An Image-in-Image communication scheme and VLSI implementation using FPGA

Santi P. Maity[1], Ayan Banerjee[2], and Malay K. Kundu[3]

[1,2] Department of Electronics and Telecommunication Engineering, Bengal Engineering & Science University, Shibpur, Howrah, India 711103
spmaity@telecom.becs.ac.in, ayan@telecom.becs.ac.in
[3] Machine Intelligence Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata, India 700 108 malay@isical.ac.in

Abstract— **The proliferation of the digitized media (audio, image and video) introduces a challenging problem of security in data transmission in the network environment. In this paper a novel algorithm that serves the purpose of covert image-in-image communication and its VLSI implementation using FPGA is proposed. Channel coding and spatial bi-phase modulation scheme are used to map gray-scale image information into a binary equivalent message that extracts message information from the noise less/noisy version of the cover data. Implementation of the proposed low cost algorithm may be speeded up signif*i*cantly by hardware realization and can be used for real time applications of multimedia data transmission. VLSI implementation using Field Programmable Gate Array (FPGA) has been developed for the algorithms and circuit can be integrated into the existing digital still camera framework.**

Index Terms: Channel coding, FPGA, spatial bi-phase modulation, VLSI design.

I. Introduction

With the advancement in digital techniques the ease of processing, storage and almost noise free transmission of digital multimedia signals over the general purpose channel are now easier; at the same time a class of problems have been emerged to maintain authenticity, integrity and security in digital data transmission [1]. Data encryption and data hiding are the two different techniques that are now being used widely to serve this security purpose. Data encryption or cryptographic techniques offer security by preventing an eavesdropper from accessing the original media. On the other hand, in data hiding scheme, security is provided through an imperceptible embedding of auxiliary message into the digital multimedia signal and the access of the original data is not protected.

In all practical data hiding methods the original image (data) is inevitably distorted by some small amount of embedding noise that can not be removed completely due to quantization, bit-replacement or truncation at the gray scales 0 and 255 [2]. Although the distortion is often quite small, it may not be acceptable for medical imagery (for legal reasons) or for military images inspected under unusual visual conditions (after extreme zoom)[3]. In this paper we propose an algorithm that is different from both data hiding and encryption principles in the sense that it neither embeds any data in the original media nor it prevents the access of the media. The algorithm allows faithful decoding of a message from a noiseless/noisy version of the cover image provided an auxiliary message is available at the user end by transmitting either (i) through a secured channel (cryptography or data encryption) or (ii) through an insecured channel subject to some distortion constraint.

In natural images a strong spatial correlation exists among the neighboring pixel values and the correlation is ref*l*ected by the successive binary run of larger lengths formed by the MSB (most signif*i*cant bit) planes of the gray levels. MSB plan offers relative stability i.e. each run length does not change much after various image processing operations, so long image visual quality is properly maintained. In this message transmission scheme the run length of MSB plane is used for signal modulation. Gray scale image is used as information signal so that decoded message preserves its recognizability when transmitted through noisy channel. Improvement in data transmission reliability is further increased by using channel coding scheme in the form of variable redundancy that is incorporated among the different bit planes based on their relative signif*i*cance. Spatial bi-phase modulation scheme reduces transmission overhead i.e. amount of data transmission that is increased in channel coding scheme.

The various signal processing algorithms can be implemented with either software or hardware where the latter offers advantages over the former in terms of less area, low execution time, and less power [4]. In any data hiding or image-in-image communication problem, if a chip is f*i*tted in the digital devices, the stego image/modulated signal can be obtained from the output video or images right at the origin. Hardware implementation using Field Programmable Gate Array (FPGA) offers real time application. FPGA has advantage of low investment cost and desktop testing with moderate processing speed [5].

The paper is organized as follows: Section II describes the algorithm and section III presents VLSI implementation using FPGA. Sections IV and V present some experimental results along with the applications, and conclusions respectively.

II. The proposed algorithm

The algorithm may be considered as a digital modulation scheme that employs synchronous detection for decoding of

message. The gray scale image is used as information bearing signal so that it not only conveys unique information but also shows a good degree of resiliency after various forms of image distortions. The cover image is considered as carrier, a binary equivalent image called as modulated signal is generated using spatial bi-phase modulation scheme. The modulated signal is transmitted either through a secured channel or through a noisy channel subject to a distortion constraint. Message is extracted at the receiver from this modulated signal using the cover image (carrier) or its noisy version (drift in carrier).

