Confidentiality and Privacy for Videos Storage and Transmission

Shaimaa ~. El-said[#1], Khalid F. ~. Hussein [*2], Mohamed M. Fouad [#3]
 # Electronics and Communication Department- Faculty OfEngineering- Zagazig University- Egypt
[1] Eng.sahmedAwindowslive.com
[3] fouadzuAhotmail.com
Microwaves Department- Electronics research institute- Egypt
[2] *khalid elgabaly@yahoo.com*

Abstract

Security and privacy issues of the transmitted data have become an important concern in multimedia technology. In this paper, we propose a computationally efficient and secure video encryption algorithm. This makes secure video encryption feasible for real-time applications without any e,tra dedicated hardware. We achieve computational efficiency by e,ploiting the frequently occurring patterns in the DCT coefficients of the video data. Computational comple,ity of the encryption is made proportional to the influence of the DCT coefficients on the visual content. On an average, our algorithm takes only 2 ms of encryption time per frame. This paper deals with the issues of color videos in-compression encryption, this done by using the optimized multiple Huffman table (OMHT) technique, will be elucidated upon in order to compress and encrypt both color images and videos. Only visual processing will be discussed (no audio processing). OMHT produces an encrypted coded sequence with an improved quality for the high compression ratios when compared to the e,isting standards in addition to the ability of encryption. The performance of the proposed technique is compared with that of the standard encryptions and it gives a very good perceptual quality especially at lower bit rates. We present e,perimental results which show that OMHT outperforms many traditional and selective techniques for video over wireless applications.

Keywords: *Videos compression, in-compression encryption technique, Optimized Multiple Huffman Tables technique, Computational cost analysis, Cryptanalysis, and Performance analysis.*

1. Introduction

The development of information technology and the rapid growth of computer networks allowed large files, such as digital images, to be easily transmitted in open networks such as the internet [1]. Data encryption is widely used to ensure security in open networks such as the internet. Each type of data has its own aspects, and different techniques should be used to protect confidential image data from unauthorized access [2]. Most of the available encryption algorithms are used for text data, however, due to large data size and real time constrains, algorithms that are good for textual data may not be suitable for multimedia data. 1n most of the natural images, the values of the neighboring pixels are strongly correlated. This means that the value of any given pixel can be reasonably predicted from the values of its neighbors [19]-[21]. Encryption is the process of transforming the information to ensure its

security [3]. With the huge growth of computer networks and the latest advances in digital technologies, a huge amount of digital data is being exchanged over various types of networks. 1t is often true that a large part of this information is either confidential or private demanding different security techniques to be used to provide the required protection [4]. Although data encryption is widely used to ensure security, most of the available encryption algorithms are used for text data. Due to large data size and real time constrains, algorithms that are good for textual data may not be suitable for multimedia data. Even though Triple-DES and 1DEA can achieve high security, it may not be suitable for multimedia applications and therefore encryption algorithms such as DES, AES, RSA and 1DEA were built for textual data. These algorithms are used perfectly to secure textual data. However, digital images are different from texts in many aspects and thus requiring different encryption algorithms [5][10].

There are number of applications for which the naive based encryption and decryption represents a major bottleneck in communication and processing. Some recent works explored a new way of securing the content, named, partial encryption or selective encryption, by applying encryption to a subset of a bitstream. The main goal of selective encryption is to reduce the amount of data to be encrypted while achieving a required level of security.

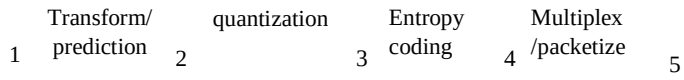


Figure 1. Candidate domains used to apply encryption to multimedia.

According to Fig. 1, there are two straight forward places to apply generic encryption to multimedia. The first possibility is to encrypt multimedia samples before any compression, stages 1 and 2, [11] [12] are examples of pre-compression selective encryption. The main problem with this approach is that the encryption often significantly changes the statistical characteristics of the original multimedia source, resulting in much reduced compressibility. The wavelet-based compression algorithm SP1HT [13] is an example of post-compression encryption scheme, stage 4 and 5. Wu et al proposed encryption scheme based on encoding with multiple Huffman tables (MHT) used alternately in a secret order [14]; is an example of in-compression selective encryption stage 4. The encryption with reasonably high level of security and unaffected compression can be achieved simultaneously in MHT technique, requiring almost negligible additional overhead. One of the major advantages by using this kind of joint encryption-compression approach is that encryption and compression can be achieved in one single step, which simplifies the system design and makes it flexible for some advanced multimedia processing such as scalability and rate shaping.

The MHT algorithms [14]-[17], aiming to increase the model space while maintaining the computational efficiency, keep the structure of the Huffman tree but enlarge the model space through tree mutation. MHT coding [14] makes use of standard coding tables. 1t is included in the final bit-stream for every image to be compressed. This approach has some disadvantages:

1. Visual degradation: very high-visual degradation can be achieved at low bitrate.

2. Cryptographic security: Gillman and Rivest [18] showed that decoding a Huffman coded bitstream without any knowledge about the Huffman coding tables would be very difficult.

However, the basic MHT is vulnerable to known and chosen plaintext attacks as pointed out in [19].

3. 1t writes all codes of the corresponding tables in the final bitstream even if only some of they were used to encode the associated events of the particular input image.

4. 1t is not adaptive technique. 1t does not make use of any statistic about the distribution of the events of the image.

To improve the security several kinds of enhanced MHT schemes have been proposed:

- By inserting random bit in the encrypted bit stream or integrating with a stream cipher [16].

- Recently another scheme via random rotation in partitioned bit streams has been reported [17].

Using known fixed tables in MHT technique generated by mutation (a method introduced in [14]) for compressing and encrypting images causes degradation in both compression ratio and security. We focus our research attention to enhancing multiple Huffman tables coding techniques.

To overcome the drawbacks of MHT technique, a new scheme for more general and efficient secure multimedia transmission, OMHT [22], is proposed. OMHT depends on using statistical-model-based compression method to generate different tables from a training set has the same data type as images or videos to be encrypted leading to increase compression efficiency and security of the used tables. Performance analysis of the newly proposed scheme OMHT shows that it can provide superior performance over both generic encryption and MHT in the security and compression.

This paper is organized as follows: Optimized Multiple Huffman tables coding technique (OMHT) is described in section 2. Section 3 presents a performance analysis and security analysis of the newly proposed OMHT technique. Conclusion is given in section 4.

