order and lsb bits are inserted. After inserting message bits in all lsb of the grayscale image then the original and the result image are compared to see the similarities.

The histogram is a crucial aspect of images and has been effectively applied to various tasks such as hashing, image retrieval, and image copy detection. A homogeneous histogram in a picture is anticipated to be produced by the optimal protection strategy. SSIM is having the range of -1 to 1, score of 1 means similar image and score of -1 means completely different images.The secret message from the lsb bits of a grayscale image can be extracted by using 8x8 block and following the lsb bits in hillbert curve order.

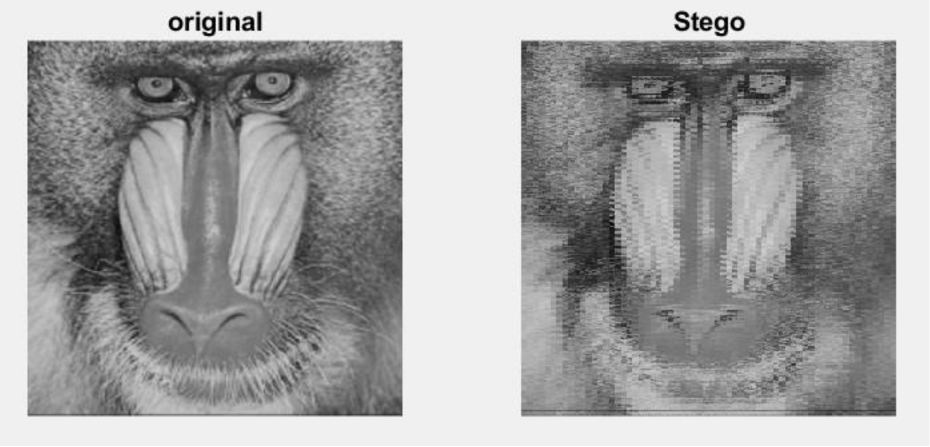
4. Image Quality Metric Results

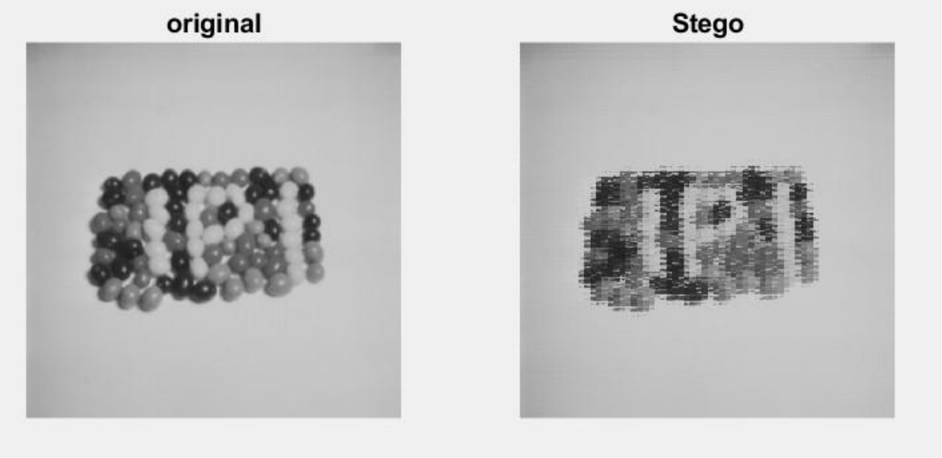
Table 1: Image Quality Metric data

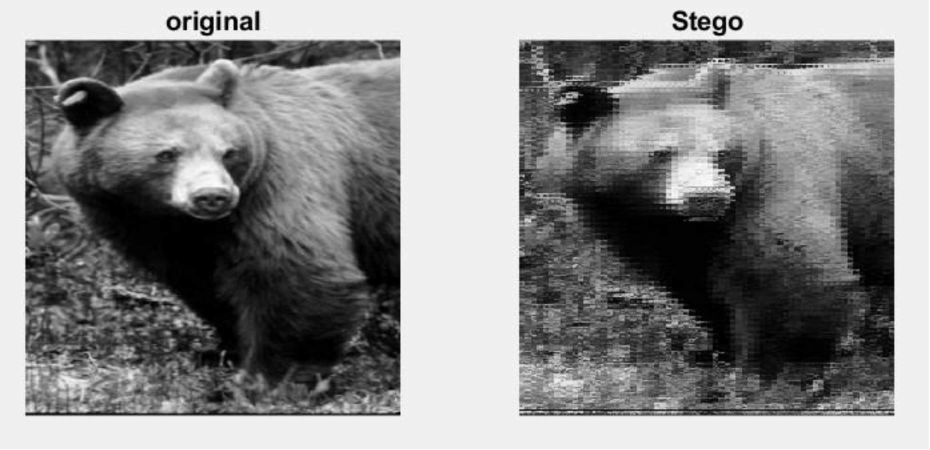
	SSIM	**PSNR**	**SNR**
Baboon.jpg	0.3596	21.8310	16.2776
Beans.jpg	0.8394	25.3548	22.35.17
Bear.jpg	0.3712	19.4984	11.9824
F14.jpg	0.8949	28.9457	22.7253
Ca2.jpg	0.3922	18.4708	13.3187
F16.jpg	0.6061	20.8478	18.0430

