Securing Multimedia Content using Joint Compression and Encryption

Amit Pande, Prasant Mohapatra, *Fellow, IEEE* and
Joseph Zambreno, *Member, IEEE*

Abstract—Algorithmic parameterization and hardware architectures can ensure secure transmission of multimedia data in resource-constrained environments such as wireless video surveillance networks, tele-medicine frameworks for distant health care support in rural areas, and Internet video streaming.

Joint multimedia compression and encryption techniques can significantly reduce the computational requirements of video processing systems. We present an approach to reduce the computational cost of multimedia encryption, while also preserving the properties of compressed video (useful for scalability, transcoding, and retrieval), which endanger loss by naive encryption. Hardware-amenable design of proposed algorithms makes them suitable for real-time embedded multimedia systems. This approach alleviates the need of additional hardware for encryption in resource-constrained scenario, and can be otherwise used to augment existing encryption methods used for content delivery in Internet or other applications.

In this work, we show how two compression blocks for video coding: a modified frequency transform (called as Secure Wavelet Transform or SWT) and a modified entropy coding scheme, (called Chaotic Arithmetic Coding (CAC)) can be used for video encryption. Experimental results are shown for selective encryption using proposed schemes.

Index Terms—Video encryption, joint coding and encryption, embedded systems, Discrete Wavelet Transform, Arithmetic Coding

1 Introduction

In this work, we discuss the design of algorithms and hardware architectures for secure transmission of multimedia data in resource-constrained environments. Some typical application scenarios include wireless video surveillance networks, telemedicine frameworks for distant health care support in rural areas, and Internet video streaming.

Wireless Video Surveillance Networks: Wireless video surveillance networks have been recently deployed in different network settings such as WiFi, WiMAX, and wireless sensor networks. These networks are deployed in private and public settings, and carry sensitive visual information. For example, the live feeds from over 10,000 cameras are used by city police departments in the NYC and Chicago areas to monitor criminal activity. It is important to protect the video feeds of these cameras from eavesdropping. However providing real-time end-to-end encryption of video data using conventional cryptographic primitives is difficult due to a) wireless network characteristics (low bandwidth, frequent packet drops), b) QoS (real-time delivery, low jitter), and c) the limited computational resources at the encoder. Apart from the need of a secure way of transmitting videos, we need computationally-efficient algorithms to save on computing power and also enable easy access of visual information from encrypted videos in databases.

Tele-medicine Frameworks: Tele-medicine is an application of clinical medicine where consultation, and even remote medical procedure and examinations are performed using interactive audio-visual media. Extending such services to remote locations (which lack high-speed connections and even electric power in under-developed countries) requires

- *A. Pande and P. Mohapatra are with Department of Computer Science, University of California, 2063 Kemper Hall, Davis CA-95616, USA. Phone: (530) 752-0870, Fax: : (530) 752-4767 e-mail: {amit, prasant}@cs.ucdavis.edu*
- *J. Zambreno is with Department of Electrical and Computer Engineering, 2215 Coover Hall, Iowa State University, Ames, IA-50011, USA. e-mail: zambreno@iastate.edu*
- *This research is supported by the National Science Foundation Grant NSF#2331914 and Grant #1019343 to the Computing Research Association for the CIFellows Project.*

efficient low-power devices. Further, the privacy of patient information and prescriptions is an important concern for these applications, considering the vulnerability of communication channels against eavesdropping and other attacks.

Internet and Mobile Video: The advent of embedded multimedia systems has already revolutionized the way we live. Video messaging, video-conferencing, video surveillance and Internet video sites such as YouTube are increasingly becoming popular and pervasive. Network traffic in next-generation cellular networks is predicted to be dominated by video [1] and it makes sense to provision for security of videos in these applications. Most mobile devices have low computational resources and limited battery resources.

Conventional encryption schemes such as those using AES and DES are not suitable for video data because of the large computational overhead. Compressed multimedia streams also exhibit well-defined hierarchical structure that can be exploited in several useful ways (e.g. scalability, random access, transcoding, rate shaping) in low and variable bandwidth scenarios - these structures would not be recognizable in traditional ciphertext.

In this work, an augmented video coding model is used for joint compression and encryption which can significantly reduce the computational requirements. We propose to build design blocks which enable security for these applications at the algorithmic level, and leave domain-specific optimization to application developers. These algorithmic optimizations map easily to fixed point hardware, allowing us to come up with efficient architectural optimizations for resource constrained scenarios. In other application scenarios, these approaches can complement the security provided by conventional schemes such as AES.

The proposed schemes are also low-cost in the sense that the required computational hardware is considerably smaller than existing approaches, and in some configurations the hardware resources are fewer than that for conventional video compression schemes.

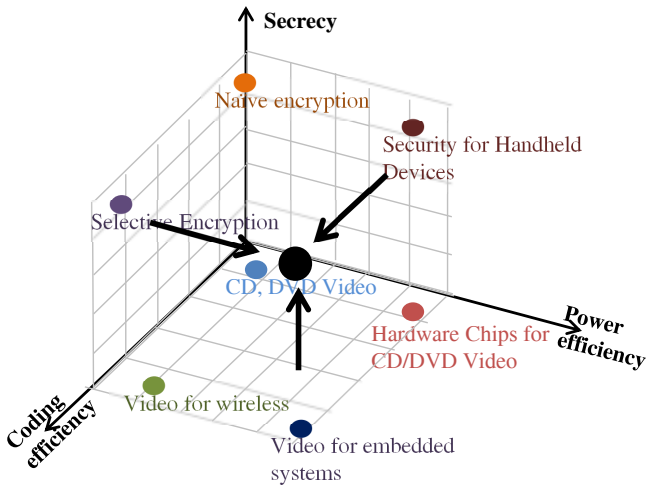

Fig. 1. The broad goal of joint approach is to develop algorithms and architectures to push the operating curve towards (1,1,1), considering all the factors (Coding Efficiency, Power Efficiency and Privacy) together during design

