International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 10 (2015) pp. 26529-26542
© Research India Publications
http://www.ripublication.com

A Novel Reversible Data Hiding Scheme Based on Enhanced Difference Pair Mapping

Greeshma Balachandran
Mtech Student, ECE Department,
Amrita Vishwa Vidyapeetham University
greesh.ab@gmail.com
Dr. S. Ravishankar
Professor, ECE Department,
Amrita Vishwa Vidyapeetham University
s_ravishankar@blr.amrita.edu

Abstract

This paper focuses on the design of a novel Reversible Data Hiding Scheme using Enhanced Difference Pair Mapping which aims at higher embedding performance. Difference Pair Mapping is developed on the difference of pixel pair values in an image. Each pixel value is modeled in a context based approach using Gradient Edge Detection Predictor. The proposed approach uses run length coding to compress and decompress the location map of the specified image which indicates location for data embedding. The lossless address generation and address recovery scheme is accomplished using run length coding. The experimental evaluation of the proposed method is simulated in Matlab 2013a. The results indicate improvement in embedding capacity with better image quality as compared to current existing techniques.

Keywords: Difference Pair Mapping(DPM); Reversible Data Hiding (RDH);Run Length Coding(RL); Gradient Edge Detection Predictor(GED).

Introduction

Information exchange in today's communication has taken a paradigm shift for data transmission. Providing security to the information transfer has also become an important issue in the design of communication system. Cryptography and Steganography are the proven methods to provide security to the information. Cryptography is used to encrypt the message to combat information hacking. Steganography or data hiding is used to hide the information in a cover media for information transmission. The cover can be a text, audio or video file. Data hiding is the ability by which hidden data can be made accessible to a specific group of people while keeping the medium confidential and secure. In general, there exists a trade-off relationship between embedding capacity, quality of cover image and faithful reproduction of retrieved data. Reversible Data Hiding (RDH) is a technique developed to embed a secret data into a cover image and extract the same from masked image. The major goal of RDH is the lossless recovery of data and image[1,2-3].

The paper is organized as follows. Section II describes the current research in the area of RDH. Section III describes in detail the proposed RDH Scheme. Section III A describes the Location Map (LM) generation and recovery for embedding and extraction of secret data. Section III B describes embedding and extraction process. Section III C gives detailed description of Gradient Edge Detection Predictor (GED) for context based image modeling. Section III D and III E explain about the modified approach to embedding and extraction of data bits. Section IV accounts for the result analysis and compares it to the previous related works. Section V concludes the proposed work by highlighting the future scope.

Related Works

The first RDH method based on histogram was proposed by Ni et al. which uses peak and valley points of the pixel-intensity-histogram to embed data bits[4]. It does not hold good for flat histogram and the embedding payload is low. Lee et al. proposed difference-histogram based RDH which can exploit the neighbouring pixel correlation better and can embed larger amount of payload with lesser amount of distortion[5]. Tai et al. proposed histogram modification based reversible data hiding which increases the Embedding Capacity (EC) but the desired optimality in the visual appearance of image is lost[6]. Tai et al.'s method is improvised by Hong, which utilises a Dual Binary Tree (DBT) which can increase EC with distortion[7]. Enhanced Difference Pair Mapping (DPM) based RDH scheme is motivated from the aforementioned works.

In the proposed RDH Scheme image redundancy is exploited for embedding larger amount of data .The limitation in this approach is lower EC since only one pixel in a pixel pair is allowed to be modified in value. The proposed scheme employs DPM based two dimensional RDH with Run Length Coding (RL) for compression and decompression of Location Map (LM) and Gradient Edge Detection Predictor (GED) for context modeling.

Proposed RDH Scheme

In the proposed methodology, the RL scheme has been exploited for compression and decompression of address of the LM. GED is a single threshold controlled predictor which utilizes both the efficiency of Gradient Adjusted Predictor (GAP) and simplicity of Median Edge Detector Predictor (MED)for context modeling[8]. This has resulted in better EC without compromising on the image quality. The block diagram in Fig 1 and Fig 2 describes the embedding Process and extraction Process. The embedding and extraction process is described in Section A,B,C,D and E.