A. Transmitter operation

The 2-D pixel values of the cover image is converted to an 1-D signal. The channel for message encoding is formed from the MSB plane of the pixel values (string 1). The gray scale message signal is first mapped to a 1-D signal and converted into a binary string. An extended binary string (string 2) is formed by incorporating variable redundancy (repeating each bit by suitable odd number of times) onto the different bit planes of the message. Higher redundancy is assigned to the higher bit plane since they contain visually significant data and less or no redundancy for lower order bit planes that contribute more subtle details in the image [6]. Strings 1 and 2 are partitioned into sub strings having equal and fixed number of digits or symbol. If there occurs more than 50% positional match of the symbols in the two respective sub strings, a bit **1** is assigned for the sub string otherwise a bit **0**. Bit **1** indicates in-phase condition of two sub strings while out of phase condition is represented by bit **0**. Assigning a binary digit, corresponding to each substring of particular number of symbols, is called here as spatial bi-phase modulation technique. The process converts a gray scale message into a binary equivalent modulated signal. To quantify the merits of the combined effect of channel coding and spatial bi-phase modulation, a new string (string 3) is formed by keeping each sub string of string 2 unchanged or complemented bitwise based on the bit value **1** or **0** of the newly obtained binary string. The number of positional mismatch in the symbols between the strings 1 and 3 are counted and are divided by the total number of symbols in the string in order to calculate the probability of error denoted by $p(e)$. If a sub string consists of l (an odd number) of symbols and the string consists of total k number of such sub strings, $P(e)$ that denotes the probability of making wrong decision for all the sub strings can be expressed as follows:

$$P(e) = (\sum_{n}^{l} \binom{l}{n} {p_e}^n (1-p_e)^{l-n})^k \qquad (1)$$

where $n = (l+1)/2$. All k number sub strings are assumed to be independent among each other. Lower value of $P(e)$ indicates low message encoding loss and the value is related with the length of the sub string l which is again related with the size of binary message.

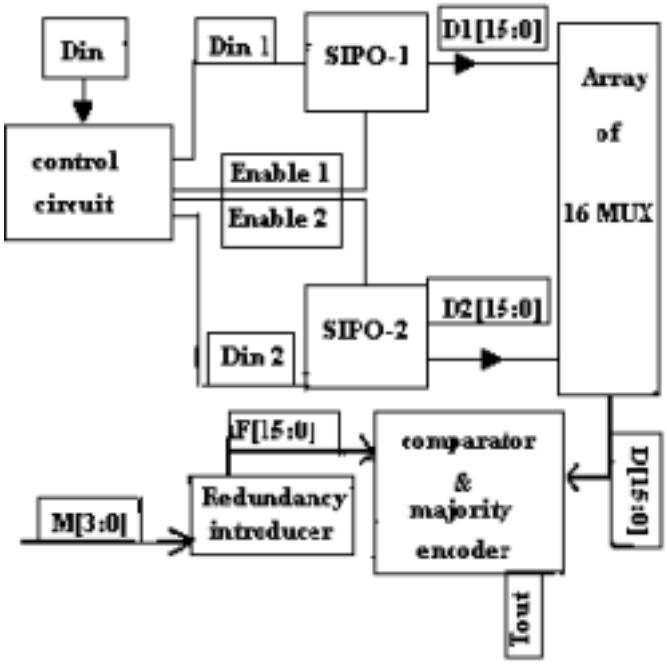

Fig. 1. VLSI architecture of transmitter

B. Receiver operation

The 2-D pixel values of the cover image or its noisy versions are converted to 1-D string. The MSB plane of this 1-D signal is picked up and partitioned into substrings of fixed and equal number of symbols. Each substring either remains unchanged or complemented based on the value of the received bit **1** or **0** respectively that represents the particular sub string. Each substring obtained after such operation is partitioned into sub substrings (smaller substrings) based on the degree of redundancies incorporated on the different message bits. Binary detection is then applied for each sub substring based on the majority decision rule i.e. if more than 50% symbols of a sub substring are **1**, decision for decoding is **1** otherwise **0**. The binary digits of all the sub substrings of a substring are then converted to the pixel values for the decoded message. The similar decoding process is applied for other sub strings and the message decoding is thus completed.