2 Basics

Multimedia compression involves large computations and large amount of data-transfers thus requiring application-specific hardware such as ASICs and FPGAs to compress and deliver the media in real-time. Video compression using hardware accelerators has gained increased attention because of the popularity of low-power embedded devices. *Thus, an efficient architectural design of multimedia compression blocks is a must to ensure real-time video delivery.*

While the compressed multimedia files typically exhibit well-defined hierarchical structure that can be exploited in several useful ways (e.g. for scalability, random access, transcoding, rate shaping), these structures are not recognizable in ciphertext, and hence, are wasted. These properties are useful to index, search and retrieve compressed multimedia from digital libraries and also for communication over heterogeneous networks. We need a paradigm where encryption does not change the compressed output, yet provides access and copy control for concerned media. *Thus, we need encryption of video data without affecting the properties of compressed bitstream, or affecting the compression performance.*

On one hand, compression and encryption operations require large amount of computational overhead, while on the other hand, there has been an increasing trend towards deployment of battery-driven low-power embedded systems such as portable mo-

bile devices (iPods, mobile phones, and cameras). *Apart from optimizations in hardware architectures, we also need to reduce the computation cost for secure multimedia transactions through algorithmic improvements.* In Figure 1 we illustrate the motivation of our proposed approach: considering coding efficiency, power efficiency and privacy in a joint design of algorithms and architectures.

Related Research Efforts:

The research in video coding for the last five decades has been commercially utilized in the form of state-of-the-art video coding standards such as MPEG-1, MPEG-2 and so on. MPEG-2 based schemes are useful for DVD quality compression. In these scenarios, coding or power efficiency are not the constraints and security was provided using end-end encryption with AES or some variants [2]. Some work has been done for resource optimization in these schemes in cases of low-bandwidth but the problem is not so acute.

Recent research in wireless networks and video surveillance [3], [4] aims at optimizing the video quality for wireless transmission and often uses the MPEG-4 format which produces a more compressed and scalable bitstream. The H.264 SVC format is the most recently used in this work. Many researchers have also tried to optimize the hardware implementations of MPEG-2 and MPEG-4 based video applications.

Recent research in video encryption over wireless and other scarce resource channels has identified the need for non-traditional approaches to video encryption besides the use of standard cryptographic ciphers. These approaches involve selective or partial encryption of video stream, chaotic encryption and shuffling in compressed bitstream etc. There has also been research to accelerate these video processing kernels in hardware such as ASIC or FPGA.

Thus, there has been little research which targets the three-fold goal of high compression, low computational cost, and secrecy. With these three goals in mind, we propose our approach in the next section.

3 OUR APPROACH

What we propose in this work is a redesign of the video compression blocks themselves to enable encryption and efficient mapping onto hardware. For example, if the video coders have an additional parameter which can be changed to provide encryption, we can use it as a keyspace for secret key generation. The required mixing of the inputs, as required by cryptographic ciphers, is automatically provided by different blocks of video coding system. Similarly, if we could design the system with the rational coefficients as a design constraint, we will obtain a hardware-amenable implementation.

The redesign of video coding blocks enables joint compression and encryption and also reduces the computational requirements of multimedia encryption algorithms. The approach modifies the compression system properties instead of the compressed bitstream itself. Moreover, the redesign is amenable to hardware acceleration over reconfigurable computing platforms. We leverage signal processing techniques to make the algorithms suitable for hardware optimizations (and encryption), and reduce the critical path of circuits using hardware-specific optimizations.

A trivial way to explain this solution (of joint encryption and compression) is to find 2^N different but similar ways to compress a video, where all of them give similar compression performance and the compressed bitstream has the same properties. For large values of 2^N, we can say that the N bit code representing the choice of compression system is the encryption key of the system. In order for such a system to be secure, the combined system must follow cryptographic requirements such as good diffusion and confusion properties [5]. The output from two closely related keys should be nearly uncorrelated and there should not exist a way to reverse-engineer the N bit key except by a brute-force attack.

This proposal also meets the requirements of property-preserving encryption because essentially we are trying to shuffle the compression parameters using the key and not modifying the input bitstream itself. Each of the 2^N compression systems provide 'property-preserving' compression.

How exactly we can augment encryption to video coding system?

This is achieved by redesign of individual video coding blocks followed by integration into a single prototype and hardware implementation (Figure 2). These modifications are briefly described below:

(a) **Augmented Prediction Model:** We propose to use a fuzzy prediction model, which selects from several past and future frames and uses

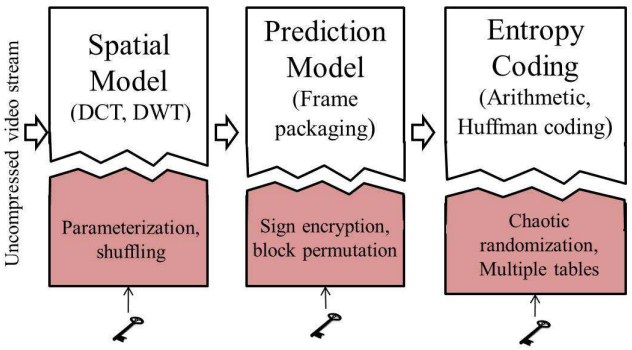

Fig. 2. Video compression system augmented with different operations to ensure real-time encryption

multiple streams, based on a key-dependent fuzzy logic instead of the traditional use of immediate neighbors. Similarly, the sign bits of the motion vectors can be encoded and/or a key-based non-linear mapping of motion vectors can be performed.

(b) **Augmented Spatial Model:** We propose to parameterize the transform filter (DWT), so that the choice of filter depends on key value. Different filters give different output coefficients while the compression efficiency of each is similar. The output sub-bands of DWT (or sub-blocks for DCT) can be re-oriented and permuted according to a key.

(c) **Augmented Entropy Coding:** Modified entropy coders can be used with multiple statistical models so that the exact choice of model is governed by a key. Similarly, arithmetic coding can be implemented using a key-based chaotic random map. The re-iterations of chaotic map make the output appear random, while the choice of map itself is governed by a key.