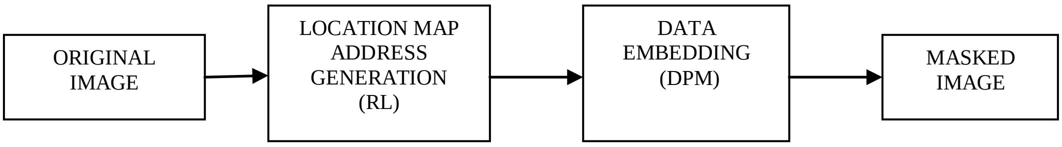


Figure 1: Enhanced RDH Scheme for Embedding

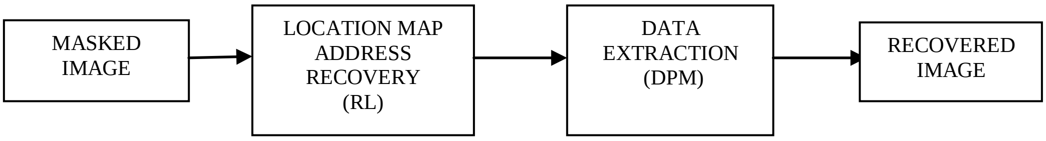


Figure 2: Enhanced RDH Scheme for Extraction

A. Location Map Address Generation and Address Recovery

The lossless address generation and address recovery is accomplished using RL which specifies the location for data embedding and extraction. Run length Encoding (RLE) is a data compression technique that can encode large runs of repeating bits by only transmitting one bit from the run and displays how many times this repetition is accomplished as shown in Fig 3.The proposed work uses Run length scheme since it can losslesssly compress the data sequence and extract the sequence without affecting the image quality[9,10]. This technique fails when used to compress language texts since they don’t have long runs of repeating elements.

The run length compression scheme accounts for a higher degree of Compression Ratio (CR) when compared to Arithmetic Coding (AC).The image quality in terms of Peak Signal to Noise Ratio (PSNR) accomplishes a higher value when run length compression and decompression is used.

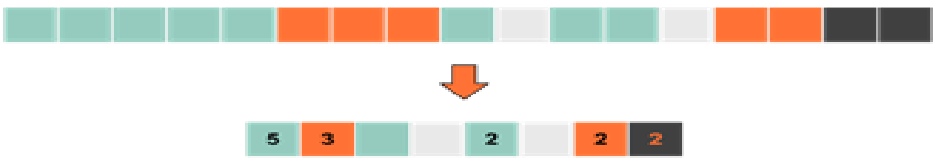


Figure 3: Run length Compression

B. Data Embedding and Data Extraction

The embedded and the extracted data bits are evaluated based on the pixel values of (x,y,z). In the proposed model, the two difference values d_1=x-y and d_2=y-z as depicted in Fig 4 are computed to determine a two dimensional difference histogram of (d_1,d_2). The computation of z is explained in Section III C.

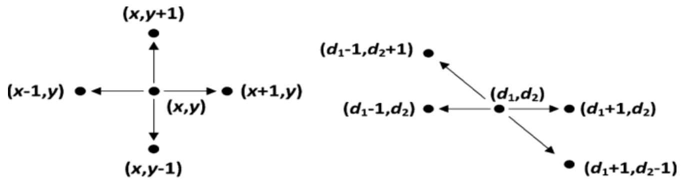


Figure 4: Determination of (d_1,d_2) from pixel pair values of (x,y,z)[11]

C. Gradient Edge Detection Predictor

The five neighbouring pixel values as depicted in Fig 5 are used for the estimation of local gradient for context modeling in GED[8]. The use of GED prediction technique has considerable advantage over GAP and MED predictors. GED predictor combines the efficiency of GAP and algorithmic simplicity of MED predictor.MED predictor cannot adapt to high noise regions whereas GAP can adapt to high noise variations with three thresholds.

Table 1: Different Cases of Difference Pair (d_1,d_2) in DPM with b={0,1} for Embedding[11]

Conditions on (d_1, d_2)	Operation in data embedding	Modification direction to the difference-pair	Modification direction to the pixel-pair	Marked value
$d_1 = 1$ and $d_2 > 0$	expansion embedding	right	right	$(x+b, y)$
$d_1 = -1$ and $d_2 < 0$	expansion embedding	left	left	$(x-b, y)$
$d_1 = 0$ and $d_2 \geq 0$	expansion embedding	upper-left	up	$(x, y+b)$
$d_1 < 0$ and $d_2 = 0$				
$d_1 = 0$ and $d_2 < 0$	expansion embedding	lower-right	down	$(x, y-b)$
$d_1 > 0$ and $d_2 = 0$				
$d_1 = 1$ and $d_2 = -1$				
$d_1 > 1$ and $d_2 > 0$	shifting	right	right	$(x+1, y)$
$d_1 < -1$ and $d_2 < 0$	shifting	left	left	$(x-1, y)$
$d_1 < 0$ and $d_2 > 0$	shifting	upper-left	up	$(x, y+1)$
$d_1 > 1$ and $d_2 < 0$	shifting	lower-right	down	$(x, y-1)$
$d_1 = 1$ and $d_2 < -1$				