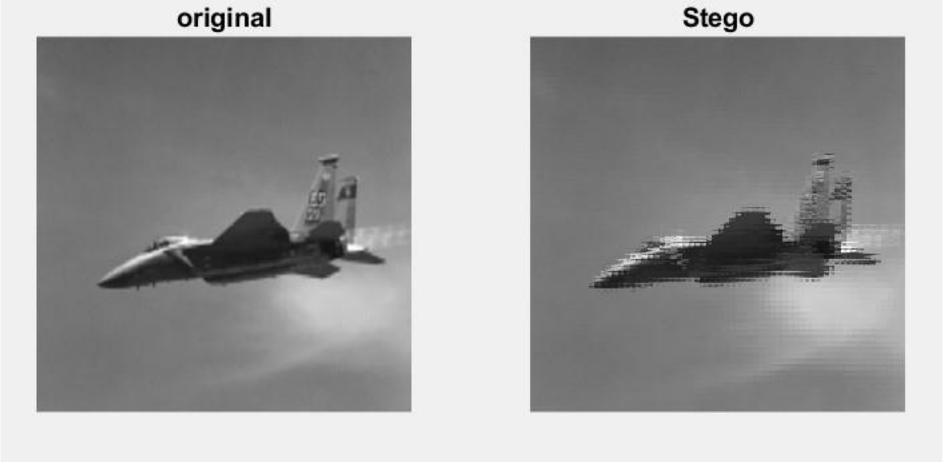
	SSIM	PSNR	SNR
Baboon.jpg	0.3596	21.8310	16.2776
Kid.jpg	0.6511	22.5994	15.0659
Lena.jpg	0.6032	22.5161	16.8345
Opera.jpg	0.5867	23.5679	17.4677
PaperMachine.jpg	0.4251	19.1059	11.9210
Peppers.jpg	0.5809	21.7303	15.9635
Total	901	839	1631

The following are few results of embedding in the petitcolas images:

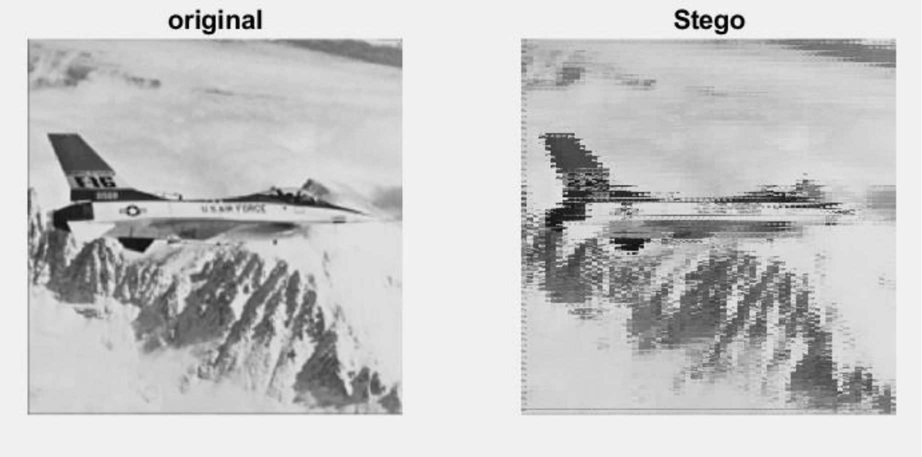




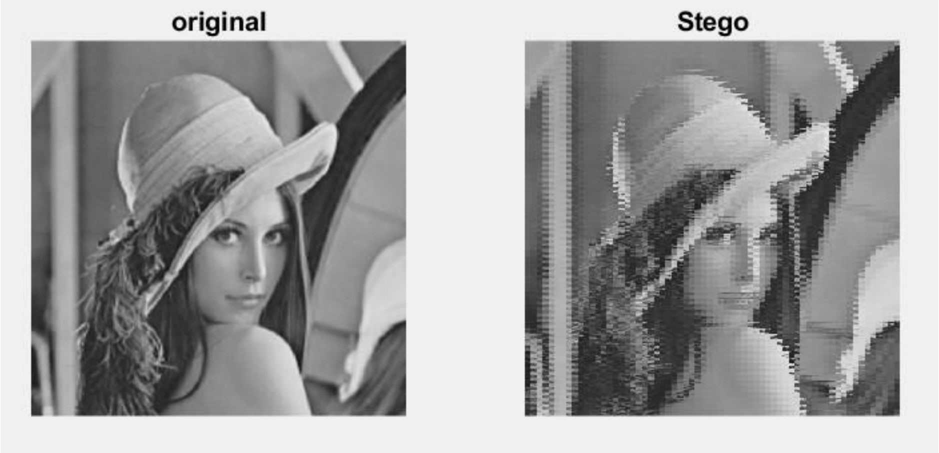


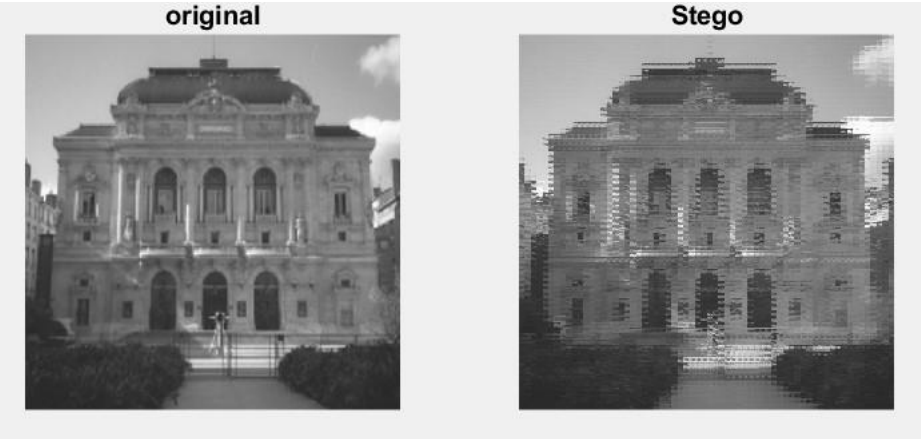

original
Stego

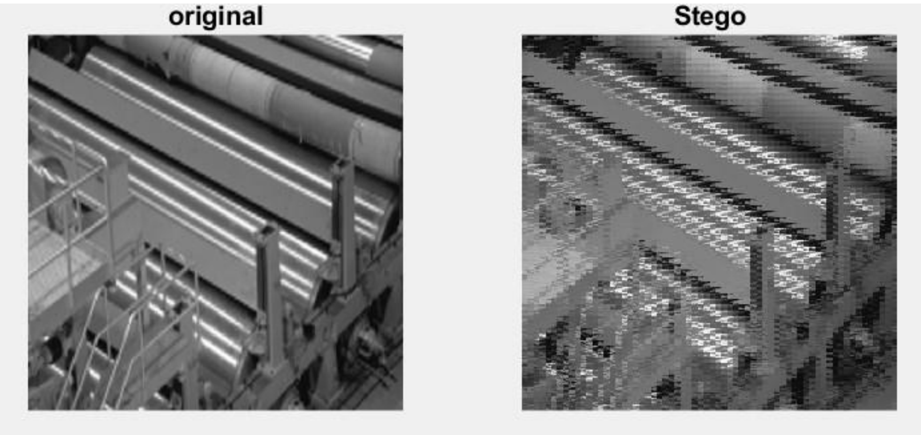
original
Stego

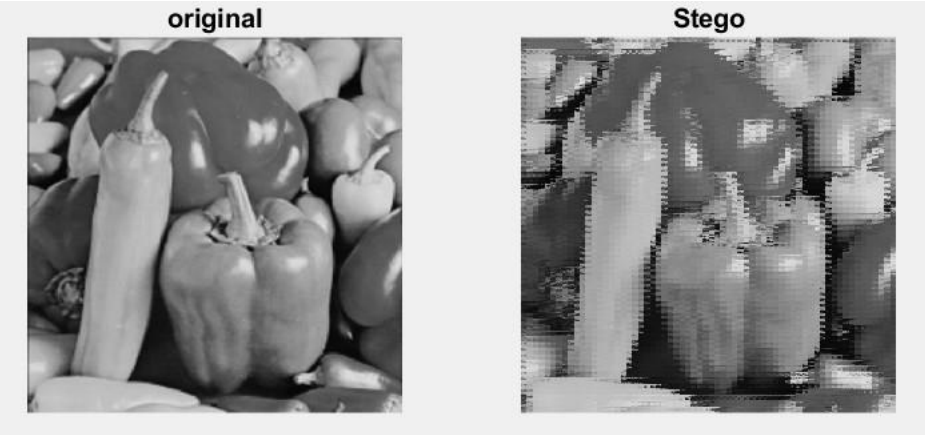

original
Stego

original
Stego


original
Stego


original
Stego


original
Stego


original
Stego

5. Conclusions

In this article, a simple lsb steganography with the order of pixels for embedding is by Hillbert curve. All the pixel's lsb are used for embedding. The SSIM quantity along with the PSNR is indicating good results. The decrease in the secret information will improve SSIM and PSNR. The results looks promising that if the selected information is embedded as watermarking or steganography then the value of PSNR will be better. This work can be extended by using various techniques of steganography and cryptography. I hope to use difference expansion method with hillbert order curve for extending this work.

6. Acknowledgement

Heartfelt thanks to Dr. S.K. Muttoo for providing me motivation, support and guidance at all times by clearing my doubts.

7. Authors' Biography

Ravi Kumar Yadav, teaching as Assistant Professor in the department of computer science, Keshav Mahavidyalaya, University of Delhi.

8. References

1. Sallee, P. (2005). Model-based methods for steganography and steganalysis. *International Journal of Image and graphics*, *5*(01), 167-189.
2. Chandramouli, R., Kharrazi, M., & Memon, N. (2004). Image steganography and steganalysis: Concepts and practice. In *Digital Watermarking: Second International Workshop, IWDW 2003, Seoul, Korea, October 20-22, 2003. Revised Papers 2* (pp. 35-49). Springer Berlin Heidelberg.
3. Cox, I. J., Kalker, T., Pakura, G., & Scheel, M. (2005). Information transmission and steganography. In *Digital Watermarking: 4th International Workshop, IWDW 2005, Siena, Italy, September 15-17, 2005. Proceedings 4* (pp. 15-29). Springer Berlin Heidelberg.