(d) **All-on-a-Chip:** Hardware-specific optimizations on augmented modules will enable us to fit the prototype on a single chip.

4 DETAILS

In this section, we discuss two augmented compression modules to make them amenable to encryption and hardware implementation. Augmented Frequency Transform is illustrated with the help of Discrete Wavelet Transform which is used in MotionJPEG2000 coder while augmented entropy coding is shown using arithmetic coding which is used in H.264 and other formats.

4.1 Augmented Frequency Transform

What is DWT? The efficient representation of time-frequency information by the Discrete Wavelet Transform (DWT) has led to its popularity for spatial modeling. DWT provides superior rate-distortion and subjective image quality performance over existing standards. Many image and video compression schemes have been derived from DWT-based structures which have become increasingly popular because of excellent compression properties. The 1-D DWT can be viewed as a signal decomposition using specific low pass (H_0) and high pass (H_1) filters. A single stage of image decomposition can be implemented by successive horizontal row and vertical column wavelet transforms. To recover the image back, we perform the inverse DWT using another set of low and high pass filters G_0 and G_1 respectively. The two most common filters for DWT are Le Gall's 5/3 filter and Daubechies 9/7 filter [6].

Existing Efforts for Hardware-Amenable Implementation. Rational binary coefficients for DWT have been designed to help in achieving a multiplier-free implementation of DWT filter coefficients. However, these multiplier-free implementations involve image reconstruction quality trade-offs. Many other researchers have also faced the problem of reducing DWT complexity.

Augmenting the DWT for Hardware Implementation. Rather than optimizing the mapping of filter coefficients, we want to re-derive the filter coefficients which map better to custom hardware.

For example, the Quadrature Mirror Filter (QMF) properties of DWT allows perfect reconstruction of image after inverse DWT operation at decoder. The condition for perfect reconstruction of image, when applying DWT boils down to:

$$G_0(z)H_0(z) + G_0(-z)H_0(-z)) = 2$$

Solving this equation using Lagrange Half Band Filters (LHBF) leads to Daubechies filter, which is the most widely-used DWT filter. However, these coefficients are irrational and lead to inefficient hardware implementation.

Tay et al. [7] derive rational coefficients for DWT, setting this as a constraint while solving these

TABLE 1
Hardware amenable implementation of DWT: Over 50% improvement in hardware requirements with increased clock frequency (Testbed: Xilinx Virtex-V XC5VLX30 FPGA)

Features	Daub 9/7	**Poly-DWT**	Tay, 01	Kotteri, 05	Chao, 03	Martina, 07	Martina, 05
Adders	15	**9**	19	15	8	19	21
Multipliers	9	**0**	0	0	4	0	0
Clock(MHz)	107	**389**	-	-	-	200	-

TABLE 2
Improved image reconstruction (PSNR values) with hardware-amenable DWT implementation

Image	**Bitrate=0.5 bpp**			**Bitrate=2 bpp**		
	Daub. 9/7	**Poly-DWT**	Martina,07	Daub. 9/7	**Poly-DWT**	Martina,07
lena	28.213	**29.46**	27.7	38.47	**38.17**	36.5
surveillance	26.1	**28.1**	26.54	38.41	**42.21**	39.21
lecture	34.35	**33.8**	32.73	48.3	**51.25**	43.71
helicoptor	33.75	**35.7**	35.01	48.59	**54.72**	47.14

equations. In our preliminary work, we used this result to build a Polymorphic DWT (Poly-DWT) architecture which uses binary rational coefficients which are amenable to hardware implementation. We also added some features which help Poly-DWT provide dynamic response to changing external conditions and thus dynamically adjust video quality and power requirements.

The hardware prototype of the proposed system on a Xilinx Virtex-V XC5VLX30 FPGA has the following features for hardware-amenable implementation [8]:

- The new architecture enables dynamic allocation of hardware resources to efficiently create a dynamic response to changing external conditions.
- Our architecture is multiplier-free. Further, its hardware requirements (9 adders) are nearly 50% that of existing architectures in the research literature (Table 1), while its image compression performance is better than fixed-point implementation of the state-of-the-art (Table 2).
- A switching scheme to allow runtime switching between 5/3 and 9/7 wavelet structures was proposed. Our architecture enables 'on-the-fly' switching of hardware resources to suit the power budget of the video processing system.

Augmenting the DWT for Both Encryption and Hardware. Engel et al. 2005 [9] parameterize DWT but the key space is small and restricted. Moreover, any hardware implementation issues are not discussed. Motivated by the successful parameterization of DWT for hardware implementation, we investigated the use of parameterization for encryption purposes.

The previous parameterization has only 2-3 rational points, but for use of parameterization for encryption, we need a family of rational coefficient filters. Zaide et al. [10] presents a parameterized construction of the filters typically used for image compression.

We obtain a new parameterization with two interesting features: (1) it has a free parameter α which can be varied as a key parameter without sacrificing the perfect reconstruction property of the video stream. The terms of α are non-linear which makes it difficult to obtain the α value back from the transform output. (2) Also, all other coefficients are rational and amenable to hardware implementation. These properties, along with the property of subband rotation were used to design a video encryption scheme optimized over hardware referred to as the Secure Wavelet Transform (SWT) [11]. A key-space of $25N + 3$ bits can be obtained from N levels of wavelet decomposition. For an image size of 512×512 pixels this upper limit of N is 9.