Table 2: Different Cases Of Difference Pair (d_1^m,d_2^m) in DPM with b={0,1} for Extraction[11]

Conditions on (d_1^m, d_2^m)	Extracted data bit b	Recovered value
$d_1^m \in \{1, 2\}$ and $d_2^m > 0$	$d_1^m - 1$	$(x^m - b, y^m)$
$d_1^m \in \{-1, -2\}$ and $d_2^m < 0$	$-1 - d_1^m$	$(x^m + b, y^m)$
$(d_1^m = 0$ and $d_2^m \geq 0)$ or $(d_1^m = -1$ and $d_2^m \geq 1)$	$-d_1^m$	$(x^m, y^m - b)$
$(d_1^m < 0$ and $d_2^m = 0)$ or $(d_1^m < -1$ and $d_2^m = 1)$	d_2^m	
$(d_1^m = 0$ and $d_2^m < 0)$ or $(d_1^m = 1$ and $d_2^m < -1)$	d_1^m	$(x^m, y^m + b)$
$(d_1^m > 0$ and $d_2^m = 0)$ or $(d_1^m > 1$ and $d_2^m = -1)$	$-d_2^m$	
$(d_1^m = 1$ and $d_2^m = -1)$ or $(d_1^m = 2$ and $d_2^m = -2)$	$d_1^m - 1$	
$d_1^m > 2$ and $d_2^m > 0$	no embedded data bit	$(x^m - 1, y^m)$
$d_1^m < -2$ and $d_2^m < 0$	no embedded data bit	$(x^m + 1, y^m)$
$d_1^m < -1$ and $d_2^m > 1$	no embedded data bit	$(x^m, y^m - 1)$
$d_1^m > 2$ and $d_2^m < -1$	no embedded data bit	$(x^m, y^m + 1)$
$d_1^m = 2$ and $d_2^m < -2$		

The GED predictor takes into account a single threshold and can adapt to high noise variations when compared to GAP and MED predictor. {A,B,C,D,E} are the neighbouring pixels as depicted in Fig 5.g_h and g_v represents the horizontal and vertical gradients.The T is an indication of the threshold limit to define the context of an image. The computation of z is defined by Equation 1.

		E	
	C	B	
D	A		

Figure 5: Causal pixels for GED predictor

if g_V-g_h>T, z=A
else if g_V-g_h< -T, z=B
else z=3*(A+B)/8 +(C+D+E)/12
Where $g_h=|D-A|+|C-B|$
$g_{V}=|C-A|+|E-B|$ (1)

D. Embedding Procedure

The embedding procedure as given in Fig 6 is characterized into three steps.

Step 1: The cover image is partitioned into groups of pixels without overlapping , except for last two rows and columns.

Step 2: The secret data bit to be embedded into the cover media must satisfy the LM criteria. Location map LM(i) is a binary value (0,1) which depends upon i where i denotes the pixel intensity value. If a pixel takes values{0,255} ,then the LM(i) is denoted as 1.Similarly for all other values of pixel-pair, LM(i) is assigned 0.The double thresholding in DPM is eliminated as GED predictor evaluates a threshold limit for context modeling when predicted value is determined. Thus secret data is embedded based on only the LM(i) values.

a) LM(i)=1,underflow and overflow occurs and data cannot be embedded into the cover media if i takes values {0,255}.
b) LM(i)=0,data can be embedded into the cover media as given in Table I, if i takes values from{1-254}.k* represents the index of last data carrying pixel pair obtained from Table I.

Step 3: The Least Significant Bit (LSB)replacement is the next criteria in the process. Auxiliary information and compressed LM is embedded as given in Table 1 to enhance the decoding process. The LM is compressed using RLE.

The GED predictor eliminates the 12 bits allocated as the pixel pair threshold.

a) The LSB of the first $3*\lceil \log_2 k \rceil$ +L image pixels are recorded as S where L is the length of compressed LM and k is a constant taking values of 262144 bits for a 512x512 gray scale image.

b) The binary sequence S is embedded as given in Table I, to the remaining pixel-pairs ranging from k*+1 to k_{end} where k_{end} is an indication of last LSB replaced pixel pair.
c) The LSB replacement of the first 3*|¯$\log_2 k$¯| +L image pixels with auxiliary information and compressed location map is accomplished. The auxiliary information of k*, L and k_{end} consists of 3*|¯$\log_2 k$¯|bits. Thus, masked image is determined.