III. VLSI ARCHITECTURE

There are two main sub-blocks, one is the transmitter and the other one is the receiver.

A. Transmitter Architecture

The VLSI architecture of the transmitter for the proposed algorithm is shown in Fig. 1. The major sub blocks are con-

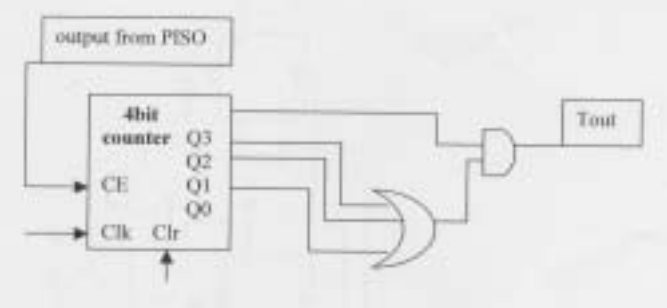

Fig. 2. Majority encoder block

trol circuit, serial-in-parallel out (SIPO) shift registers, multiplexer and majority encoder. In order to speed up the process, two SIPO operates in parallel. When one SIPO takes external input data, the other SIPO feeds data to the majority encoder unit and vice versa. Two data sets, each of sixteen bit length from the two SIPO, are fed to the multiplexer that outputs one data set to the majority encoder block. In the majority encoder block the other input is coming from the extended message obtained after adding redundancy. Two data sets of 16 bit are fed into an array of 16 XNORs for similarity comparison. The output from the similarity comparator block is then passed through parallel-in-serial-out shift register and fed into a 4 bit binary up counter to enable its clock. The output of the counter is fed into an encoder. For a string of 16 bits, if the similarity comparator output is **1** in 9 bit positions or more, then the counter value will be upgraded by the same amount. The counter is reset after every 16 bit string. The counter value is fed to an encoder block. The encoder is designed so that if its input value is nine or more, it will give at its output **1**, otherwise **0**. This encoder output is final modulated output from the transmitter.

Fig. 2 shows the detailed combinational logic for the majority encoder block. Here the output from the PISO (Parallel-in-serial-out) shift register is used as the clock enable signal for the 4 bit counter. Thus whenever the output from the PISO shift register is **1**, the output of the counter is incremented by **1** value. In order to distinguish between two substrings of length 16 bits a **Clr** terminal is used for the counter to reset it after 16 clock cycles. The terminal count of another 4 bit counter after being passed through a D flip-flop is fed to the **Clr** terminal. Thus the transmitter output **Tout** is obtained at an interval of 16 clock cycles with an initial delay of 32 clock cycles.

Fig. 3 shows the detailed design of the control circuit. The circuit is basically used to control the flow of data to the two serial-in-parallel-out shift registers. A 5 bit counter is used together with some combinatorial circuit to generate the necessary control signals. The data channel **din** is used as the input of the SIPO-1 for the first 16 clock cycles of the counter. At the 16-th clock cycle the data channel **din** is transferred to the input of the SIPO-2 and the enable signal meant for SIPO-1 is made high for one (1) clock cycle only. At the 31-st clock cycle the enable-2 is obtained by passing the terminal count of the counter through a D flip-flop. Thus as the counter again starts counting from zero the enable-2 line remains high for 1 clock cycle only and the data channel is now transferred to SIPO-2. One set of data (of size 16 bits) is read from the data file through one SIPO shift register, the previous set of data can be obtained from another SIPO shift register and it leads to faster operation.

B. Receiver Architecture

The VLSI architecture of the receiver for the proposed algorithm is shown in Fig. 4. In the receiver unit the main sub-blocks are (i) control unit, (ii) SIPO-1 and SIPO-2, (iii) controlled complementer, and (iv) redundancy remover. Control unit, SIPO-1 and SIPO-2 work in same fashion as in the transmitter. The controlled complementer block consists of sixteen 2:1 multiplexer and sixteen inverters to generate sixteen image bits. The design of the controlled complementer circuit is shown in Fig. 5.