The hardware prototype of the proposed system on a Xilinx Virtex-V XC5VLX330 FPGA has the following features for hardware-amenable implementation:

(a) The DWT kernel was parameterized to incor-

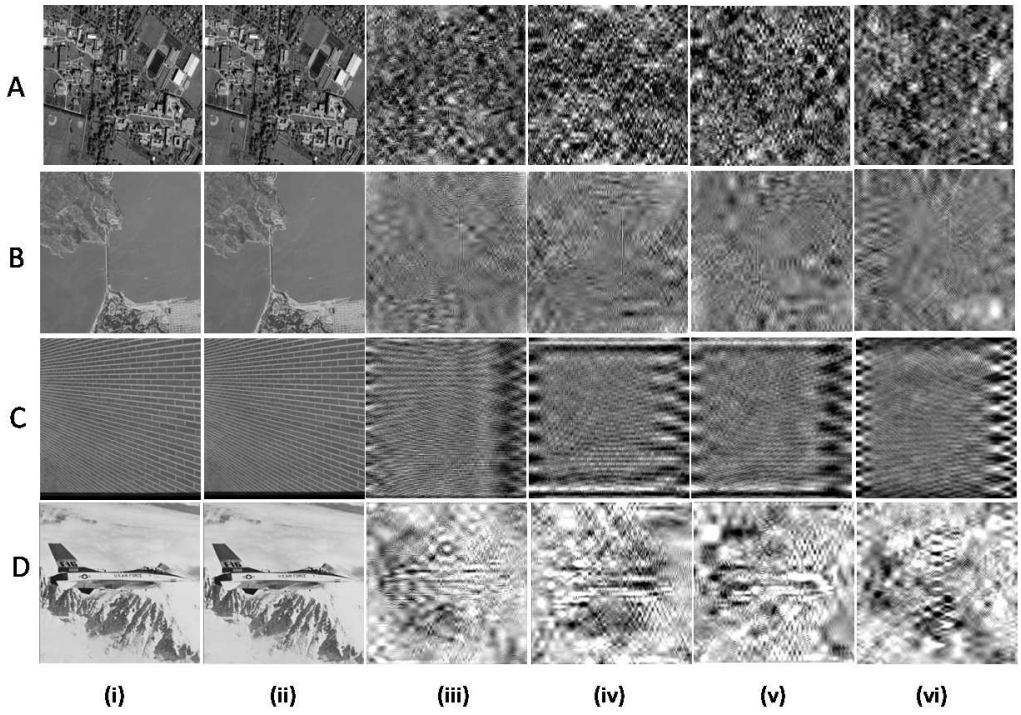

Fig. 3. Video encryption with augmented DWT (i)- Original image encrypted with key-0, (ii)- Image decrypted with same key, (iii)-(vi)- Image decrypted with randomly generated keys.

TABLE 3
A first architecture for video encryption on DWT (Testbed: Xilinx Virtex XCVLX330 FPGA)

	SWT	Martina, 07	Daub. 9/7	Jou, 01	Vishwanath, 95	Huang, 04
Multiplier	**0**	0	16	12	36	12
Adder	**11**	19	15	16	36	16
Critical Path	$4T_a + T_l$	$5T_a$	$T_m + 4T_a$	$T_m + 2T_a$	$T_m + 4T_a$	$4T_m + 8T_a$
Frequency	**114**	200	107	-	-	-
Encryption	**YES**	NO	NO	NO	NO	NO

Note: T_m, T_l and T_a are the time delay in multiplier, look-up table and adder circuits respectively.

porate the encryption feature and promise reasonable security for real-time embedded multimedia systems.

(b) A zero computation overhead subband reorientation scheme added to parameterization, leads to efficient image encryption (see Figure 3)

(c) An optimized hardware implementation of the SWT architecture is presented. The proposed hardware implementation has low critical path and thus achieves a high clock frequency (Table 4.1).

4.2 Augmented Entropy Coding

What is Arithmetic Coding (AC)? Entropy coding schemes are used to compress data, in a lossless manner, to a maximum level with the assumption of an independent and identically distributed random variable distribution. The two most popular entropy coding techniques are Huffman coding and arithmetic coding. Of these, Huffman coding is computationally cheap, while arithmetic coding yields better compression. Arithmetic coding involves recursive partitioning of the range [0,1) in accordance with the relative probabilities of the occurrence of the input symbols.

Existing efforts in AC-based encryption In [12], a chaos-based adaptive arithmetic coding technique was proposed. The arithmetic coder's statistical model is made varying in nature according to a pseudo-random bitstream generated by coupled chaotic systems. Many other techniques based on

varying the statistical model of entropy coders have been proposed in literature, however these techniques suffer from losses in compression efficiency that result from changes in entropy model statistics and are weak against known attacks [13]. Recently, Grangetto et al. [14] presented a Randomized Arithmetic Coding (RAC) scheme which achieves encryption by inserting some randomization in the arithmetic coding procedure at no expense in terms of coding efficiency. RAC needs a key of length 1-bit per encoded symbol. Kim et al. [15] presented a generalization of this procedure, called as Secure Arithmetic Coding (SAC). The SAC coder builds over a Key-Splitting Arithmetic Coding where a key is used to split the intervals of an arithmetic coder, adding input and output permutation to increase the coder's security. Successful attacks have been demonstrated against these SAC schemes.

Augmenting Encryption to AC. We generalize the arithmetic coder (as we did in the case of DWT) to get multiple ways of encoding without losing compression efficiency. An interesting observation made in [16] is the equivalence between arithmetic coding and chaotic maps. We first interpreted Arithmetic Coding (AC) in terms of iterations over piece-wise linear chaotic maps and then defined a family of such maps, each yielding the same compression efficiency. We next developed a data (image or video) encryption scheme based on arithmetic coding, which we refer to as Chaotic Arithmetic Coding (CAC). CAC uses a key to make the exact choice of map from the family of predefined maps to perform AC.