The embedding procedure as given in Fig 6 is explained with an example:

When pixel pair(x,y)=(132,132) and z=131,the difference pair evaluates to (d_1,d_2)=(x-y,y-z)=(0,1).The pixel pair(x,y) is expanded to carry one bit of data. If the LM(i) consists of a value 0,then:

a) if the embedded data bit is b=0,then pixel pair remains as (x,y).
b) if the embedded data bit is b=1,then difference pixel pair evaluates to be(-1,2) and pixel pair is described as (x,y+1)=(132,133).

E. Extraction Procedure

The data extraction and image restoration of recovered media from the masked media as given in Fig 7 is characterized into six steps.

Step 1: The LSB of first 3*|¯$\log_2 k$¯| pixels of masked media is determined to obtain the indexes of k* and k_{end}.

Step 2: The next L LSB determines the length of compressed LM. The LM(i) is determined by decompressing L. The RL decompression technique generates the decompressed location map.

Step 3: The LSB sequence is extracted using Table II and the cover media is partitioned into groups of pixels without overlapping , except for last two rows and columns.

Step 4: The LM when accounts for a value 0, and pixel-pair restoration as described in Table II is accomplished. If location map is 1, there is no data extraction and pixel-pair remains intact.

Step 5: The embedded message starting from the k* pixel-pairs is extracted using Table II.

Step 6: The message bit is extracted and the original image is recovered from the masked media.

The extraction process as given in Fig 7 is described below:

When pixel pair of masked image is (x,y)=(133,132) and Z=131,the difference pair evaluates to $(d_1,d_{2)}$=(x-y,y-z)=(1,1).

a) The recovered image pixel value evaluates to (132,132) and extracted bit is 0.

The Difference Pair Mapping Technique embeds data into the original cover media as well as extracts the data bit on recovery. The image quality is guaranteed throughout the process described.