2. Optimized Multiple Huffman Tables Technique

OMHT joint compression-encryption technique based on using adaptive lossy image compression (AL1C) technique [23], where the source image is converted into a vector then applied to the one dimensional discreet cosine transformer (1D-DCT) acts as source symbols reduction technique. From the output vector only the DCT's coefficients that carry a predetermined energy percent of the DCT output coefficients are selected, and other coefficients are neglected. Then the selected coefficients undergo the inverse discreet cosine transform (1DCT). The 1DCT output has a wider range of pixels' values, so quantization, is used to remove this excess information. The quantizer output is then processed by Huffman encoder which generates compressed output stream. 1n order to get the original image at the receiver, the previous stages are processed in the reverse direction. Experiments proved that AL1C compression techniques achieve better compression ratio with higher peak signal to noise ratio at low bit-rates than that of JPEG without the great visual degradation that appears in case of JPEG at low bit rate.

OMHT technique can process video sequence in two scenarios; in the first scenario (model 1) the video sequence is considered as group of consecutive frames each frame is compressed and encrypted separately as it be a stand alone image. As shown in Fig. 2 and Fig. 3 the OMHT encryption technique OMHT process takes two parallel paths A, and B, so it takes no additional time to add encryption to the compressed bitstream as both traditional and selective encryption techniques. 1t uses statistical-model-based compression to generate different tables from a training set of videos. This leads to increased compression efficiency and security. Multiple tables are generated for each one of videos types: A training dataset is constructed up as a number of selected frames of the same color palette as those to be encrypted. Each dataset is divided into N subsets each with equal number of frames. Random K frames are used to form single Huffman table. Concatenate all pixels values of all K images in single vector. Then, calculate the occurrence probability of each value, draw the Huffman tree, then generate Huffman table. Now we obtain D different (not fixed) Huffman tables for each dataset. The number of tables that can be obtained from N frames is:

$$D = C_{K}^{N}$$

2.1. Preparing of the Optimized Tables

Following is the procedure of preparing and using the optimized multiple Huffman tables and how it is used to both encode and encrypt videos. These procedures are common in model 1 and model 11 as shown in Fig. 2 and Fig. 3(path B).

Step 1: Frames training set are divided into datasets. Each dataset's frames have the same properties.

Step 2: Each dataset contains N image.

Step3: The input video's frames are compared to datasets to select the dataset that has the same properties.

Step4: Randomly choose M subsets each subset contains K frames from the dataset. Concatenate all frames of each subset and calculates the pixels probabilities. Then draw Huffman tree and find the Huffman table contains the different pixels' values and their associated variable codewords. Now we have M different tables to be used.

Step5: Number the generated M tables from 0 to M-1.

Step6: Generate a random vector P (the secret order) its length equal to the length of video frame under consideration. Each element value in P ranges from 0 to M-1.

Step7: For the ith encountered symbol (coefficient to be encoded), use table Pi (mod n) to encode it.

Step8: Tables are saved at each decoder and the order (P) by which the tables are generated and used is kept secret.

2.2. Video Compression using OMHT

OMHT model 1 processes the video sequence frames by frame and each frame is considered as a stand alone image as shown in Fig. 2(path A). Following is the procedure of compressing a video sequence using model 1:

1. Converting video sequence of any type into RGB frames.
2. Each nXm frame is converted into a single vector by cascading consecutive rows to form a single vector.

3. This vector is exposed to a discrete cosine transformer to transform the spatial domain of frame's pixels into their frequency domain in which the frame energy concentrated in small number of coefficients.
4. The output of the DCT process is a vector that have the same length of the image (number of pixels in the frame), but with many values approximated to zeros. A selector is used to select the number of transmitted coefficients (Tc) that contains a specific percentage of the frame energy where as the other coefficients are canceled.
5. The selected coefficients are returned back into spatial domain using 1nverse Discrete Cosine transformer.
6. The 1DCT output is quantized using Range Mapping technique (RM) as discussed later in this section. This quantization technique is reversible; this means that the dequantized values can be turned back to their original values leading to no quantization losses.

OMHT model 11 processes the video sequence frames by transmitting the first frame of the video without any processing except encrypting it using optimized tables encoding then calculating the motion vectors between the consecutive frames, compress them, and encrypting them using optimized tables encoding. As shown in Fig. 3(path A). Following is the procedure of compressing a video sequence using model 11:

1. Converting video sequence of any type into RGB frames.
2. Each nXm frame is converted into a single vector by cascading consecutive rows to form a single vector.
3. Calculate the difference vector between each two consecutive frames (motion vectors) by subtracting each frame from its previous one as shown in Fig. 4.
4. The first frame FO is only encrypted by encoding it using optimized tables.
5. Each motion vector is exposed to a discrete cosine transformer to transform its components from the spatial domain into their frequency domain in which their energy concentrated in small number of coefficients.
6. The output of the DCT process is a vector that have the same length of the difference vector (number of pixels in the MV), but with many values approximated to zeros. A selector is used to select the number of transmitted coefficients (Tc) that contains a specific percentage of the MV energy where as the other coefficients are canceled.
7. The selected coefficients are returned back into spatial domain using 1nverse Discrete Cosine transformer.
8. The 1DCT output is quantized using Range Mapping technique (RM).

1n Range Mapping (RM) quantization technique [23], the range of the pixel values at the 1DCT output is mapped to another range that can be encoded by a lower number of bits. Hence, each pixel value at the 1DCT output is mapped to another value in the new range. As the new (mapped) values can be represented by a smaller number of bits, symbol reduction is achieved. The RM quantization procedure can be described as follows:

- Calculate the minimum and maximum of the pixel values at the 1DCT output, (Umin , Umax).
- Determine another range (Vmin, Vmax) that can be encoded by a smaller number of bits.

- Map each pixel value (u) of the image at the 1DCT output to its corresponding value (v) in the new rang through the following equation:

$$\sim = \frac{(\min)(\max \min)}{\min \mathbf{U}_{\max} \mathbf{U}_i} \mathbf{2} \sim \sim$$

- The mapped pixel values (v) together with range limits Umin , Umax , Vmin and Vmax are transmitted to the receiver (or saved in the storage unit).
- At the receiver, the original pixel values (u) are restored from the mapped values (v) according to the following equation:

$$\sim = \frac{(\mathbf{v} \mathbf{V}_{\min})(\mathbf{u}_{\max} \mathbf{U}_{\min})}{\mathbf{V}_{\max} \mathbf{V}_{\min}} + \mathbf{u}_{\min}$$

The basic principle for OMHT (model 11) for video compression and encryption as shown in Fig.5 is the image-to-image prediction. The first frame FO is called an 1-frame and is self-contained, having no dependency outside of that image. The following frames may use part of the first image as a reference. An image that is predicted from one reference image is called a P frame. The appropriate GOP (GOP = 4, e.g. 1PPP 1PPP ...) depends on the application. By decreasing the frequency of 1-frames, the bit rate can be reduced. By removing the B-frames, latency can be reduced.