Let us consider a scenario where we have a string $S = x_1, x_2, ...x_N$ consisting of N symbols to be encoded. The probability of occurrence of a symbol $s_i,\ i \in 1, 2, ...n$ is given by p_i such that $p_i = N_i/N$ and N_i is the number of times the symbol s_i appears in the given string S. We next consider a piece-wise linear map (ρ) with the following properties:

It is defined on the interval $[0, 1)$ to $[0, 1)$.It can be decomposed into N piece-wise linear parts such that each part maps the region on x axis $[beg_k, end_k)$ to the interval $[0, 1)$. This mapping is one-one and onto. Each linear map is associated uniquely with one symbol and doesn't intersect with another. The mapping between linear map and assigned symbol from dictionaryis defined arbitrarily but one-one relationship must hold. The width of the map (on x-axis) is equivalent to probability of assigned symbol,

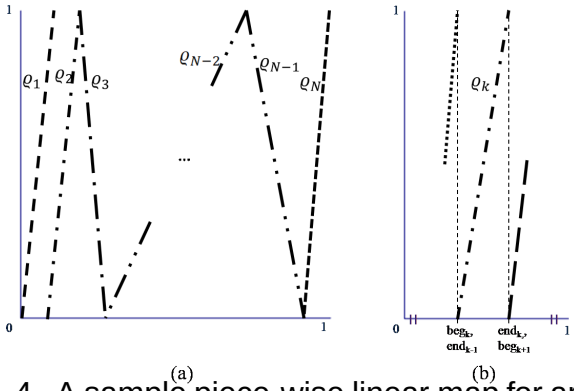

Fig. 4. A sample piece-wise linear map for arithmetic coding like compression (a) The entire map is shown (ρ) (b) A single linear part of the map (ϱ_k) is zoomed. It can have a positive or negative slope depending on choice

leading to Shannon optimal compression efficiency for large strings.

Figure 4 shows a sample map fulfilling these properties. Figure 4(a) shows the full map with different parts $\varrho_1, \varrho_2, ...\varrho_N$ present while Figure 4(b) zooms into individual linear part ϱ_k. The maps are placed adjacent to each other so that each input point is mapped into an output point in the range $[0, 1)$.

There are N different piece-wise maps (for encoding a symbol from dictionary of N symbols, called as N-ary AC), each with two possible orientations (with positive or negative slope). Thus, the number of total permutations possible is given by $N!2^N$. Thus, for N-ary arithmetic coding or arithmetic coding with N symbols, it is possible to have $N!2^N$ different mappings each leading to same compression efficiency. Since we can arbitrarily choose any 1 of the $N!2^N$ maps, the key space for encoding a single bit of data is $\left\lceil \log_2(N!2^N) \right\rceil$ bits, where $\lceil\rceil$ represents the greatest integer function. For N=2, it gives 8 mappings. If we increase N to 4 this value increases to 384. Thus, without any sacrifice in computational efficiency or coding rates, CAC is able to achieve a huge keyspace which can be effectively used for data encryption.

AC is more commonly implemented in binary mode to reduce the computational requirements of video coders. There are eight equivalent modes of skewed binary maps which can be used for BCAC or Binary CAC.

Implementation Efficiency For a normal binary arithmetic coder, at each iteration the starting in-

terval $[I_s, I_e)$ is updated at one end. On encoding a '0' the final interval becomes $[I_s + p(I_e - I_s), I_e)$ while on encoding a '1' the final interval becomes $[I_s, I_s + p(I_e - I_s)$. Thus, every iteration requires one multiplication and two addition operations. The decoding procedure for a binary arithmetic coder involves updating the interval $[I_s, I_e)$ at one end depending on whether the last decoded symbol was a '0' or a '1'. Thus, every iteration again requires one multiplication and two addition operations.

For chaotic arithmetic encoder, both end of interval are updated at every iteration using a linear transformation $x = my + b$ thus requiring two multiplications and two additions for encoding. The decoding is simple as it involves iteration on the chaotic map according to the linear transformation $y = nx + c$ involving a multiplication and an addition operation. There are some additional table lookups (an 8-input LUT required for BCAC to choose the exact chaotic map) involved in chaotic coding to choose the right chaotic map at every iteration which can be efficiently implemented in software or hardware. Thus, CAC encode requires more computations than BAC encode while CAC decode requires less computations than BAC decode.

Our BCAC coder is similar to a BAC coder except the variable slope and intercept of the lines of chaotic map which is decided by choice of map. These values can be mapped to look-up tables and the remaining operation can be optimized similar to BAC.

Comparison with BAC+AES

The arithmetic operations required for one bit encoding or decoding using BAC is 2 adders and 1 multipliers (discussed in Section 4.2). AES-128 bits require 40 sequential transformation steps composed of simple and basic operations such as table lookups, shifts, and XORs. It needs approximately 336 bytes of memory and approximately 608 XOR operations or roughly 3 bytes memory and 5 XOR operations per bit of encoding.

BCAC coder requires 2 adders and 2 multipliers for encoding and only 1 adder and 1 multiplier for decoding. Thus, the hardware requirements of BCAC coder are much less than BAC and AES combined. The BCAC decoder is particularly simpler than AC decoder (even without AES), which is desired for most common video applications which involve real-time decoding in mobile and embedded devices.

Compatibility with H.264

The popular H.264 codec implements Content-Adaptive Binary Arithmetic Coding (CABAC) to achieve high compression efficiency. A CABAC coder has three parts: Binarizer, context modeler and a BAC coder. We can directly replace the BAC coder with BCAC coder, as proposed in earlier section to introduce augmented entropy coding into H.264 without introducing any coding losses. Because there is no change in dictionary values or probability value p_i of coder, there will be no direct effect on video coder by replacing BAC with BCAC.

A detailed analysis of CAC scheme is given in [17].

5 SECURITY

5.1 Key-space

The keyspace for augmented DWT operation depends on number of DWT decompositions and the degree of parameterization of one filter coefficients. There are 3N+1 different sub-bands obtained by a N level wavelet decomposition of an image/ frame (19 sub-bands for N=6). Three bits are each required to describe the orientation of each sub-band, leading to $9N + 3$ bits. Further, we divide the filter parameter in range $[1, 3]$ into 2^6 values. One level of wavelet decomposition involves successive filtering with row and column filters. If we have N levels of decomposition using DWT, we can choose different α values for all $2N$ filters (represented by $12N$ bits). SWT operation allows a keyspace of $21N + 3$ bits. For SD and HD images/ video content, the value of N is sufficiently large ($N > 9$) to give a large keyspace.