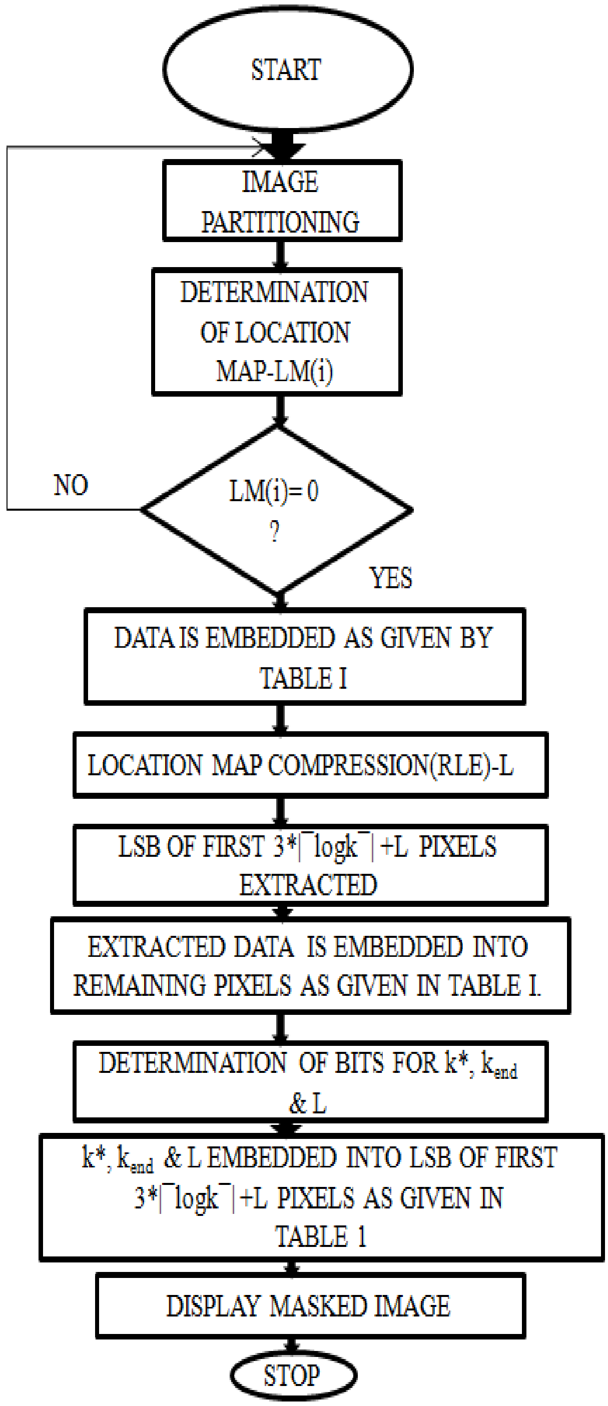


Figure 6: Flow Chart for Embedding Procedure

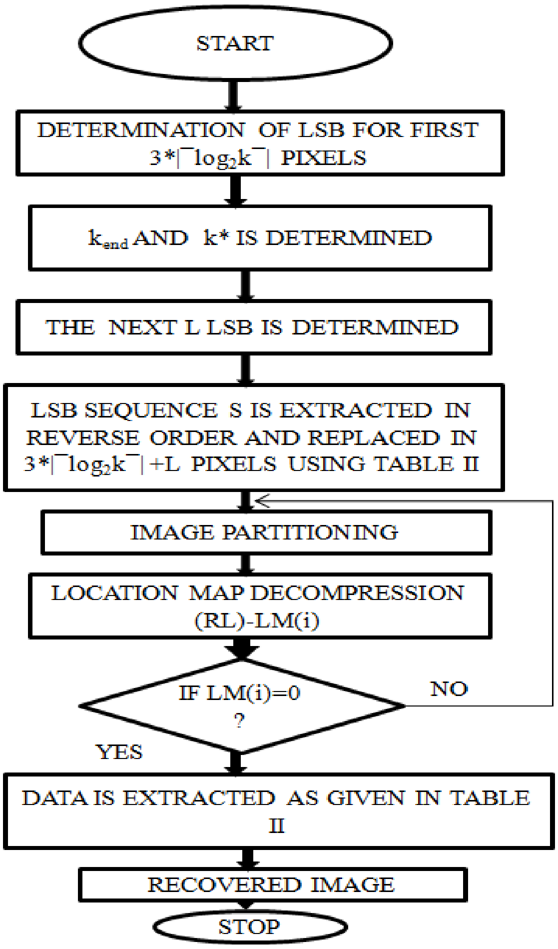


Figure 7: Flow Chart for Extraction Procedure

Simulation Results and Analysis

The six gray scale images of size 512x512 are used as a cover image in the proposed RDH scheme. The code for implementation of RDH algorithm is performed in Matlab 2013a. The simulated results of the proposed approach are compared with the Li et al. approach [11]. Li et al. uses arithmetic coding for location map compression and decompression and employs gradient adjusted predictor for image context modeling. The results of the proposed approach clearly indicate overall performance of reversible data hiding scheme. The simulation results are tabulated in Table III and IV. The experiment is repeated for an embedding capacity of 4000, 6000, 8000 and 10000 bits. The simulated results of masked image for two approaches Li et al. and modified Li et al. with RL is very much similar with marginal difference in PSNR value for all the six images as shown in Fig 10.With the proposed approach using RL coding for data compression and decompression and GED predictor for modeling the context of image, results show significant improvement in PSNR (dB) for same input image and different embedding capacity.

The Compression Ratio (CR) is determined for RL and AC compression scheme. The original location map is determined for the six images as given in Fig 7 which comprises of 262144 bits. The Compressed length of location map is determined as 36 bits in RL and 39 bits in arithmetic coding for all the six images given in Fig 7.Thus the compressed number of bits is increased by a factor of 3 in RL when compared to AC. The Compression Ratio is 65536/9 for RL and 262144/39 for AC. The RL compression scheme accounts for a higher degree of CR when compared to AC.