Another important aspect of OMHT is the bit rate mode that is used. 1n OMHT systems, it is possible to select the mode, CBR (Constant Bit Rate) or VBR (Variable Bit Rate), to be used. The optimal selection depends on the application and available network infrastructure.

With limited bandwidth available, the preferred mode is normally CBR as this mode generates a constant and predefined bit rate. The disadvantage with CBR is that image quality will vary. While the quality will remain relatively high when there is no motion in a scene, it will significantly decrease with increased motion.

With VBR, a predefined level of image quality can be maintained regardless of motion or the lack of it in a scene. This is often desirable in video surveillance applications where there is a need for high quality, particularly if there is motion in a scene. Since the bit rate in VBR may vary--even when an average target bit rate is defined--the network infrastructure (available bandwidth) for such a system needs to have a higher capacity.

The foremost advantage of OMHT model 1 is that it is more error resilience than model 11 since it is based on intra-frame coding. As all frames are coded independently from one another, transmission errors in one frame do not propagate to subsequent frames. While OMHT model 11 is based on inter-frame coding. The correct decoding of a predictive frame (P or B frames) depends on the truthful decoding of the reference frames (e.g. previous 1 or P frames). As a direct consequence, transmission errors can affect several frames.

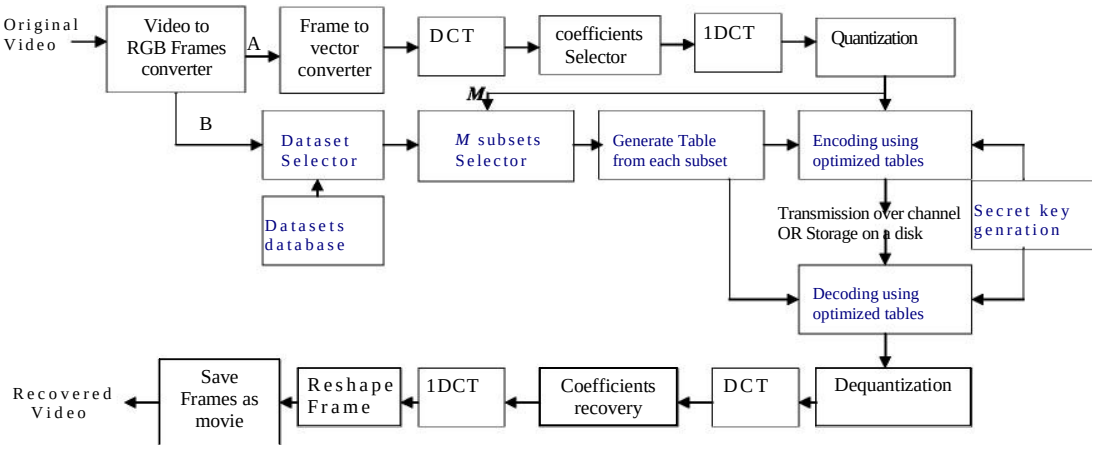


Figure 2. Block Diagram of Optimized Multiple Huffman Tables technique (model I)

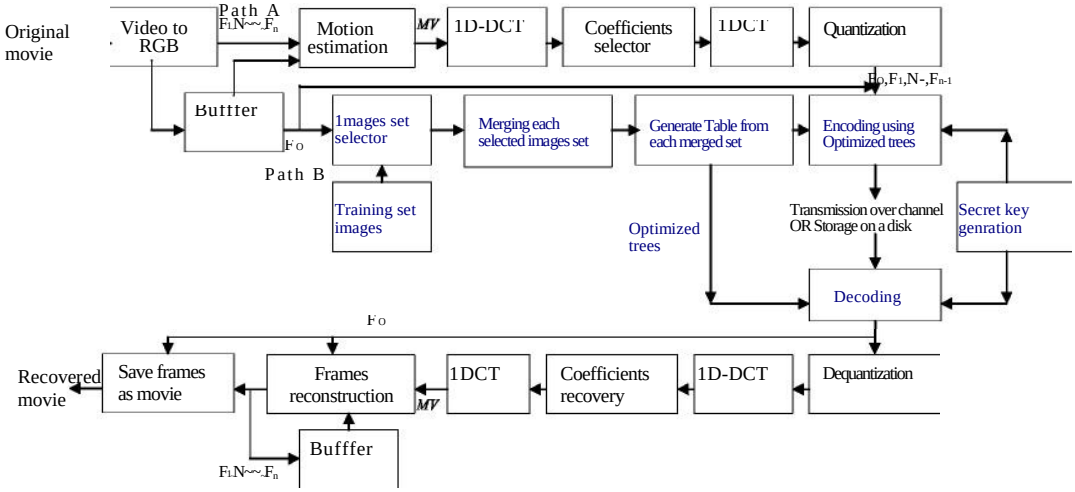


Figure 3. Block Diagram of Optimized Multiple Huffman Tables technique (model II)

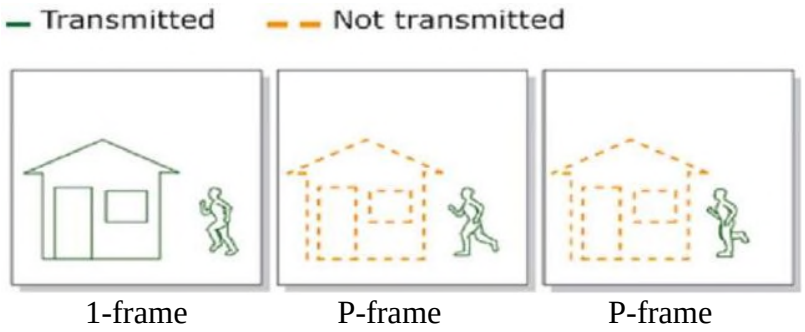


Figure 4. Three Consecutive Frames of Video Sequence

3. E,perimental Results

1n this section the performance of the proposed technique is analyzed from the viewpoints of compression and encryption. OMHT coding technique used to both encode and encrypt video sequence is tested under two scenarios; the first scenario is encoding video sequence frame by frame, and the other is encrypt the first frame and then compress and encrypt the motion vectors between successive frames (the difference between two consecutive frames). The training set was obtained from video sequences in C1F and QC1F format.