The N-bit BCAC coder requires a keyspace of 3N bits. The same key may be re-used for different iterations, hence we can have the key-size depending on the length of BCAC coder.

5.2 Sample Image encryption scheme

In this experiment, we use BCAC and SWT schemes to encrypt only the most significant DWT coefficients for an image. We considered sample images of dimension 512x512 pixels and encoded only the

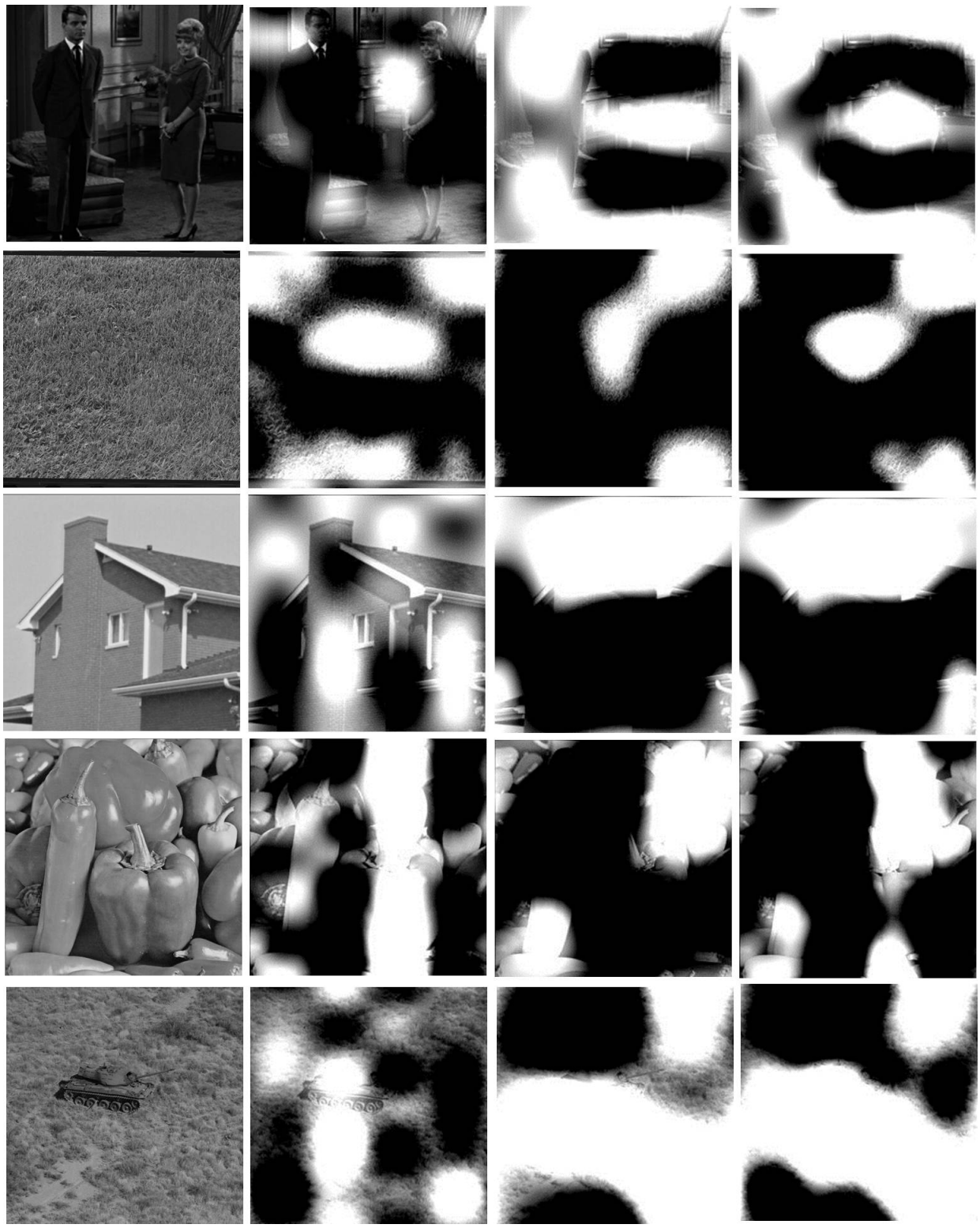

Fig. 5. Image reconstruction results with selective encryption. The images are selected from USC SIPI database. A selective encryption of 0.4% significant DWT coefficients (corresponding to 6th level decomposition) was performed using SWT and BCAC schemes. The first column shows sample images, the second column shows reconstruction results with SWT, third shows results with BCAC and fourth column shows results with SWT+BCAC

coefficients corresponding to 6^{th} level decomposition using DWT. Thus, the keylength for SWT is only 24 bits. The keylength for BCAC was chosen to 368 bits. (Each of the 16 DWT coefficients were quantized to 23 bits and then encoded, the choice of 23 bits made according to dynamic range of coefficients). The results for this simple experiment are presented in Figure 5 and in Table 4. It can be observed that selective encryption of 0.4% coefficients alone can lead to considerable degradation of perceived image quality. We performed SWT and BCAC over the 6th level wavelet coefficients alone (4 bands of 16x16 pixels). The key for SWT is 24 bits while that for BCAC encryption is 368 bits (16 pixels values each quantized into 23 bits each, 3 bits per pixel). SSIM or Structural Similarity metric measures the structural differences between the original image and the decrypted image. A value of 1 indicate strong similarity to original image while a value close to 0 indicate no-similarity between original and decrypted image. PSNR metric measures pixel-wise differences between two images (in decibels). A value higher than, say 40 indicate large similarity between original and decrypted image while a low value close to 0-10 dB or less indicates a huge differences or large noise between the two images. It can be seen in the results that BCAC leads to higher degradation in percieved image quality.