4. Cox, I., Miller, M., Bloom, J., Fridrich, J., & Kalker, T. (2007). *Digital watermarking and steganography*. Morgan kaufmann.
5. Eggers, J. J., Baeuml, R., & Girod, B. (2002, April). Communications approach to image steganography. In *Security and watermarking of Multimedia Contents IV* (Vol. 4675, pp. 26-37). SPIE.
6. SANS Security Essentials, (2001). Encryption and Exploits, Vol. 1.4,Chap 4.
7. Zaidoon, K.A.A., Zaidan, A. A., Zaidan B. B., Alanazi, H.O. (2010).Overview: Main Fundamentals for Steganography. Journal of computing, Vol. 2, Issue 3.
8. Zielinska E., Mazurcsyk W., Szczypiorski K. (2012, February). TheAdvent of Steganography in computing environments",http://arxiv.org/abs/1202.5289.
9. Mohamed, M. H., Mofaddel, M. A., El-Naser, A., \& Tarek, Y. (2023). Comparison Study Between Simple LSB and Optimal LSB Image Steganography. Sohag Journal of Sciences, 8(1), 29-33.
10. Tian, J. (2002, December). Reversible watermarking by difference expansion. In Proceedings of workshop on multimedia and security (Vol. 19). Juan-les-Pins: ACM.
11. Sagan, H. (2012). Space-filling curves. Springer Science, \& Business Media.
12. Nguyen BC, Yoon SM, Lee HK. MultiBit Plane Image Steganography. IWDW, LNCS, vol. 4283, pp. 61–70; 2006.

13. Mahimah, P., \& Kurinji, R. (2013,December). Zigzag pixel indicator based secret data hiding method. In 2013 IEEE International Conference on Computational Intelligence and Computing Research (pp. 1-5). IEEE.
14. Hilbert, D. (1891). ber stetige Abbildung einer Lmie auf ein Flchenstck. *Math Annalen*, *38*, 459-460.
15. Lawder, J. K., & King, P. J. (2000, July). Using space-filling curves for multi-dimensional indexing. In *British National Conference on Databases* (pp. 20-35). Berlin, Heidelberg: Springer Berlin Heidelberg.
16. http://opac.sub.uni-goettingen.de/DB=1/PPN?PPN=235181684