An Efficient Hardware Architecture for Multimedia Encryption and Authentication using the Discrete Wavelet Transform

Amit Pande and Joseph Zambreno
Department of Electrical and Computer Engineering
Iowa State University, Ames, IA 50011
Email: {amit, zambreno}@iastate.edu

***Abstract*—This paper introduces a zero-overhead encryption and authentication scheme for real-time embedded multimedia systems. The parameterized construction of the Discrete Wavelet Transform (DWT) compression block is used to introduce a free parameter in the design. It allows building a keyspace for lightweight multimedia encryption. The parameterization yields rational coefficients leading to an efficient fixed point hardware implementation. A clock speed of over 240 MHz was achieved on a Xilinx Virtex 5 FPGA. Comparison with existing approaches was performed to indicate the high throughput and low hardware overhead in adding the security feature to the DWT architecture.**

Index terms: Parameterization, Discrete Wavelet Transform, Multimedia encryption, Watermarking.

I. Introduction

The Discrete Wavelet Transform (DWT) has enabled research in image and video coding [1], [2], [3], [4] and has become a part of multiple next generation multimedia compression and transmission standards [5], [6]. The increasing importance of the DWT in image and multimedia compression applications has inspired the development of efficient hardware for implementations. Figure 1 shows some constraints in the design of a DWT filter. It must provide a high compression ratio and image reconstruction quality so as to serve the end user requirements. Some other desired features include low hardware cost, low power requirements and high throughput of the system.

Recent research [7], [8] addresses these issues by choosing a simple DWT filter design with fewer adders and other hardware, thereby optimizing the hardware architecture for high performance. In [9], we provide a DWT-based architecture to enable real-time video-streaming applications such as those used in tele-medicine [10], remote laboratories [11], educational video streaming [12] and video surveillance. The DWT filters presented in [7] give a polymorphic hardware support for the real-time requirements of multimedia applications. A summary of state-of-the-art implementations of the DWT on custom hardware is provided in the references [13], [14] and [15]. Existing implementations do not suit the security demands of real-time multimedia systems.

However, a parameterized DWT implementation can fulfill these requirements. The re-design of the DWT filter can meet the security requirements in addition to providing a perfect

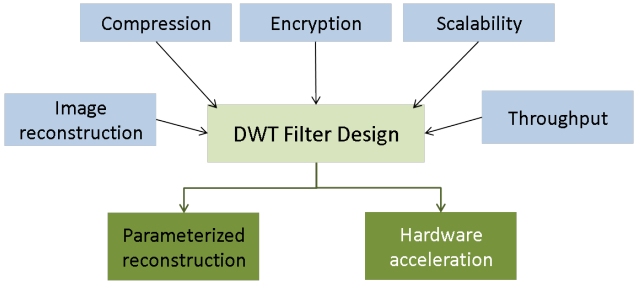

Fig. 1. DWT Filter Design Constraints

image reconstruction and high compression framework for video compression. This paper introduces a new layout and configuration scheme for the parameterized DWT that enables lightweight multimedia encryption and authentication.

The existing popular encryption algorithms such as AES and RSA have large computational requirements. Hardware implementations of AES are often pipelined, leading to a significantly large latency for real-time applications (31 cycles for AES [16]). Most of the present schemes use variations of the basic scheme presented in Figure 2(a). Encryption algorithms such as AES, DES or IDEA are typically applied over the full or partial output bit stream obtained from the compression engine.

Video compression and data encryption are both computationally expensive tasks. The scheme presented in Figure 2(a) restricts a custom hardware design for the DWT that requires low power consumption and hardware usage. Such a design also limits an efficient delivery of scalable video streams. These restrictions can be alleviated by developing a scheme that integrates both encryption and compression operations into one without any significant computational overheads. This concept is presented in Figure 2(b). A light-weight encryption block is built into the compression engine.

Next, we consider an example to explain the significance of lightweight multimedia encryption schemes for embedded systems.