5.3 Sample Video Encryption Scheme

We use a sample video codec in Matlab which implements frequency transform (DCT), frame prediction (I and P frames in MPEG style) and we introduced entropy coding into the system (BCAC). Figure 6 presents results with video samples. A DCT based block-based video codec was used with a GOP size of 10. Each video has CIF resolution. The four videos: waterfall (least motion), calendar, highway and foreman (highest motion) were used from standard video databases[1]. It can be seen that encrypting only the motion vectors give degradation along those regions of video where motion is present. When we encrypt the residual also using CAC, we obtain high degrees of encryption.

6 SUMMARY

In this work we explored the potential of joint compression and encryption schemes for securing the multimedia content. We illustrated the potential for hardware savings and efficient encryption using this design with two examples - Secure Wavelet Transform and Chaotic Arithmetic Coding.

There is tremendous potential in this field for future research and deployment in state-of-the-art video codecs such as H.264/ SVC. There is possibility of developing such encryption schemes for motion compensation and estimation, and implementation on embedded device architectures.

1. http://trace.eas.asu.edu/yuv/

AUTHORS

Amit Pande is a Project Scientist in Department of Computer Science, University of California Davis. His research interests are in multimedia coding, communications, encryption, forensics; hardware acceleration and wireless networks. He is a recipient of NSF Computing Innovation Fellowship (2010-12), Best Paper Award at WPMC 2011, Design Contest Winner at VLSI Design Conference 2012 and 2009. He completed his PhD dissertation titled "Algorithms and Architectures for Secure Embedded Multimedia Systems" under Dr. Zambreno at Iowa State University in 2010. It was awarded with Research Excellence Award 2010 and Zaffarano Award 2010 (honorable mention). His undergraduate major project at IIT Roorkee titled "Ensuring Multimedia QoS for scarce resource networks" was awarded with Institute Silver Medal and Agilent Engineering and Technology Award 2010.

Joseph A. Zambreno has been with the Department of Electrical and Computer Engineering at Iowa State University since 2006, where he is currently an Associate Professor and co-director of the Reconfigurable Computing Lab (RCL). Prior to joining ISU he was at Northwestern University in Evanston, IL, where he graduated with his Ph.D. degree in Electrical and Computer Engineering in 2006, his M.S. degree in Electrical and Computer Engineering in 2002, and his B.S. degree summa cum laude in Computer Engineering in 2001. While at Northwestern University, Dr. Zambreno was a recipient of a National Science Foundation (NSF) Graduate Research Fellowship, a Northwestern University Graduate School Fellowship, a Walter P. Murphy Fellowship, and the EECS department Best Dissertation Award for his Ph.D. dissertation titled "Compiler and Architectural Approaches to Software Protection and Security." He is a recent recipient of the NSF CAREER award (2012), as well as

TABLE 4
Results for Selective encryption of 6th level DWT decomposition. Image reconstruction quality is evaluated for images from USC SIPI database using PSNR and SSIM metric

Image	SWT		BCAC		SWT+BCAC	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
Tank	0.35	4.26	0.19	−4.39	0.06	−5.33
Couple	0.147	6.819	0.07	−0.86	0.35	−1.35
Girl	0.088	2.88	0.01	−3.31	−0.02	−6.47
Grass	0.132	2.56	−0.161	−3.37	−0.132	−4.63
Peppers	0.19	0.24	0.12	8.35	0.08	5.15
House	0.16	−4.17	.035	−4.21	0.03	−5.5

the ISU award for Early Achievement in Teaching (2012) and the ECpE departments Warren B. Boast undergraduate teaching award (2009, 2011).

Dr. Prasant Mohapatra is currently the Tim Bucher Family Endowed Chair Professor and the Chairman of the Department of Computer Science at the University of California, Davis. He was/is on the editorial board of the IEEE Transactions on Computers, IEEE Transactions on Mobile Computing, IEEE Transaction on Parallel and Distributed Systems, ACM WINET, and Ad Hoc Networks. He has been on the program/organizational committees of several international conferences. He served as the Program Vice-Chair of INFOCOM 2004 and the Program Chair of SECON 2004, QShine 2006, WoWMoM 2009, WoWMoM 2012 and ICCCN 2012. He has been a Guest Editor for IEEE Network, IEEE Transactions on Mobile Computing, IEEE Communications, IEEE Wireless Communications, and the IEEE Computer. Dr. Mohapatra received his doctoral degree from Penn State University in 1993, and received an Outstanding Engineering Alumni Award in 2008. Dr. Mohapatra's research interests are in the areas of wireless networks, sensor networks, Internet protocols, and QoS. He is a Fellow of the IEEE.

REFERENCES

[1] O. Oyman, J. Foerster, Y. joo Tcha, and S.-C. Lee, "Toward enhanced mobile video services over wimax and lte [wimax/lte update]," *Communications Magazine, IEEE*, vol. 48, no. 8, pp. 68 –76, Aug. 2010.

[2] M. Baugher, D. McGrew, M. Naslund, E. Carrara, and K. Norrman, "The secure real-time transport protocol (SRTP)," United States, 2004.

[3] G. Gualdi, A. Prati, and R. Cucchiara, "Video streaming for mobile video surveillance," *Multimedia, IEEE Transactions on*, vol. 10, no. 6, pp. 1142 –1154, Oct. 2008.

[4] D. Hu, S. Mao, Y. Hou, and J. Reed, "Scalable video multicast in cognitive radio networks," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 3, pp. 334 –344, Apr. 2010.

[5] C. E. Shannon, "Communication theory of secrecy systems," *Bell Systems Technical Journal*, vol. 28, pp. 656–715, 1949.

[6] D. I. Cohen A. and F. J.C., "Biorthogonal Bases of Compactly Supported Wavelets," *Commun. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.

[7] D. Tay, "Rationalizing the coefficients of popular biorthogonal wavelet filters," *IEEE Trans. Circuits & Systems for Video Technology*, vol. 10, no. 6, pp. 998–1005, Sept. 2000.