Security analysis of the encryption algorithm is commonly needed for evaluating and comparing the performance of encryption algorithms. Performance analysis based on cryptanalysis can prove the complexity for the attacker deciphering the encryption algorithm in theory, but cannot provide the visual security degree of the cipher-images. 1n order to develop an objective assessment algorithm on visual security degree of visual media, current security assessments methods of image and video encryption were deep studied and divided into three kinds: assessment based on cryptographic analysis, assessment based on subjective evaluation, and assessment based on video quality assessment.

3.1. Assessment Based On Video Quality

The target of this assessment is to prove that using encryption with the compression in a single step does not affect the video's visual quality. Video quality can be evaluated by measuring compression ratio, peak signal to noise ratio, mean square error, and signal to noise ratio of compressed images. The quantitative comparison is made using Peak Signal to Noise Ratio (PSNR) and Structural Similarity 1ndex (SS1M). These quality metrics are calculated for each frame and finally the average of them is taken. Peak Signal to Noise Ratio (PSNR) is closely related to MSE and is given in:

$$PSNR_{dB} = 20 \log_{10} \frac{(2^n - 1)}{\sqrt{\text{MSE}}}$$

Here n is the number of bits per image sample. PSNR is a measure of the similarity of an image that is computed by measuring the pixel difference between the original image and the compressed image. A PSNR relates to the mathematical similarity of two images. Values for PSNR range between infinity for identical images to 0 for images that have no commonality. The PSNR measure suffers from a number of limitations. For a given image or image sequence, high PSNR usually indicates high quality and low PSNR usually indicates low quality. However, a particular value of PSNR does not necessarily equate to an 'absolute' subjective quality.

PSNR ratings do not necessarily correlate with 'true' subjective quality. The SS1M index is a different approach for video quality assessment. This method differs from the other methods, which all are error based, by using the structural distortion measurement instead of the error. The idea behind this is that the human vision system is highly specialized in extracting structural information from the viewing field and it is not specialized in extracting the errors. Thus, a measurement on structural distortion should give a better correlation to the subjective impression. Fig. 5 shows the block diagram of the structural measurement system[27] where it studies the similarity of luminance, contrast, and structure of the reconstructed video with respect to the original one. Finally, the three components are combined to yield an Overall similarity measure:

Where x, and y are the original and reconstructed videos respectively. L, C, and S are luminance, contrast, and structure similarities respectively.

1n this experiment, we have taken different kinds of video sequences to be tested, Suzie (144x176x3), Miss$_P$America (144x176x3), Akiyo (144x176x3) , Foreman (288x352x3), Mobile (144x176x3), tennis (144x176x3), Football (288x352 x3), and Flower (288x352x3) each with 100 frames. The first four sequences contain the low frequency components where the change in gray levels in each frame is small. The other four sequences contain the high frequency components where the change in gray levels in each frame is large.

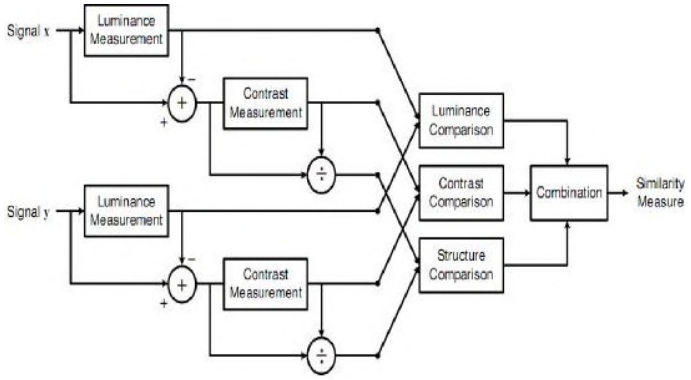


Figure 5. Diagram of the structural similarity (SSIM) measurement system [27].

1n case of using OMHT technique without AL1C compression technique; only encoding the frames with the optimized tables, the size of the output bitsream using model 1 is the same as that of the standard Huffman encoder and less than that of MHT technique, while using model 11 gives much lower size than that obtained by using standard Huffman or MHT technique, as shown in Table 1. Fig. 6 shows the increasing in the compressed file size due to using some kinds of selective encryption techniques and from using MHT, and OMHT techniques. 1t is obvious that OMHT does not affect the size of the compressed bit stream.

Suzie and Miss$_P$America video sequences, are of the most common test video sequences in video compression research, consists primarily of low frequency content. As human eye is less sensitive to smooth variation more compression can be achieved using OMHT modeling without loss of quality as shown in Fig. 7 and Fig. 8.

Table 2 shows the compression ratios, PSNR and SS1M values for MPEG-4 compression technique and OMHT model 1 and model 11 at absolute thresholds, to verify that the proposed encryption technique does not affect the compression quality. From the table we can observe that, using OMHT models we can achieve security without affecting compression ratio or the frame quality especially at low bitrate. Here, the Compression Ratio (CR) can be calculated as the ratio of the input file size to the output file size.

Table 1. Compression performance comparison between the standard Huffman coding, MHT, and OMHT (model I, and model II)

File	Huffman	MHT		OMHT(t)		OMHT(~~)	
	size	oip size	increase%	olp size	increase%	oip size	increase %
Cats	242811	23811	-1.94	22933	-5.877	192675	-20.65
Chart	176158	18394	4.42	17645	0.16888	124533	-29.3
Gold	19640	21147	7.67	20065	2.7648	15223	-22.49
Hotel	19949	21727	8.91	21132	5.5981	12935	-35.16
Water	118464	10398	-12.22	11692	-1.318	96635	-18.43
Tools	86864	92396	6.37	89533	2.981	67154	-22.69
woman	233316	23411	0.34	23332	0.00385	185943	-20.3
Seismic	11544	10942	-5.22	11162	-3.4223	8639	-25.17
Bike	248619	26206	5.41	24819	-0.170	185431	-25.42

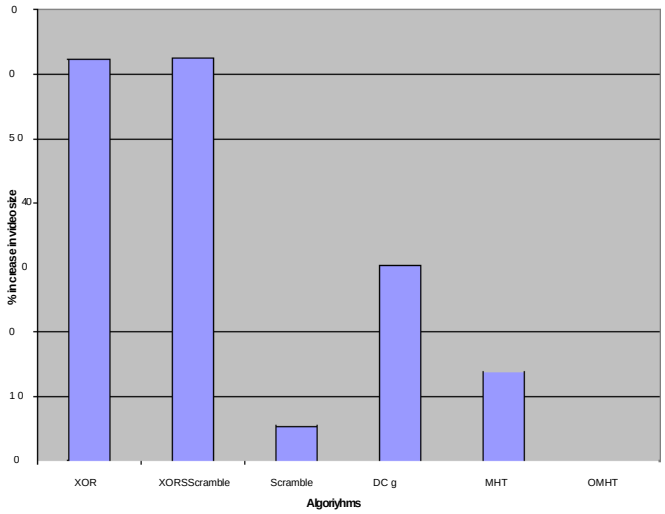


Figure 6. Encryption Overhead on the Compression

Table 2. Comparison between compression performance of standards and that of the proposed models at low bitrate