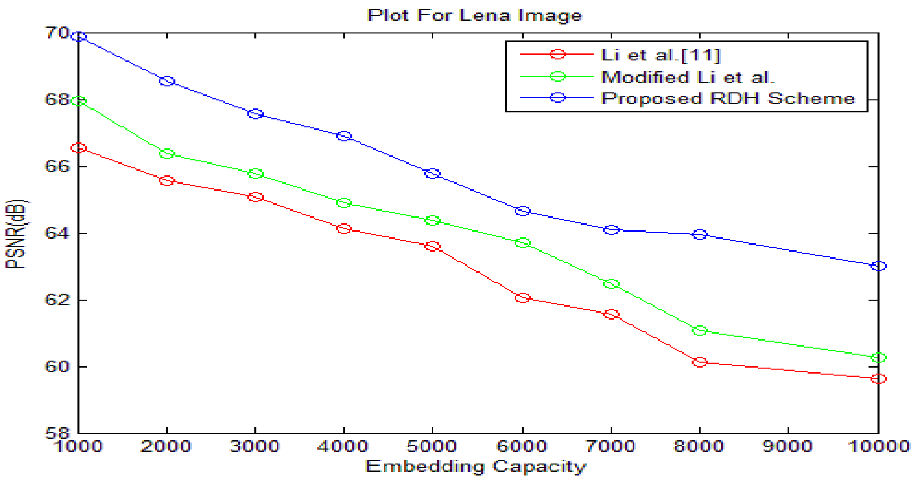
Figure 7: Original Images

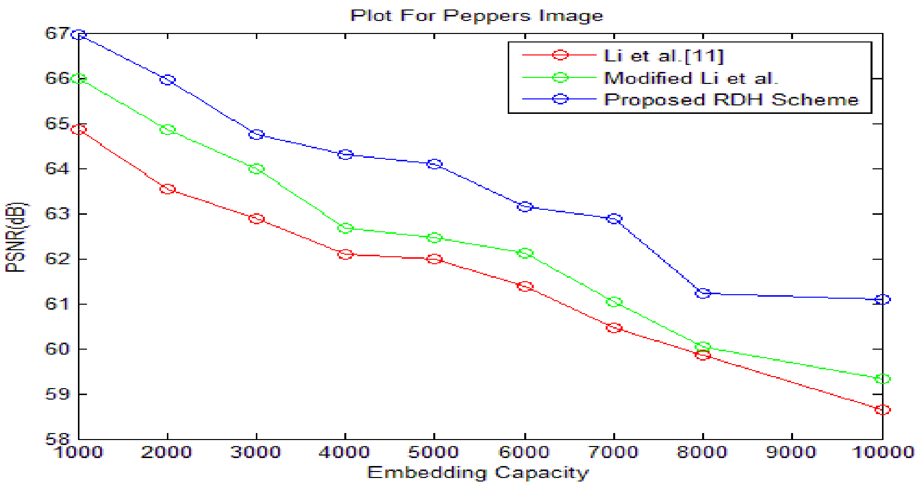
Figure 8: Masked Images

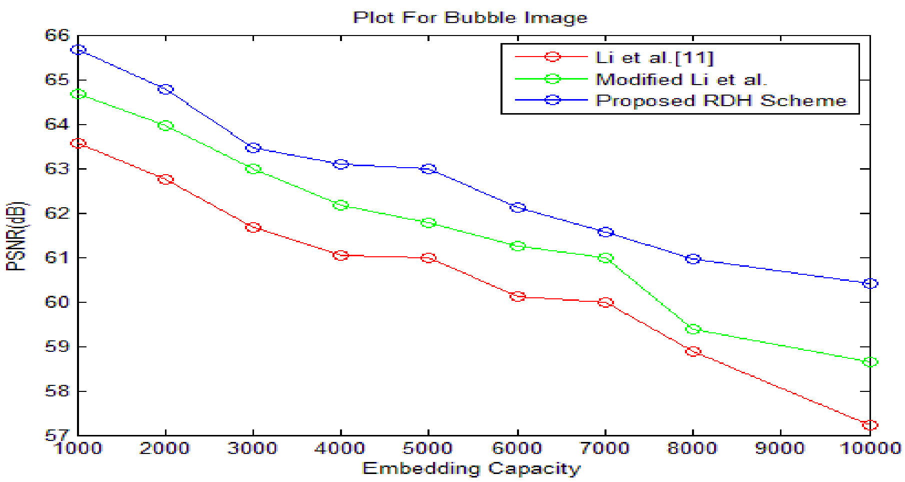
Figure 9: Recovered Images

Table 3: PSNR of Masked Image at Different EC

Image	Psnr(Db) Ec=4000			Psnr(Db) Ec=6000			Psnr(Db) Ec=8000			Psnr(Db) Ec=10000		
	Li Et Al. [13]	Rl (Gap)	Rl (Ged) Propo-sed	Li Et Al.[13]	Rl (Gap)	Rl (Ged) Propo-sed	Li Et Al. [13]	Rl (Gap)	Rl (Ged) Propo-sed	Li Et Al.[13]	Rl (Gap)	Rl (Ged) Propo-sed
Lena	64.14	64.90	66.89	62.04	63.70	64.65	60.14	61.08	63.95	59.62	60.27	63.01
Peppers	62.09	62.67	63.28	61.38	62.12	63.14	59.86	60.05	61.23	58.64	59.34	61.09
Bubble	61.03	62.18	63.10	60.13	61.25	62.13	58.89	59.38	60.96	57.23	58.65	60.41
Flower1	60.64	61.78	62.75	59.45	60.76	61.13	57.13	59.07	60.74	56.87	57.98	60.05
Field1	59.67	60.13	61.23	58.13	59.78	60.68	56.65	57.54	59.76	55.45	56.14	59.04
Field2	58.34	60.03	60.34	57.76	59.89	60.06	56.62	59.44	59.84	54.78	55.35	59.01
Average	60.985	61.948	62.934	59.815	60.253	61.965	58.215	59.426	61.08	57.098	58.288	60.43


Plot For Lena Image
Li et al.[11]
Modified Li et al.
Proposed RDH Scheme
PSNR(dB)
Embedding Capacity
70
68
66
64
62
60
58
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000