In Figure 3, a surveillance aircraft (A) is sending aerial surveys and other important information to the ground troops (B), crucial for their attack on the enemy base (C). In this scenario, typical encoding schemes would require large computational

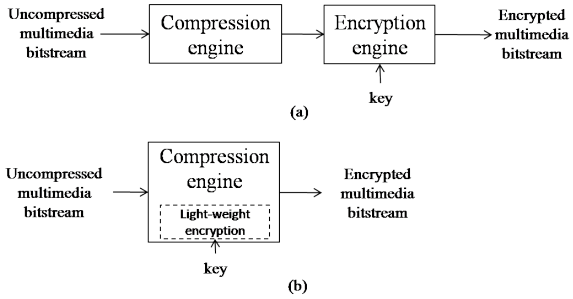

Fig. 2. (a) Traditional scheme for multimedia encryption and (b) Lightweight multimedia encryption scheme

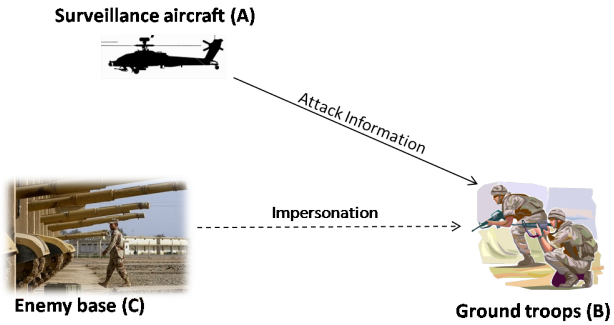

Fig. 3. An example scenario for proposed lightweight multimedia encryption scheme

resources and hence high power consumption making them unsuitable for real-world embedded systems. Moreover, such conventional ciphers would incur a large latency in image transmission which can be critical for ground troops' (B) operation. Some of the crucial security issues involved in this case are as follows:

1) The message (image) sent by A must not be easily perceptible to B.
2) B must be able to authenticate the incoming message (from A) to avoid impersonation from C.

The power-constraints of A serve as bottleneck for the use of strong ciphers like AES, and RSA. However, the power-constraints of the adversary C in a real-world scenario, and the time-crucial nature of transmitted information allows us to propose a solution based on lightweight encryption and watermarking authentication [17], [18]. The lightweight encryption scheme can provide a reasonable degree of security with little or no overhead in power or other requirements.

In this paper we present a new parameterized construction of a DWT filter with rational coefficients. The parameterized construction can be used to build a key scheme while the rational coefficients of the DWT enable an efficient hardware architecture using fixed point arithmetic. The DWT, an essential part of modern multimedia compression algorithms, thus serves as a compression-cum-encryption block. The main contributions of this work can be summarized as follows:

1) We introduce the concept of the parameterized DWT architecture for multimedia encryption. The new DWT architecture implements DWT as an encryption operation.
2) We optimize and pipeline the hardware architecture to achieve a high clock frequency of 242 MHz with minimum hardware requirements.
3) We provide some experimental results of image encryption and watermarking using the parameterized DWT operation.

The rest of the paper is organized as follows: Section II gives a brief introduction to the DWT. Section III provides the parameterized construction of the DWT to yield a free parameter a in DWT operation. A keyspace is built for video encryption using the DWT and is explained in Section IV. The rational coefficients in the parameterized DWT allow us to build an efficient hardware architecture which is explained in Section V. Section VI presents the experiments in multimedia security and synthesis results for the DWT architecture on a Xilinx Virtex 5 FPGA. Section VII concludes the paper and provides future work.

TABLE I
APPROXIMATE COEFFICIENTS FOR THE DAUBECHIES 9/7 FILTER

i	$h_1(i)$	$h_2(i)$
± 4	0.026748757411	0
± 3	-0.016864118443	-0.045635881557
± 2	-0.078223266529	-0.0287717631145
± 1	0.266864118443	0.295635881557
0	0.602949018236	0.557543525

II. PRELIMINARIES

Prior works in signal processing establish that the 1-D DWT can be viewed as a signal decomposition using specific low pass and high pass filters [19]. A single stage of image decomposition can be implemented by successive horizontal row and vertical column wavelet transforms. Thus, one level of DWT operation is represented by filtering with high and low pass filters across row and column successively. After each filtering downsampling is done by a factor of 2 to remove the redundant information.