V1DEO	Technique	bpp	PSNR(dB)	SS1M
SUZ1	MPEG-4	0. 21	39.7	0.927
	MJ2K	0. 25	39.1	0.92
	OMHT (*model I*)	0. 23	39.15	0.9
	OMHT (*model*	0.2	39.47	0.92
M1SS_AMER1CA	MPEG-4	0.2	33.8	0.93
	MJ2K	0.28	34	0.946
	OMHT(*model I*)	0.21	34.6	0.938
	OMHT (*model*	0.24	34.1	0.926
FOREMAN	MPEG-4	0.13	35.8	0.94
	MJ2K	0.129	35	0.939
	OMHT (*model I*)	0.12	35.6	0.921
	OMHT (*model*	0.1	36	0.93
AK1YO	MPEG-4	0.1	34.8	0.929
	MJ2K	0.11	34	0.93
	OMHT (*model I*)	0.123	34.6	0.924
	OMHT (*model*	0.11	35.1	0.94
FOOTBALL	MPEG-4	0.12	31.3	0.94
	MJ2K	0.132	32.9	0.939
	OMHT (*model I*)	0.138	33	0.921
	OMHT (*model*	0.1	32.5	0.93
TENN1S	MPEG-4	0.1	32.9	0.929
	MJ2K	0.11	33.6	0.93
	OMHT (*model I*)	0.123	33.5	0.924
	OMHT (*model*	0.11	32.6	0.92

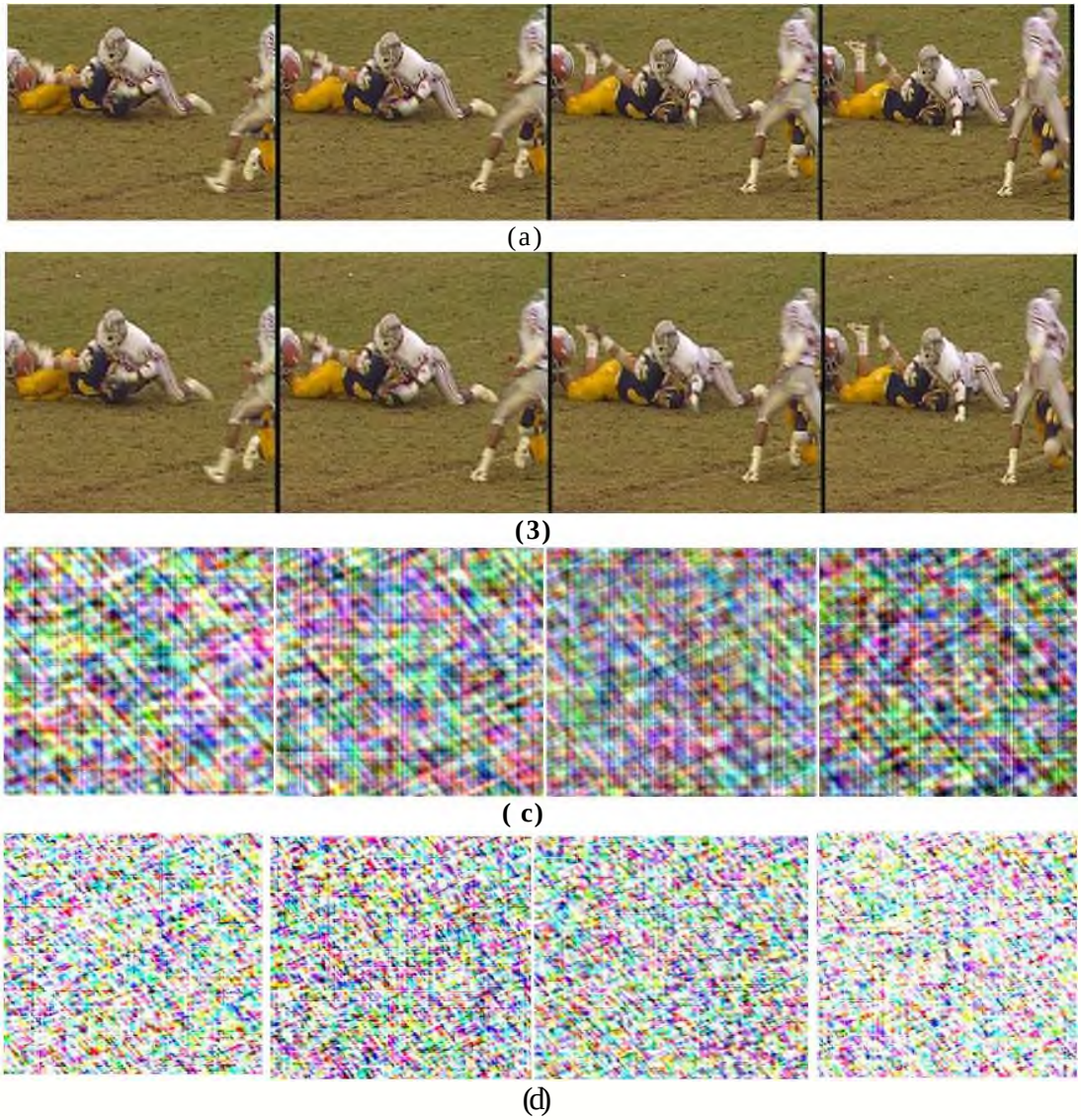
The improvement in picture quality can be observed when we fix the compression ratio to a particular value and compare the quality for various methods of compression. Table 2 also shows the compression performance of OMHT model 1, and model 11 on several videos. 1t can be observed that video coder with OMHT model 1 give better perceptual quality than that of model 11, while model 11 gives higher compression ratios and better security than model 1.

Figure 7. (a) First frame of Suzie video sequence (b) the reconstructed frame using MPEG-4 (C) The reconstructed frame using OMHT (model I) (d) The reconstructed frame using OMHT (model II)

Figure 8. First frame of Miss_America video sequence the reconstructed frame using MPEG-4, OMHT model I , and OMHT model II respectively

Fig. 9 shows the compression and encryption performance of OMHT model 1. Fig. 9(a) is the first four original frames of the Football. C1F sequence, Fig. 9(b) is the reconstructed frames decoded and decrypted using the correct secret key, the calculated PSNR=33dB. Fig. 9(C) shows the frames decoded using OMHT with wrong secret key it gives PSNR=6.356 dB, and Fig. 9(d) Frames decoded using JPEG their PSNR=4.3 dB.