Plot For Peppers Image
Li et al.[11]
Modified Li et al.
Proposed RDH Scheme
PSNR(dB)
Embedding Capacity
67
66
65
64
63
62
61
60
59
58
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000


Plot For Bubble Image
Li et al.[11]
Modified Li et al.
Proposed RDH Scheme
PSNR(dB)
Embedding Capacity
66
65
64
63
62
61
60
59
58
57
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000

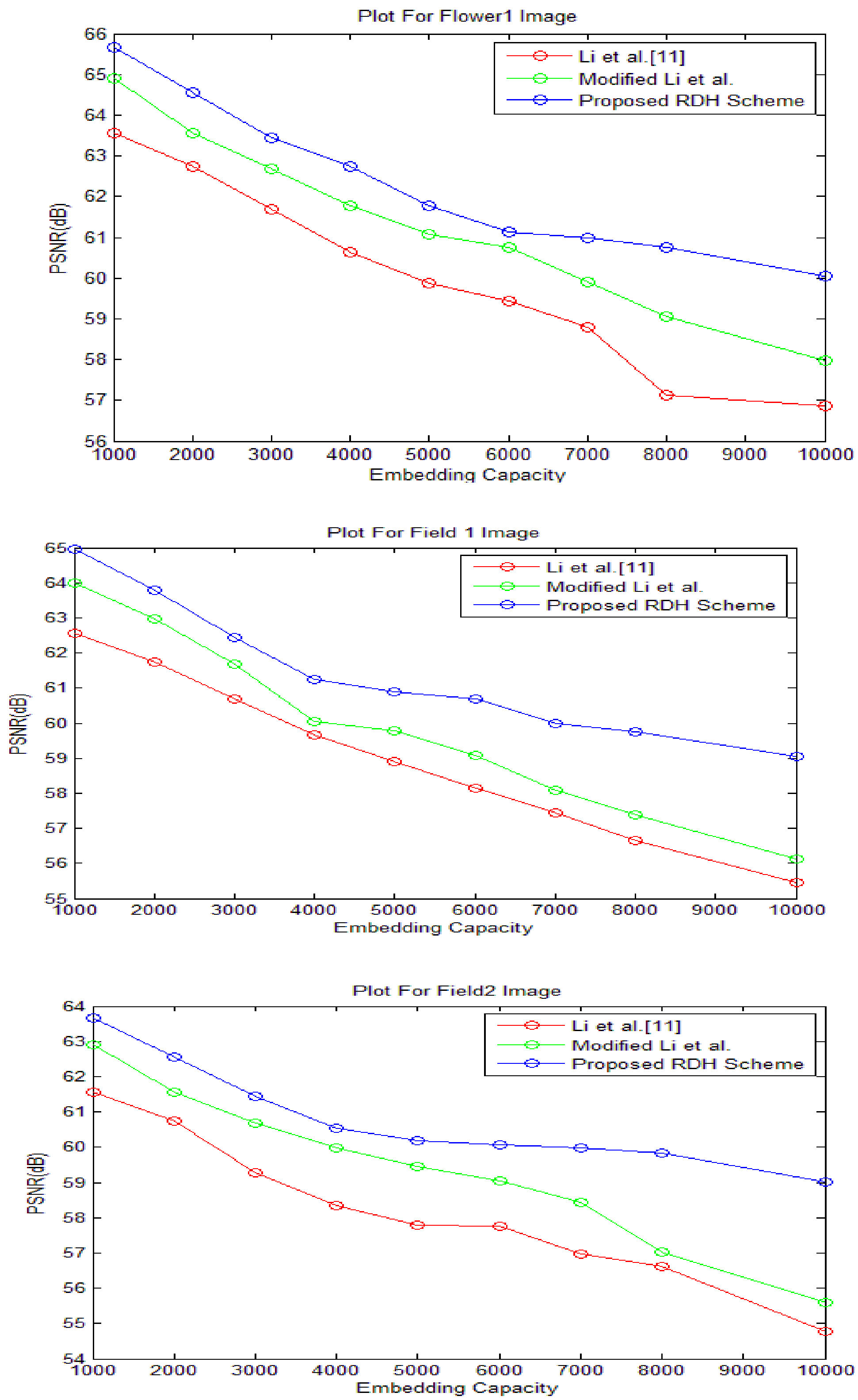


Figure 10: Performance Comparison of Images with Proposed RDH Scheme

Table 4: SSIM of Masked Image at Different EC

Image	Ssim Ec=4000			Ssim Ec=6000			Ssim Ec=8000			Ssim Ec=10000		
	Li Et Al. [13]	Rl(Gap)	Rl (Ged) Propo-sed	Li Et Al. [13]	Rl (Gap)	Rl (Ged) Propo-sed	Li Et Al. [13]	Rl (Gap)	Rl (Ged) Propo-sed	Li Et Al. [13]	Rl (Gap)	Rl (Ged) Propo-sed
Lena	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Peppers	1.0000	1.0000	1.0000	0.9999	1.0000	1.0000	0.9999	1.0000	1.0000	0.9998	0.9999	1.0000
Bubble	1.0000	1.0000	1.0000	0.9998	1.0000	1.0000	0.9997	1.0000	1.0000	0.9997	0.9998	1.0000
Flower	0.9997	1.0000	1.0000	0.9996	1.0000	1.0000	0.9996	1.0000	1.0000	0.9995	0.9997	1.0000
Field1	0.9996	0.9997	1.0000	0.9995	1.0000	1.0000	0.9994	0.9995	1.0000	0.9993	0.9995	1.0000
Field2	0.9996	0.9998	1.0000	0.9994	0.9996	1.0000	0.9995	0.9996	1.0000	0.9994	0.9996	1.0000
Average	0.9998	0.9999	1.0000	0.9997	0.9999	1.0000	0.9989	0.9998	1.0000	0.9996	0.9997	1.0000

As the embedding increases the PSNR decreases substantially which is depicted in Fig 10. The proposed RDH Scheme maintains a higher degree of PSNR(dB) on embedding of data into the cover media. The quality of the image is guaranteed throughout the process. The structural similarity of the masked media to the cover media is determined by a factor SSIM. The SSIM maintains high precision throughout with maximum accuracy as given in Table IV.

Conclusion

The proposed reversible data hiding scheme is more efficient than the earlier reported methods for any given embedding capacity. The proposed algorithm is computationally efficient in comparison to the algorithm proposed by Li et al.[11]. The results clearly show that the proposed algorithm for RDH can be easily implemented on a real time system employing GED predictor. GED predictor determines the predicted value using a single threshold which further eliminates the determination of pixel pair threshold for LSB replacement in embedding as well as extraction of data sequence as defined in Li et al.