The function of the multiplexer is to select any one out of two available input (let A and B) at a time and send this to the output. Two 2-input AND gates (let G1 and G2) and one 2-input OR gate (G3) perform the function. A is connected to the input terminal of G1 and B is connected to the input terminal of G2. One select signal S is connected to the second input terminal of G2 directly and in inverted manner to the second input terminal of G1. The output terminals of G1 and G2 are connected to the two input terminals of G3. When S=0, G1 yields A and G2 yields 0 due to basic AND operation. G3 will produce A at the circuit output due to basic OR operation. By similar logic, when S=1 circuit output is B.

The controlled complementer circuit receives 16 data bits D[15:0] in parallel and coded bit **Tout** as inputs. 16 data bits D0.....D15 are connected in parallel to the first input terminal of the 16 multiplexers. *Tout* is connected to the select terminal of the muxs. Second input terminal of each mux receives the inverted logic level of the first input. Each mux is designed in such a manner that when select input is **1** the first input will be selected, otherwise the second input will be selected. When **Tout=1** it implies that data and images bits are the same, otherwise opposite to one another. The logic reveals that when **Tout=1** data bit appears at the output of the circuit as the image bit. Thus 16 image bits F[15:0] are obtained from 16 multiplexers.

The output from the controlled complementer is fed to the redundancy remover unit. The 16-bit substring is divided into 4 parts of 9 bits, 5 bits and two lower bits. There are two 4 bit counter, one for removing redundancy from 9 bits and the other one from the next 5 bit. The string of 9 bits is fed to a counter to enable its clock and if the counter value is 5 or more, it is further encoded to give an output **1**,