The two most common DWT filters used in image compression are the Le Gall's 5/3 filter and the Daubechies' 9/7 filter [5], accepted in the JPEG2000 standard. The Le Gall's filter has rational coefficients and its hardware implementation requires less resources. The Daubechies' 9/7 filter has better compression performance, however, it has irrational coefficients and leads to lossy compression. Applying a 2-D DWT to an image of resolution $M \times N$ results in four images of dimensions $\frac{M}{2} \times \frac{N}{2}$. Subsequent levels of the DWT-based decomposition yield a multi-resolution structure suitable for image compression.

III. PARAMETERIZED DWT DERIVATION

This section discusses the rational coefficient parameterized construction of the 9/7 DWT filter to serve as the backbone for the new DWT architecture. The irrational coefficients in

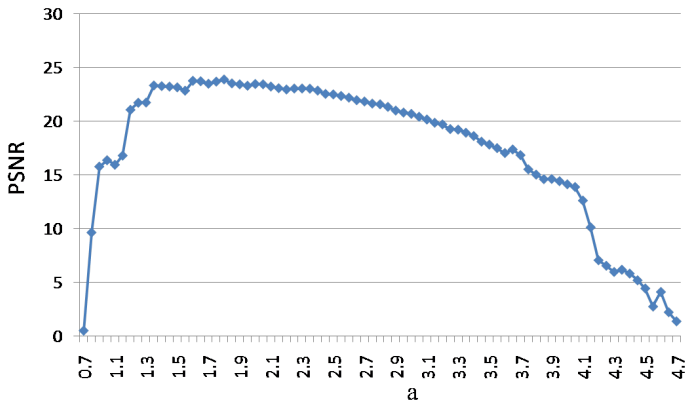

Fig. 4. Variation in PSNR with parameter a

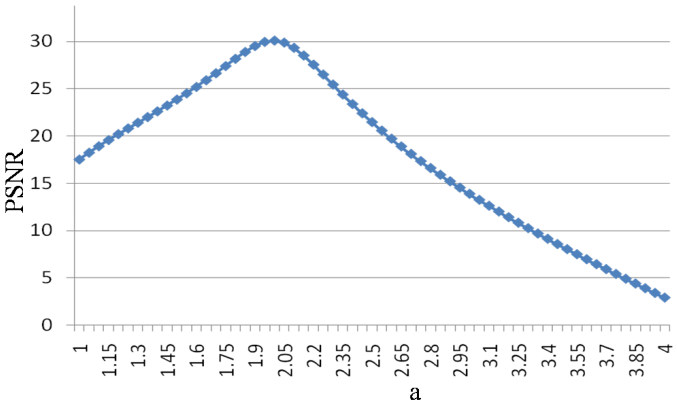

Fig. 5. PSNR of reconstructed image encoded with $a = 2.0$ and decoded at variable values of a

Daubechies' 9/7 filter limit its precision of implementation on fixed point hardware such as FPGAs and ASICs.

The Bi-orthogonal Wavelet Filter Banks are used in image compression because of their excellent image compression properties. They must satisfy Perfect Reconstruction (PR) condition and are desired to have a large number of Vanishing Moments (VMs) to have good approximation property [20]. The Daubechies 9/7 filter has good compression property and is being used in wavelet-based image compression standards. It has 4 VMs each for the analysis and synthesis low pass filters. Liu et al. [21] discusses the derivation of rational coefficients filter for the DWT with arbitrary number of taps. The parameterization is achieved by reducing two VMs in the filter expression to introduce a free parameter a in the design.