Fig. 10 shows the compression and encryption performance of OMHT model 11. Fig. 10(a) is the first four original frames of the Football. C1F sequence, Fig. 10(b) is the reconstructed frames decoded and decrypted using the correct secret key, the calculated PSNR=32.5dB. Fig. 10(C) shows the frames decoded using OMHT with wrong secret key it gives PSNR=5.5 dB, and Fig. 10(d) Frames decoded using JPEG their PSNR=3.8 dB.



(a)

(3)

(c)

(d)

Figure 9. (a) and (b) are the original and reconstructed frames of "Football. CIF" sequence using OMHT *model I* technique respectively. (C) Frames decoded using OMHT with wrong secret key (PSNR=6.356 dB). (d) Frames decoded using JPEG (PSNR=4.3 dB)

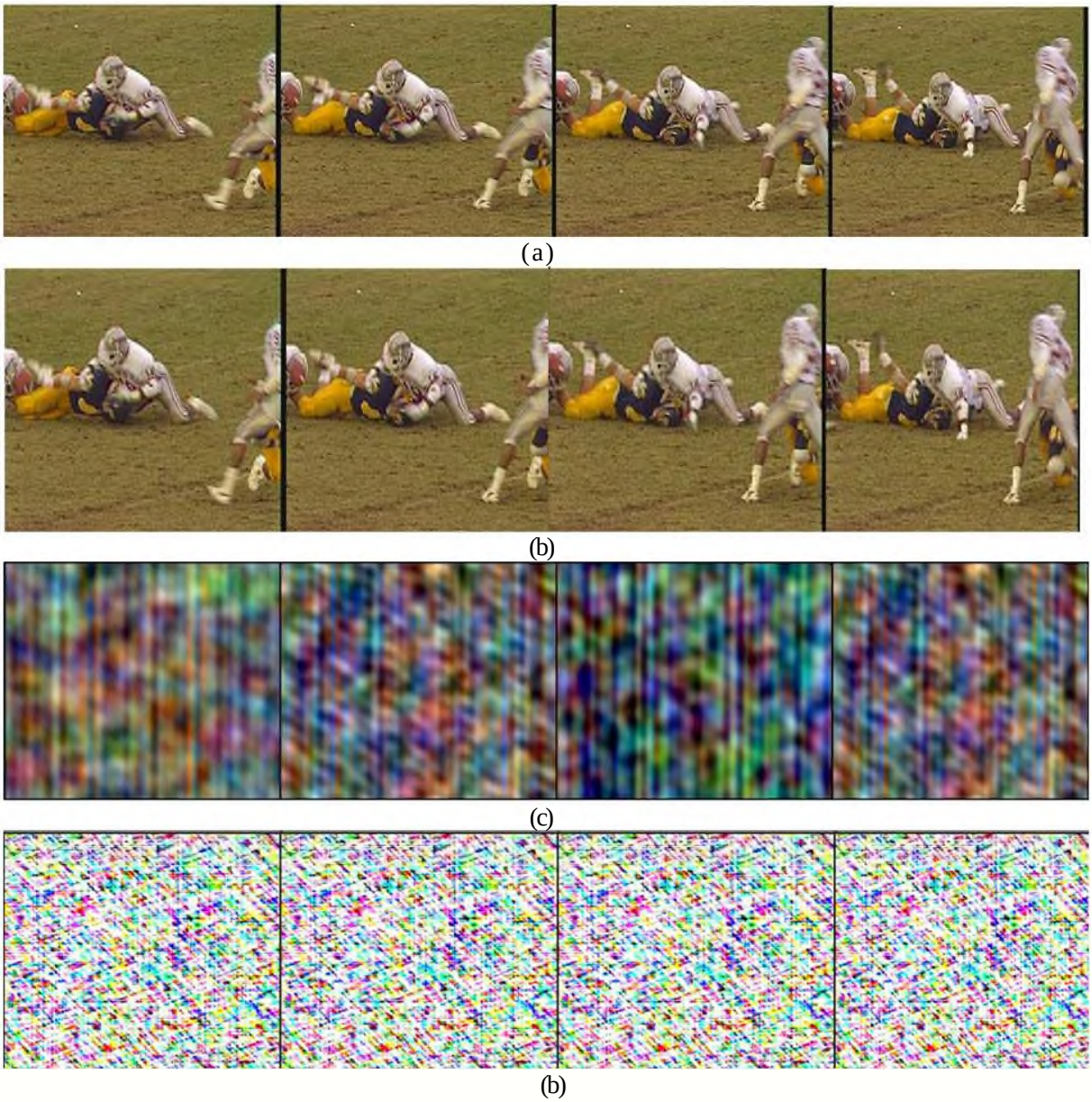
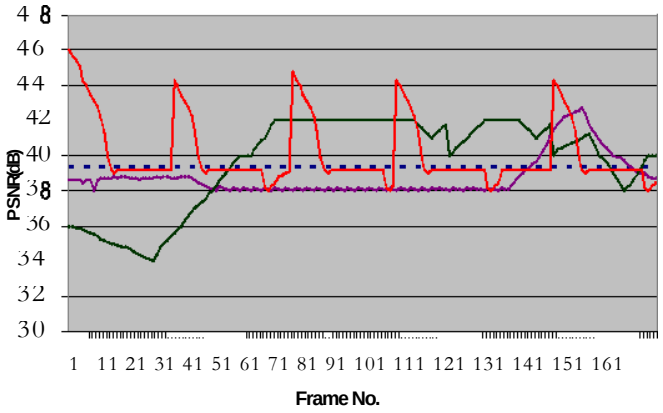


Figure 10. (a) and (b) are the original and reconstructed frames of "Football. CIF" sequence using OMHT model II technique respectively. (C) Frames decoded using OMHT with wrong secret key (PSNR=5.5 dB). (d) Frames decoded using JPEG (PSNR=3.8 dB)

Fig. 11 shows a PSNR comparison between MJ2K, MP4, and OMHT for Suzie video, without transmission errors. OMHT model 11 is assumed to has a GOP that consists of 1 frame followed by 30 P frames. From the figure it is obvious that OMHT model 1 has nearly the same PSNR for all frames in the video sequence. OMHT model 11 has very high PSNR for the 1 frame then it reduces gradually through the GOP. Both MPEG 4 and motion JPEG 2000(MJ2K) have PSNR varies with the frame number.