[8] A. Pande and J. Zambreno, "Poly-DWT: Polymorphic wavelet hardware support for dynamic image compression," *ACM Transactions on Embedded Computing Systems*, 2011.

[9] D. Engel and A. Uhl, "Parameterized biorthogonal wavelet lifting for lightweight JPEG 2000 transparent encryption," in *Proc. ACM Work. on Multimedia and Security (MM&Sec) 2005*. ACM, 2005, pp. 63–70.

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

[11] A. Pande and J. Zambreno, "The secure wavelet transform," *Real-Time Image Processing, Springer*, to appear.

[12] R. Bose and S. Pathak, "A novel compression and encryption scheme using variable model arithmetic coding and coupled chaotic system," *IEEE Trans. Circuits and Systems I*, vol. 53, no. 4, pp. 848–857, April 2006.

[13] G. Jakimoski and K. Subbalakshmi, "Cryptanalysis of some multimedia encryption schemes," *IEEE Trans. Multimedia*, vol. 10, no. 3, pp. 330–338, April 2008.

[14] M. Grangetto, E. Magli, and G. Olmo, "Multimedia selective encryption by means of randomized arithmetic coding," *IEEE Trans. Multimedia*, vol. 8, no. 5, pp. 905–917, Oct. 2006.

[15] H. Kim, J. Wen, and J. Villasenor, "Secure arithmetic coding," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 2263–2272, May 2007.

[16] N. Nagaraj and P. G. Vaidya, "One-time pad, arithmetic coding and logic gates: An unifying theme using dynamical systems," *CoRR*, vol. abs/0803.0046, 2008.

[17] A. Pande, P. Mohapatra, and J. Zambreno, "Using chaotic maps for encrypting image and video content," in *IEEE International Symposium of Multimedia*, 2011, pp. 171 – 178.

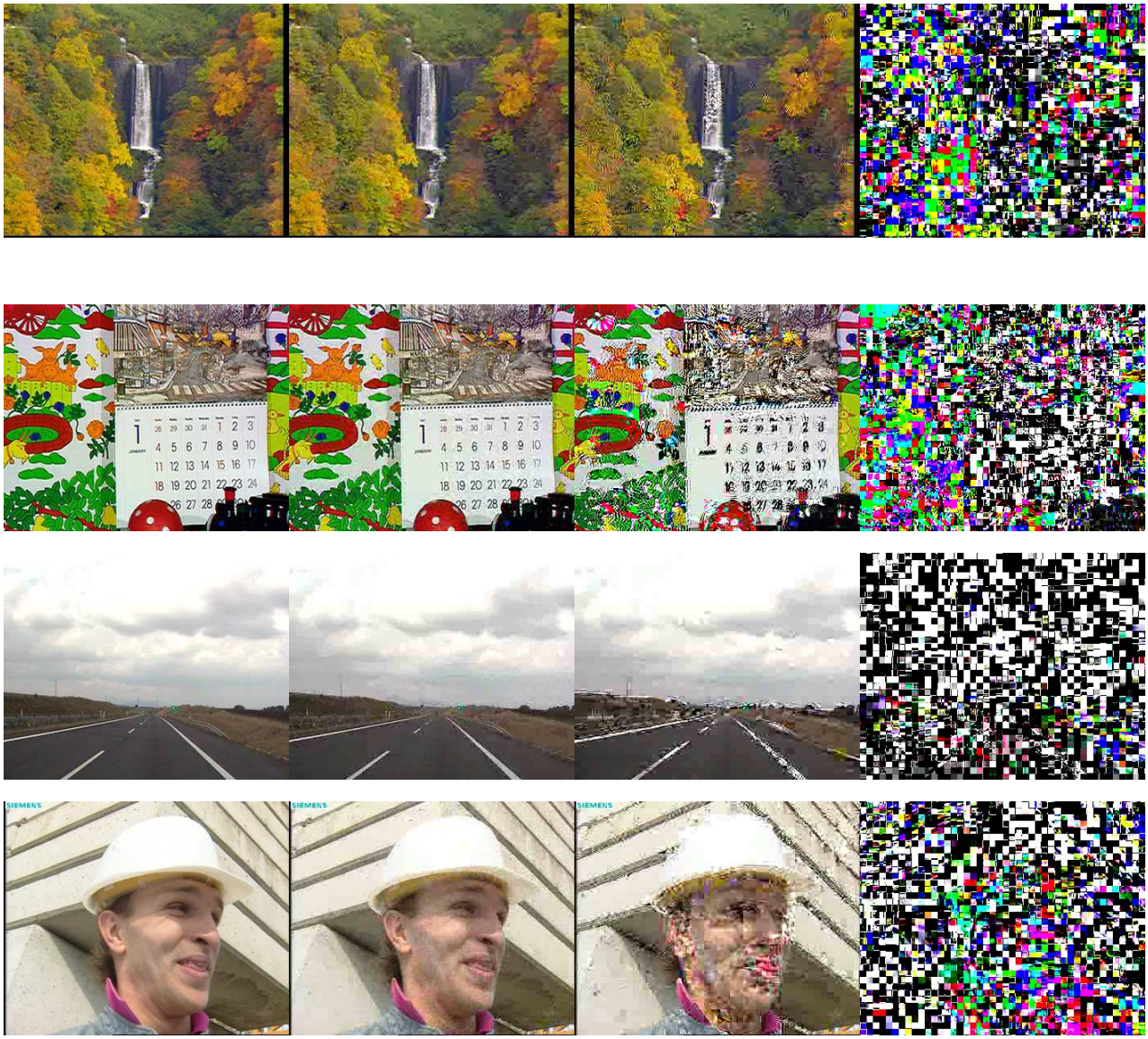

Fig. 6. Video reconstruction results with CAC scheme. (a) Original Video frame, (b) Compressed and reconstructed Video frame, (c) Only Motion Vectors (MV) Encrypted with BCAC and (d) MV + Residue encrypted with BCAC. Frame of each video sequence is shown in the figure.