Let $H_1(z)$ and $H_2(z)$ denote the analysis and synthesis low pass filter coefficients. On introducing a free parameter a in the equations for $H_1(z)$, the corresponding value of $H_2(z)$ is obtained by solving for conditions for linear phase, PR and low pass filter [21].

$$\begin{aligned} H_1(z) &= \left(z^{\frac{1}{2}} + z^{\frac{1}{2}}\right)^4 \times \left(a + (1-a)\left(z^{\frac{1}{2}} + z^{\frac{1}{2}}\right)^2\right) \\ H_2(z) &= \left(z^{\frac{1}{2}} + z^{\frac{1}{2}}\right)^2 \times Q(z) \end{aligned}$$

where

$$Q(z) = \sum_{n=0}^{3} q_n \times \left(z^{\frac{1}{2}} + z^{\frac{1}{2}}\right)^{2n}$$

Liu et al. [21] find an approximate expression for the rational representation of these coefficients (q_n). They are given by:

$$\begin{aligned} q_0 &= 1; \\ q_1 &= 5 - 2 \times a \\ q_2 &= 4 \times a^2 - 14 \times a + 16 \\ q_3 &= 36 \times a - 8 \times a^2 - 60 + 32/a \end{aligned}$$

Simplifying these expressions, we get the following expression for $H_1(z)$ and $H_2(z)$.

$$\begin{aligned} H_1(z) = (-9/64a + 1/32a^2 + 15/64 - 1/8/a)(z^4 + 1/z^4) \\ +(-1/16a^2 + 11/32a - 11/16 + 1/2/a)(z^3 + 1/z^3) \\ +(1/8 - 1/2/a)(z^2 + 1/z^2) \\ +(-11/32a + 1/16a^2 + 15/16 - 1/2/a)(z + 1/z) \\ +(9/32a - 1/16a^2 - 7/32 + 5/4/a) \end{aligned}$$

$$\begin{aligned} H_2(z) = (1/32 - 1/32a)(z^3 + 1/z^3) + (1/8 - 1/16a)(z^2 + \\ 1/z^2) + (7/32 + 1/32a)(z + 1/z) + (1/4 + 1/8a) \end{aligned}$$

The rational terms in the expressions for these filters can be implemented in hardware using shifts and adds instead of multiplication operations. This is a big savings over the original Daubechies filter in terms of hardware requirements. However, we need to perform multiplication with the free parameter a and its exponents. This filter is implemented in our DWT architecture and is explained in the following sections.

IV. Multimedia Security using the DWT

In this section, we give a brief summary of the security perspective of parameterized DWT filter. The overall scheme for multimedia authentication and encryption as well.

A. Building the Keyspace

The number of DWT operations N in an image of size $M \times M$ pixels is bounded by the limit $N \leq log_e(M)$. For example, we can obtain up to nine levels of wavelet decomposition for an image of size 512×512 pixels. One level of wavelet decomposition involves two filtering operations: one each along the row and column directions. Thus, we can choose up to 18 different a values, one each for the 18 different instances of DWT kernels being used in the operation.

Figure 4 shows the variation in PSNR of the reconstructed image at a bitrate of 0.2 bpp using the SPIHT coder with variations in the parameter a. The test image *Lena* was used for this simulation. The variations of a beyond the range of 1 to 3 yields a poor PSNR value, indicating poor compression of the coefficients. Thus, the parameter a can be varied