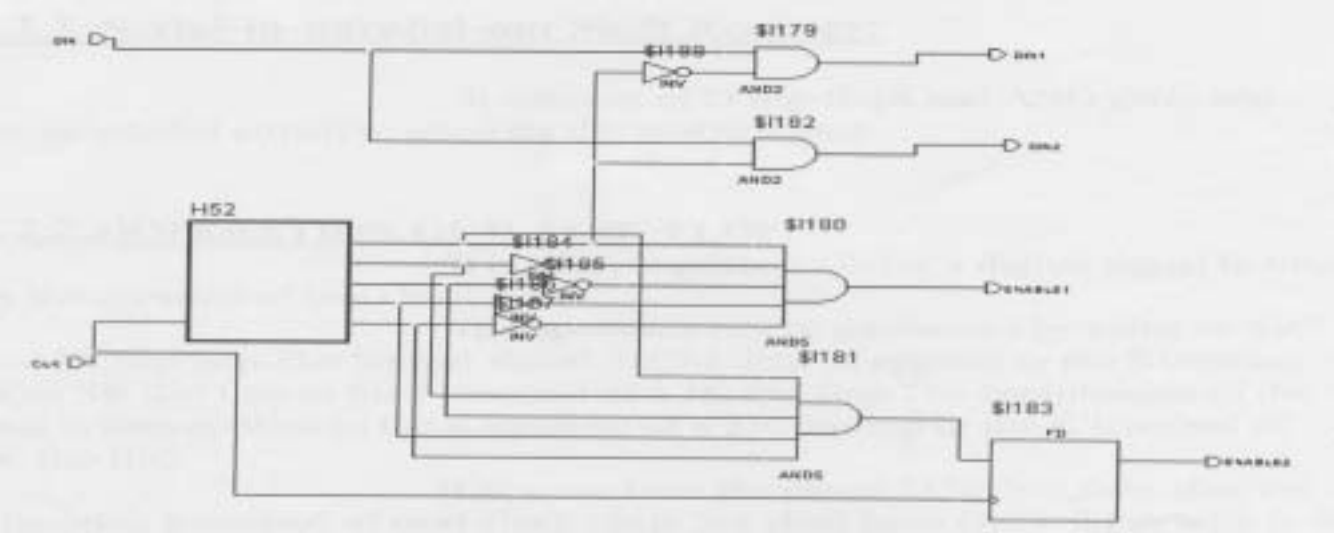

Fig. 3. Control circuit

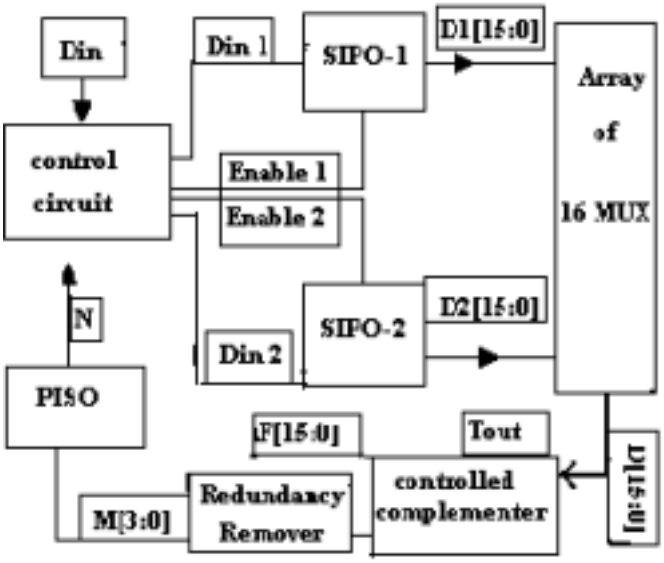

Fig. 4. VLSI architecture of receiver

otherwise **0**. Similarly, redundancy is removed from the bit string of five bits. Thus each substring of 16 bit is converted to a 4 bit after redundancy removal. By converting each 16 bit substring to 4 bit string, the image is decoded.

IV. Results and discussion

The message signal is a gray scale image of size (64×64), 4 bits/pixel and the cover image (carrier) is also a gray-scale image of size (256×256) 8 bits/pixel. Each sub string consists of 16 symbols and the size of the binary modulated signal is (64×64). The performance of the algorithm has been tested both in terms of the various forms of image distortions and noise corruption of the modulated signal analogous to the concept of drift in local oscillator carrier frequency and channel noise in synchronous detection of digital communication. It has been seen that extracted messages are quite recognizable even after greater depth of degradations like linear and non linear spatial filtering, sharpening, histogram equalization, lossy JPEG and JPEG 2000 compression, noise addition, rescaling etc. occurred in the cover image. It is also found that decoded message is well recognized even after 30% change in binary modulated signal. The quality of the extracted message is quantified by mutual information value $I(X;Y)$ where random variables X and Y represents the transmitted and decoded message respectively. If $p(x_i)$ represents the probability of occurrence of the i-th pixel value in the transmitted image and $p(y_j/x_i)$ represents the channel transition matrix, $I(X;Y)$ that represents the average amount of information received from the signal degradation, can be expressed as follows [7]:

$$I(X;Y) = \sum_i \sum_j p(x_i)p(y_j/x_i) \log \frac{p(y_j/x_i)}{\sum_i p(x_i)p(y_j/x_i)} \tag{2}$$

where i, j represent the index of the symbols.

Fig. 6(a), 6(b) and 6(c) represent the cover image Fishing Boat [8], the message and the extracted message respectively where there is neither any degradation for the cover image nor any form of noise disturbance for the modulated signal. The entropy of the message, shown in Fig. 6(b), is 0.867 while the $I(X;Y)$ value for the decoded message, shown in Fig. 6(c), is 0.624. The $I(X;Y)$ values for the decoded messages are 0.51 and 0.56 when the cover image is mean and median filtered with PSNR (peak signal to noise ratio) values 22.56 dB and 25.36 dB respectively. The values for the same index are 0.48 and 0.41 when the cover