[11]. The proposed methodology maintains a higher degree of PSNR(dB) of 60.43 and SSIM of 1.0000 for an EC of 10000 bits of data sequence. Further, it is proposed to implement the RDH scheme on an FPGA resource.

References

[1] Y. Q. Shi, “Reversible data hiding ,” in *Proc. IWDW*, 2004, vol. 3304, pp. 1–12, ser. Springer LNCS.

[2] Y. Q. Shi, Z. Ni, D. Zou, C. Liang, and G. Xuan, "Lossless data hiding: fundamentals, algorithms and applications," in *Proc. IEEE ISCAS*,2004, vol. 2, pp. 33–36.

[3] R. Caldelli, F. Filippini, and R. Becarelli, "Reversible watermarking techniques: An overview and a classification," *EURASIP J. Inf. Security*,vol. 2010, Article 134546..

[4] Z.Ni, Y.Q. Shi, N. Ansari, andW. Su, "Reversible data hiding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 3, pp. 354–362, Mar 2006.

[5] S. K. Lee, Y. H. Suh, and Y. S. Ho, "Reversible image authentication based on watermarking," in Proc. IEEE ICME, 2006, pp. 1321–1324.

[6] W. L. Tai, C. M. Yeh, and C. C. Chang, "Reversible data hiding based on histogram modification of pixel differences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 6, pp. 906–910, Jun. 2009.

[7] W. Hong, "Adaptive reversible data hiding method based on error energy control and histogram shifting," *Opt. Commun.*, vol. 285, no. 2, pp. 101–108, 2012.

[8] Aleksej Avramovic, Branimir Reljin, "Gradient Edge Detection Predictor for Image Lossless Compression",52nd International Symposium ELMAR-2010,15-17 September 2010,Zadar,Croatia.

[9] S.R. Kodituwakku and U. S. Amarasinghe, "Comparison Of Lossless Data Compression Algorithms For Text Data ", Indian Journal of Computer Science and Engineering, Vol 1, No 4,pp-416-425.

[10] Rakesh Karmakar and Rajib Sarkar, "Implementation of Run Length Encoding on FPGA Spartan 3E", International Journal of Computer Science And Technology IJCST Vol. 5, SPL - 1,Jan - March 2014.

[11] Xiaolong Li , Weiming Zhang , XinluGui , and Bin Yang, " A Novel Reversible Data Hiding Scheme Based on Two-Dimensional Difference Histogram Modification, IEEE Transactions On Information Forensics And Security " , Vol. 8 , No. 7, pp. 1091-1100 ,July 2013.