OMHT (/) OMHT(II) MJ2K MPEG-4

Figure 11. PSNR comparison between MJ2K, MP4, and OMHT for Suzie video, without transmission errors.

1n Table 3 the main Characteristics Comparison between MJ2K, MPEG-4 compression techniques, and OMHT in-compression encryption technique is done to prove that using the OMHT technique with its two models has nearly the same characteristics of the standard compression techniques in addition to providing high security level for video transmission or storage.

Table 3. Main Characteristics Comparison between MJ2K, MPEG-4 compression techniques, and OMHT in-compression encryption technique

	MJ2K	MPEG-4	OMHT-1	OMHT-11
Encryption	no	no	yes	yes
1ntra-coding	wavelet	2D-DCT	1D-DCT	1D-DCT
1nter-coding	no	yes	no	yes
Coding efficiency	low	high	medium	high
Main artifacts	Blur, ranging	Blocking, ranging	Blur, ranging	Blur, ranging
Error resilience	high	low	high	low
Complexity	low	high	medium	high
Coding delay	low	high	medium	high

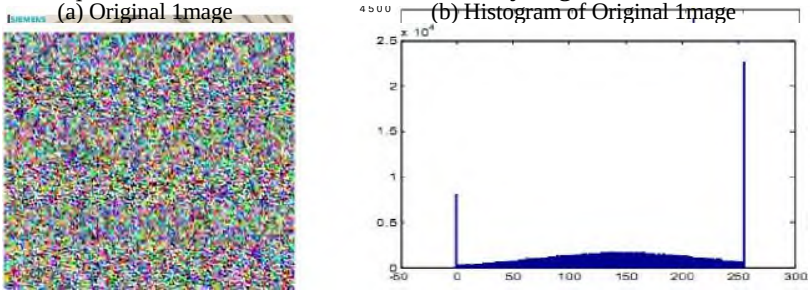
3.2. Assessment Based On Subjective Evaluation

This assessment measures the unrecognizable degree of cipher videos. 1t measures the encryption strength performance of the proposed OMHT technique, Foreman frame in Fig. 12(a) and its histogram in Fig. 12(b). This frame is compressed and encrypted using the OMHT uses multiple Huffman tables, generated from a large set of training video frames that have the same type of the test video, not predetermined fixed tables, in a secret order. Fig. 12(c), and Fig. 12(d) show the test frame and its histogram after decoding it with another technique as JPEG. Fig. 12(e), and Fig. 12(f) show the test frame and its histogram after decoding it with OMHT technique and the same encoding tables but without knowing the secret order (secret key).

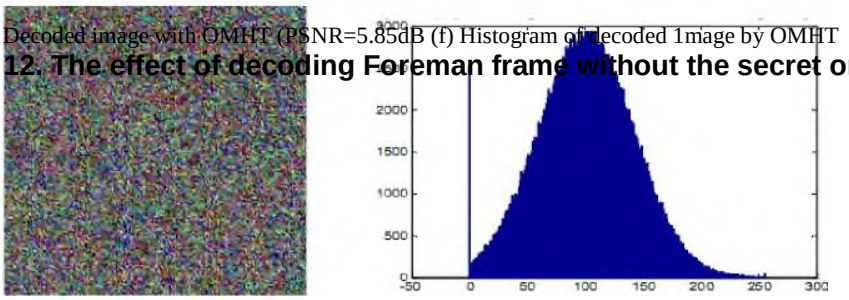
Fig. 13 shows the effect of several encryption techniques on perceptual quality of the decoded first frame of Akiyo sequence Fig.13(a) shows the original frame, Fig.13(b) shows the reconstructed frame encrypted by selective encryption that encrypts only one MSB and parity bits the PSNR of this technique equals 18.8dB, Fig.13(c) shows the reconstructed

frame encrypted by selective encryption that encrypts the eight MSB and parity bits are encrypted it gives PSNR=9.9dB, Fig.13 (d) shows the reconstructed frame encrypted by OMHT model 1 it gives PSNR=7.4dB, and Fig.13(e) Reconstructed frame encrypted by OMHT model 11 gives PSNR= 5dB. By comparing the PSNR obtained above, we can conclude that OMHT technique can degrade the perceptual quality of cipher frames compared to the other techniques, and hence it increases the security degree.

(a) Original 1mage (b) Histogram of Original 1mage



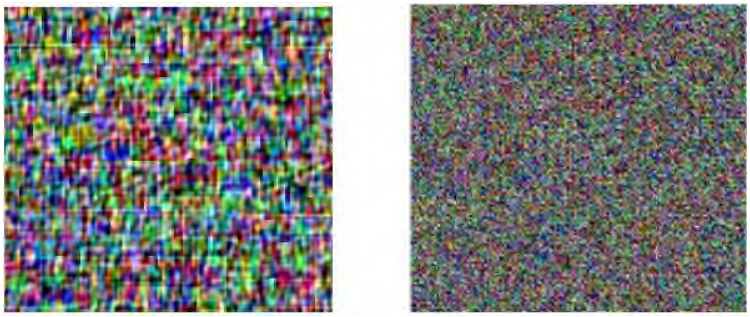
(c) Decoded image with JPEG (PSNR=5.2 dB) (d) Histogram of decoded 1mage by JPEG



(e) Decoded image with OMHT (PSNR=5.85dB (f) Histogram of decoded 1mage by OMHT

Figure 12. The effect of decoding Foreman frame without the secret order

(a) (b) (c)



(d) (e)

Figure 13. The effect of several encryption techniques on visual appearance of the decoded Akiyo frame: (a) Original frame. (b) Reconstructed frame encrypted by selective encryption (1MSB and parity bits are encrypted PSNR=18.8dB). (c) Reconstructed frame encrypted by selective encryption (8MSB and parity bits are encrypted PSNR=9.9dB). (d) Reconstructed frame encrypted by OMHT model I (PSNR=7.4dB). (e) Reconstructed frame encrypted by OMHT model II (PSNR= 5dB).

3.3. Assessment based on cryptographic analysis

Cryptanalyst analyzes the possibility of deciphering the cipher images through the use of cryptanalysis attack. To evaluate the security of the proposed scheme it is important to consider how cryptanalysts would attack the cryptosystem. Attack models are usually classified into three categories. 1n each model, every detail of the encryption/decryption algorithm is assumed to be known to cryptanalysts. The security depends solely on the hidden key.