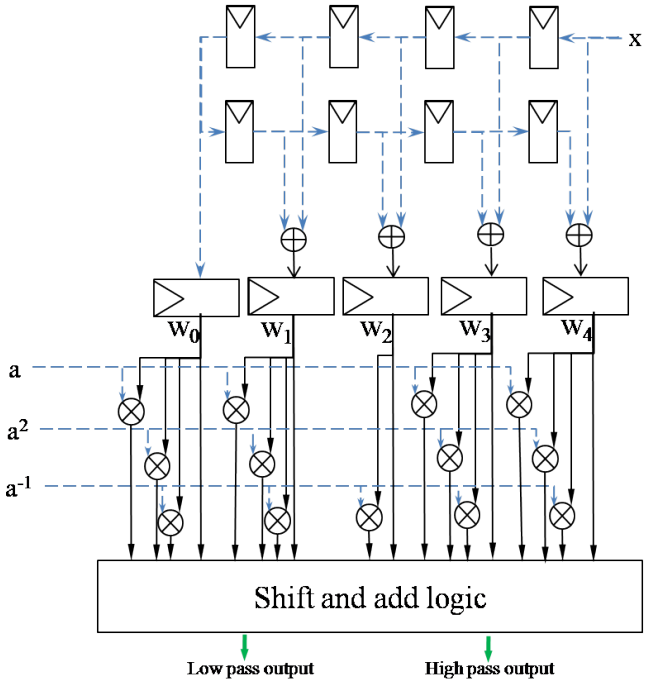

Fig. 6. Overview of Discrete Wavelet Transform architecture

between numerical range of 1 to 3 while yielding satisfactory compression.

We can use 8 bits to divide the interval from 1 to 3 into subintervals of 0.008. Hence, it will take 8 bits to represent one a parameter. This gives us a keyspace of 144 bits for a 9 level decomposition.

Figure 5 gives the PSNR performance of image encoded with $a = 2.0$ and decoded with a variable value of a.

B. Multimedia Security

The main advantage of the lightweight encryption scheme is that, while maintaining competitive compression performance and providing security, it comes at extremely low computational overhead. [17] uses a wavelet filter parameterization scheme to provide key dependency to a blind watermarking algorithm. Similarly, the 144-bit keyspace can be used to encrypt input frames. This level of security can suffice for the soft encryption requirements of mobile multimedia applications [18] and surveillance applications as previously mentioned.

C. Multimedia Authentication

Multimedia authentication can be achieved by embedding a watermark in a suitable DWT subband level. A survey of common watermarking schemes is presented in [22].

V. DWT Architecture and Design

Figure 6 gives the overview of our parameterized DWT architecture. The input data (one pixel input per cycle) x is pipelined for eight cycles. We observe that z^i and z^{-i} values in the expressions for $H_1(z)$ and $H_2(z)$ have the same coefficients. Thus, these values can be added together to simplify

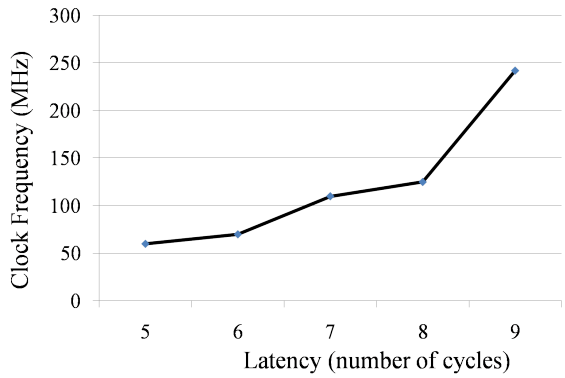

Fig. 7. Variation in Clock Frequency versus System latency (in number of cycles)

further computations. In Figure 6, eight of the nine inputs are passed through four adders to reduce the number of variables to five. These values (labeled w_0, w_1, w_2, w_3 and w_4) are multiplied with a, a^2 and a^{-1} to get the necessary intermediate values which are input to the shift and add logic. In this block, we perform shifts and add operations to implement additions and multiplications with rational fractions. The high and low pass filter coefficients are the final output of the DWT filter.

We performed several optimization steps to reduce the cost of the underlying hardware. They are summarized below:

1) Division by binary coefficients (e.g. $1/64, 1/16, 1/4$) was performed using arithmetic shift operations. This eliminates the need for multipliers in the circuits and reduces the number of multipliers in the circuit from 69 to 23.
2) Observe in low and high pass filter coefficients (see previous section) that the coefficients of $z^{\pm k}$ are the same. Thus they both can be grouped together to reduce the hardware complexity. These coefficients are labeled as w_0, w_1, w_2, w_3 and w_4 in Figure 6. This optimization gives a tremendous savings in hardware. It reduces the number of adders in the design from 70 to 41 and the number of multipliers from 23 to 13.
3) The input stream was pipelined. As shown in Fig. 6 our architecture takes one pixel (or channel input) as the input and outputs the low and high pass signal coefficients with a finite latency. Increasing the system latency allows us to achieve a higher clock speed (and hence higher throughput). This trade-off is plotted in Figure 7.