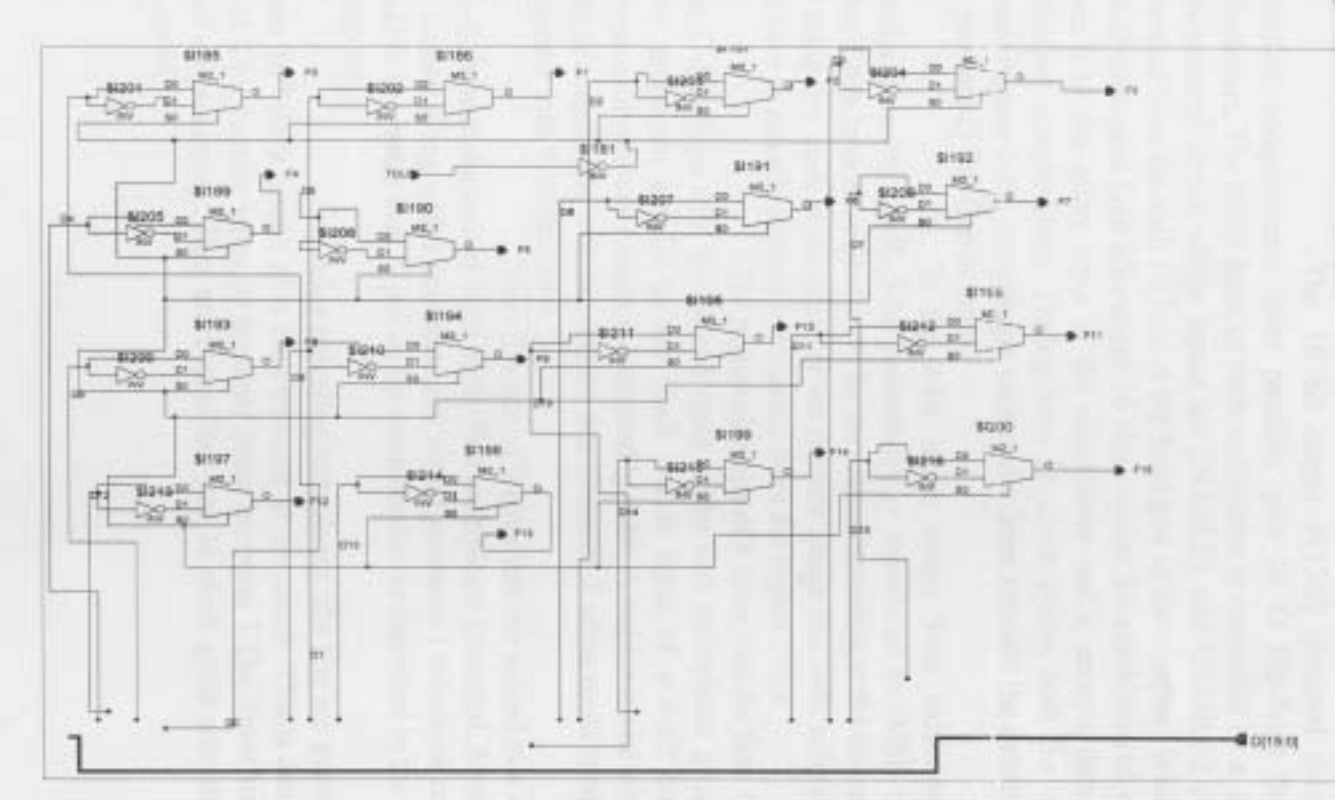

Fig. 5. Controlled complementer circuit

Fig. 6. (a) Cover image; (b) message signal; (c) decoded message

image is compressed by JPEG and JPEG 2000 operations at quality factor 25.

Noise sensitivity of the proposed algorithm for various possible signal processing operations are shown in Figs. 7(a)-(t). The signal processing operations include various linear and non linear filtering such as mean, median and gaussian filtering, histogram equalization, image sharpening, additive gaussian and speckle noise, dynamic range change, image rescaling, cropping, least significant bit manipulations, lossy compression operations such as JPEG, JPEG 2000, image dithering, shearing, warping and wiener filtering etc. available in checkmark package [9]. Table I shows the message encoding loss i.e. loss in the process of binary message formation based gn the size of sub string length. The results in the table reflects the fact that probability of error for 3-rd bit is higher compared to 4-th bit as amount of redundancy incorporated for the latter is higher compared to the former.

TABLE I
PROBABILITY OF ERROR IN SINGLE BIT, 4-TH AND 3-RD BIT

Length of sub string	Prob. of bit error $p(e)$	Prob. of wrong decision in 4th/3rd bit $P(e_1)$ & $P(e_2)$
4	0.153442	0.001663& 0.019957
16	0.178758	0.003734&0.033853
64	0.259437	0.017599& 0.074379
256	0.35301	0.073181& 0.184462

The hardware design is implemented using XILINX SPARTAN series FPGA. The chip used is XCS05 which contains 100 CLB (configurable logic block) out of which 85 are consumed, 40 CLBs for the transmitter and 45 CLBs for the receiver. Important specifications are summarized below in Table II.

TABLE II
SPECIFICATION OF HARDWARE REALIZATION

Implementation	CLB count	clock freq. (max)	clock cycle
XCS05	85	80 MHz	17 cycles/message pixel value (4 bits)

V. CONCLUSIONS

An algorithm for covert image-in-image communication and its VLSI realization using FPGA is proposed in the paper. Faithful decoding of a message is possible provided an auxiliary message is available at the receiver end by transmitting either through a secured channel or through a channel subject to 30% noise impairment to the auxiliary message. The algorithm can be applied to unchangeable image, multi-owner host image sharing, medical imagery, low power verification systems, military images inspected under unusual visual conditions etc. Algorithm requires few simple computations and VLSI implementation using FPGA allows it application for real time multimedia data transmission. Current work is going on to develop the dedicated digital system using this FPGA chip.