VI. Experiments

We targeted a Xilinx XC5VLX330 FPGA for our experiments, using ModelSim 6.4 for simulation purposes and Xilinx ISE 10.1 for synthesis. The area and performance results for the parameterized DWT implementation are summarized in this section.

Direct implementation of the Daubechies 9/7 filter gave a clock frequency of 107 MHz, while requiring 16 9x9 bit multiplier units. However, a fixed point implementation of the

	Our Design	[7]	[8]	Daubechies 9/7
Slices Flip flops	649	245	-	210
Multipliers	13	0	0	16
Adders	41	9	19	15
Registers	92	208	-	144
Clock Frequency	242.85 MHz	390 MHz	200 MHz	107

TABLE II
HARDWARE UTILIZATION OF THE DWT ARCHITECTURE ON XILINX VERTEX XCVLX330 FPGA

DWT leads to image reconstruction error and gives no security promise.

Our new architecture inputs an eight bit block every cycle (one pixel value). The initial parameterized DWT design obtained a clock frequency of approximately 60 MHz, due to its long critical path. The critical path of the circuit lies from the w_i registers to the final low pass output. We then pipelined this computation into several stages and obtained a faster implementation. By adding 4 pipelining stages we obtained a clock frequency of 242 MHz. The design used 13 10-bit multipliers, 41 adders (20 18-bit adders and 21 9-bit adders). The hardware requirements of our implementation are summarized and compared with other implementations in Table II. It is noteworthy that ours is the first implementation of the parameterized DWT filter in hardware (to the best of knowledge of the authors). Thus, this comparison with other reported architectures only indicate the overheads involved in building a secure DWT scheme.

Figure 8 shows the image performance of the parameterized DWT. We took three sample images: the first and third being an aerial survey of some landscape while the second image is a snapshot of Shakespeare's written text (Scene II from Julius Caesar). The results are presented when an encryption (or image compression) was performed with the a parameter set to 2.0 and decryption (or image reconstruction) was performed with different a values. We can see that the images decrypted with the wrong key values (Fig. 8 (b, d, e)) have poor visual quality. These images miss many important details of the original scene or text. In this experiment, we have visualized the impact of only using the parameterized DWT and a single key for all levels of decomposition.

The degradation of image details for wrong key values can be crucial for applications such as the defense application considered in the introduction. Multimedia processing tasks such as object recognition and tracking are of significance to video surveillance applications and require these high level details. Thus, lightweight encryption provided by the DWT can be useful for a variety of commercial applications.

Figure 9 shows the watermark authentication performance of the parameterized DWT. The watermark was inserted into the fourth level decomposition with DWT in high pass coefficients. It was observed that while the original watermark was obtained when decoding with same value of a (2.0 in this case), a distorted watermark was obtained when we decoded with wrong values of a. These initial investigations indicate that we can build a watermarking authentication scheme using our parameterized DWT.

VII. CONCLUSION AND FUTURE WORKS

This paper introduces a multimedia encryption and watermark authentication framework based on parameterized construction of DWT. The parameterization enables an efficient, pipelined, high throughput implementation in hardware. The qualitative and quantitative results in terms of both hardware performance and image security promise a secure framework for real-time multimedia delivery over embedded systems.

Future works includes further exploration into the security promises of the parameterized DWT operation and the trade-offs involved in compression performance of such systems.

The parent-child coding gain in the DWT-based coders was quantified by Marcellin et al. [23] indicating that the coding gain due to subband rotation dependencies is not considerable. However, the rotations of subbands will considerably affect the visual performance of images and significantly improve the keyspace for multimedia encryption. The idea of parameterization can also be extended to other multimedia encoding blocks to obtain a more powerful integrated-encryption-scheme for embedded multimedia systems.

REFERENCES

[1] A. Said and W. Pearlman, "An image multiresolution representation for lossless and lossy image compression," *IEEE Transactions on Image Processing*, vol. 5, pp. 1303–1310, 1996.

[2] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, Dec. 1993.

[3] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, Jul 2000.

[4] R. Qiu and W. Yu, "An Efficient Quality Scalable Motion-JPEG2000 Transmission Scheme," Department of Computer Science, Washington University in St. Louis, Tech. Rep. WUCS-01-37, November 2001. [Online]. Available: citeseer.ist.psu.edu/qiu01efficient.html

[5] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system: an overview," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 4, pp. 1103–1127, Nov 2000.

[6] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 9, pp. 1103–1120, Sept. 2007.

[7] A. Pande and J. Zambreno, "Design and analysis of efficient reconfigurable wavelet filters," in *Proceedings of the IEEE Intl. Conf. on Electro Information Technology*, 2008, pp. 337–342.

[8] M. Martina and G. Masera, "Multiplierless, folded 9/7 5/3 wavelet VLSI architecture," *IEEE Transactions on Circuits and Systems II*, vol. 54, no. 9, pp. 770–774, Sep. 2007.

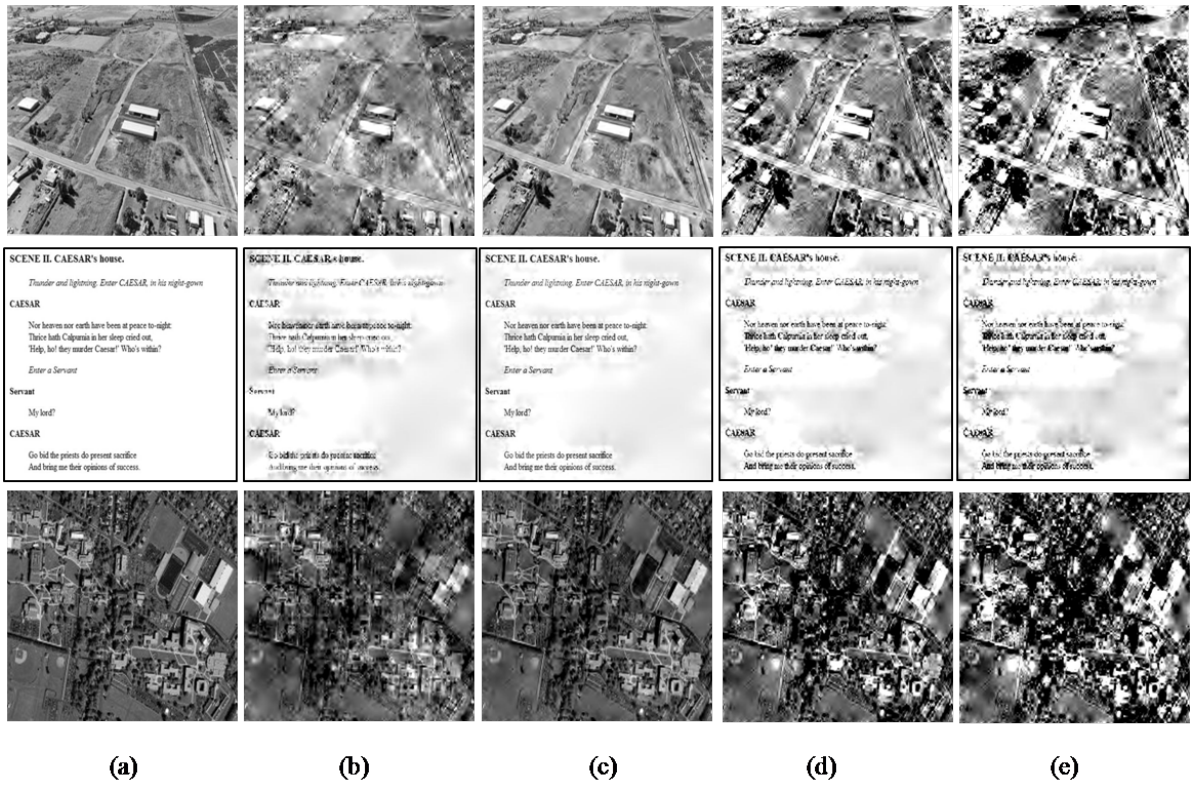

Fig. 8. Image reconstruction with different keys. (a) show the original images which are then encrypted with $a = 2$, (b)-(e) show reconstruction with $a = 1, 2, 3$ and 3.5 respectively