VI. REFERENCES

[1] B. Macq, Special issue on identification and protection of multimedia information, *Proc. IEEE* 87: 1673-1687, 1997.

[2] M. Golijan, J. J. Fridrich, R. Du, Distortion-Free data embedding for images, *4th Information Hiding Workshop*, LNCS vol. 2137, Springer-Verlag, New York, pp. 27-41, 2001.

[3] C. C. Chang, J. C. Chuang, An image intelligent property protection scheme for gray-level images using visual secret sharing strategy, *Pattern Recognition Letters*, vol. 23, pp. 931-941, 2002.

[4] N. J. Mathai, D. Kundur, and A. Sheikholeslami, Hardware implementation perspectives of digital video watermarking algorithms, *IEEE Transaction on Signal Processing*, vol. 51, pp. 925- 938, April 2003.

[5] W. K. Chen, *The VLSI Handbook*, CRC Press, August 2000.

[6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley, New York, 1992.

[7] B. P. Lathi, *Modern Digital and Analog Communication Systems(Third Edition)*, Oxford University Press, New Delhi, India,1999.

[8] http:// www.cl.cam.ac.uk/ fapp2/watermarking.

[9]http://www.cl.cam.ac.uk/ fapp2/watermarking/stirmark.

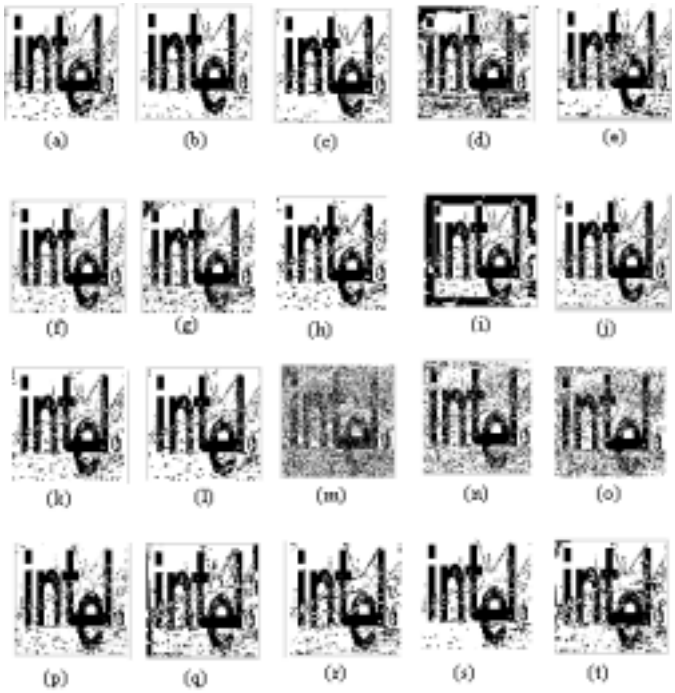

Fig. 7. (a) Decoded message after three times mean filtering of the cover image (carrier) using window size 3x3, (b) Decoded message after three times median filtering of the cover image using window size 3x3, (c) Decoded message after five times gaussian filtering of the cover image with variance 1, window size 9x9, (d) Decoded message after histogram equalization of the cover image, (e) Decoded message after image sharpening, (f) Decoded message after noise addition, (g) Decoded message after change in dynamic range from 252- 4 to 200-50, (h) Decoded message after image rescaling, (i) Decoded message after image cropping operation, (j) Decoded message after least significant bits manipulation, (k) Decoded message after JPEG compression at quality factor 30, (l) Decoded message after JPEG 2000 compression at quality factor 30, (m) Decoded message after image dithering, (n) Decoded message after additive gaussian noise with variance 0.01, (o) Decoded message after speckle noise with variance 0.04, (p) Decoded message after sample down up operation, (q) Decoded message after image shearing operation, (r) Decoded message after stirmark operation, (s) Decoded message after wiener filtering, (t) Decoded message after image warping operation.