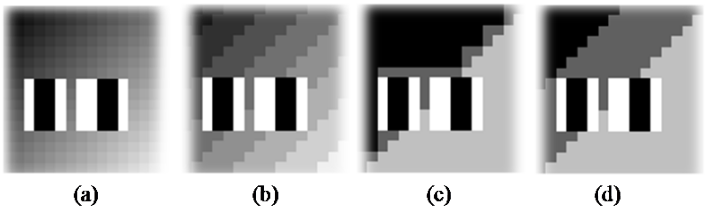

Fig. 9. Watermark retrieved for test image at 0.1 bpp with encoder $a = 2$. (a) shows the original watermark, (b-d) show the watermarks detected with $a = 2.1, 3, 3.5$ respectively

[9] A.Pande and J. Zambreno, "Polymorphic wavelet architecuture over reconfigurable hardware," in *IEEE Intl. Conf. on Field Programmable Logic and Applications*, 2008, pp. 471–474.

[10] M. Leeser, S. Miller, and Y. Haiqian, "Smart camera based on reconfigurable hardware enables diverse real-time applications," in *Proceedings of the IEEE Symposium on Field-Programmable Custom Computing Machines (FCCM)*, Apr. 2004, pp. 147–155.

[11] A. Mittal, A. Pande, and P. K. Verma, "Content-based network resource allocation for mobile engineering laboratory applications," in *Proceedings of the Intl. Conference on Mobile Learning*, 2007, pp. 146–152.

[12] A. Mittal, S. Gupta, S. Jain, and A. Jain, "Content-based adaptive compression of educational videos using phase correlation techniques," *ACM/ Springer Multimedia Systems*, vol. 11, no. 3, pp. 249–259, 2006.

[13] P. Tseng, Y. Chang, Y. Huang, H. Fang, C. Huang, and L. Chen, "Advances in hardware architectures for image and video coding - a survey," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 184–197, Jan. 2005.

[14] C. Chakrabarti, M. Vishwanath, and R. M. Owens, "'A Survey of Architectures for the Discrete and Continuous Wavelet Transforms."

[15] T. Acharya and C. Chakrabarti, "A Survey on Lifting-based Discrete Wavelet Transform Architectures," *Journal of VLSI Signal Processing Systems*, vol. 42, no. 3, pp. 321–339, 2006.

[16] A. Hodjat and I. Verbauwhede, "A 21.54 Gbits/s fully pipelined AES processor on FPGA," pp. 308–309, April 2004.

[17] M. Brachtl, A. Uhl, and W. Dietl, "Key-dependency for a wavelet-based blind watermarking algorithm," in *Proceedings of the 2004 workshop on Multimedia and security MM&Sec 2004:*. New York, NY, USA: ACM, 2004, pp. 175–179.

[18] D. Engel and A. Uhl, "Parameterized biorthogonal wavelet lifting for lightweight JPEG 2000 transparent encryption," in *MM&Sec '05: Proceedings of the 7th workshop on Multimedia and security*, 2005, pp. 63–70.

[19] G. Strang and T. Nguyen, "Wavelets and filter bank," 1996.

[20] M. Vetterli and J. Kovačevic, *Wavelets and subband coding*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.

[21] Z. Liu and N. Zheng, "Parametrization construction of biorthogonal wavelet filter banks for image coding," *Signal, Image and Video Processing*, vol. 1, no. 1, pp. 63–76, 2007.

[22] D. Zheng, Y. Liu, J. Zhao, and A. E. Saddik, "A survey of RST invariant image watermarking algorithms," *ACM Comput. Surv.*, vol. 39, no. 2, p. 5, 2007.

[23] M. Marcellin and A. Bilgin, "Quantifying the parent-child coding gain in zero-tree-based coders," *Signal Processing Letters, IEEE*, vol. 8, no. 3, pp. 67–69, Mar 2001.