- Ciphertext-only attack
  The cryptanalyst has only the ciphertext to work with; brut-force exhaustive key search attack. The strength of resisting this attack relies on a larger key space.
  As described previously the available number of tables (D) that can be generated from N images by using random K images for generating a signal table is calculated as:

$$D = C_K^N$$

  Since the length of the random vector by which the tables are used is equal the length of coefficients vector to be encrypted equals n, and there are M random tables chosen each time to be used, there would be $(M)^n$ different orders to use the Huffman tables. As a result the size of the key space is:

$$size(keyspace) = C_M^D x (M)^N$$

  Let each dataset contains one thousand images of the same type. 1f random ten images are selected to form single Huffman table, we can obtain about 3762*1020 different tables. 1f there is 7000 coefficients (min. number of coefficients after compressing the image values), we need to generate a random vector that contains 7000 elements each varying from 0 to M. The resulting size of key space would at least be:

$$size(keyspace) = C_{7000}^{3762x10^{20}} x (7000)^{7000}$$

  So, it is practically impossible to perform brute-force search in a key-space of this size.

- Known-plaintext attack
  1n this attack model, the cryptanalyst has a string of symbols and its corresponding encrypted bitstream. The goal is to obtain which tables are used and in which order. The first step is to guess which bits in the bitstream correspond to which symbol in the original data. Since symbols are encoded with different Huffman tables would produce code words with slightly different lengths, synchronization between plaintext and ciphertext is extremely difficult.

- Chosen-plaintext attack
  Cryptanalyst could chose the plaintext and obtain the corresponding ciphertext. His goal is to obtain the secret key to decrypt the ciphertext in the future. 1n most encryption schemes the secret key remains unchanged for along time. These techniques are vulnerable to this kind of attack. The attacker inserts one single symbol into the cipher and observes the corresponding output codeword. 1n this case the attacker would have no synchronization problem at all. 1n OMHT, the secret key is the order by which we choose the Huffman tables and using it. OMHT technique is designed to receive symbols as a whole chunk and output the corresponding codewords all together. To increase the security the chunk size should be as big as possible.

3.4. Assessment based on computational cost analysis

The evaluation of the computational speed of ciphers usually consists of the analysis of the key-setup cost, the encryption cost and the decryption cost [25]. The encryption and the decryption costs are usually similar, and they are more important than the key-setup cost because one single key-setup can often be followed by thousands of encryption/decryption operations. 1n the following, we analyze these costs of our OMHT encryption scheme, and compare them with those of MHT and modern ciphers.

a) Key-Setup cost: For a single frame; the key-setup process includes all the computation and memory allocation operations prior to actual encryption of the first bit in the plaintext. The computational cost of OMHT key-setup is dominated by the construction of optimized multiple Huffman tables, generation of the secret order by which those tables are used, and comparing the test video with datasets. OMHT takes about 10 operation per table generation, single operation for secret key generation, and L operation for comparison. The total number of operations equal 10XMXL+1+L, where L, M is number of datasets and number of subsets respectively. For L=4, M=20, the net Key-Setup cost =805 operations. For MHT technique it takes 20 operations per table entry, the total cost would be 20xtxm, where t and m are the table size and the number of selected tables, respectively. For the example of JPEG dc coefficient encryption as shown in the previous subsection, the key-setup cost would be around 2000 operations (t=13 and m=8).Compared with the ciphers listed in Table 4, the key-setup cost of OMHT encryption is much smaller than MHT and other ciphers.

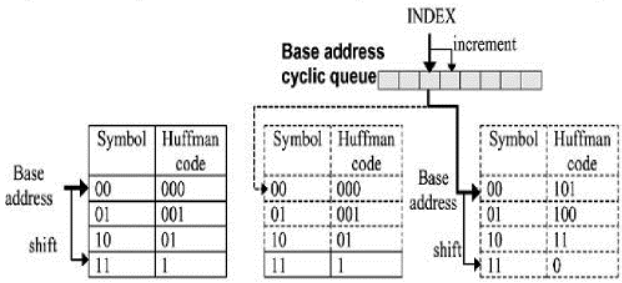
b) Encryption/Decryption cost: The net computational cost of the OMHT is the same as the basic MHT-encryption scheme [14] is less than one CPU operation per encrypted bit as explained below. When a symbol is to be encoded with a normal Huffman coder, the shift amount is added to the base address of the table to obtain the address of the desired Huffman code. This process is illustrated in Fig.14 (a). 1n the basic MHT system, we store the base addresses of the tables in a cyclic queue according to the order that they are used. When a symbol is to be encoded/encrypted, the base address is first loaded from the

memory, and then the shift-amount is added to it. Afterwards, the index to the cyclic queue of base addresses should be increased by one. Then, the index should be compared with the end of the queue in order to decide whether it should be reset to the beginning of the queue. Therefore, the computational difference between OMHT cipher/encoder and a normal Huffman coder is one memory-load, one addition and one comparison operation for each symbol encoded. The encoding process of the proposed cipher/encoder is shown in Fig.14 (b). Since each symbol in the original data usually corresponds to more than 3 bits in the Huffman bitstream, then encryption cost of our algorithm is less than one CPU operation per encrypted bit, which is around 20 times smaller than the well-known AES as listed in Table 4.

Recently, a new cryptographic cipher named COS [26] with a very fast speed is gaining popularity. 1t is around 4-5 times faster than AES. Compared to COS, the encryption cost of OMHT is still several times smaller.

Table 4. Computational Costs of AES Finalists on a Pentium-MMX Machine. The Figures in This Table are Translated from [27] by Assuming Two CPU Instructions are Executed in Every Clock Cycle in a Pentium-MMX CPU

Cipher Type	Key-setup Cost (CPU instructions)	Encryption Cost (CPU instructions/bit)
MARS	9416	25
RC6	10372	22
Rijndael	35484	20
Serpent	26308	28
Twofish	37692	20



(a) (b)

Figure 14. (a) Normal Huffman Coder Adds the Shift Amount to the Base address of the Table to Obtain the Address of the Desired Huffman Code. (b) OMHT Loads the Base Addresses of Huffman Tables from a Cyclic Queue, and the Index to the Queue is Increased by One After Coding of Each Symbol.

4. Conclusion

From the viewpoint of compression, the experiments' results reveal that the proposed OMHT technique achieves better compression and security performance than both MHT, and JPEG 1mage Compression Standard. The experiments on the proposed technique and JPEG also reveal that the resultant CR of lossy OMHT technique can be increased over a wide range for the same image according to the number of quantization levels chosen with a slightly decrease of its PSNR without a great noticeable visual degradation at low bitrate, where as increasing compression ratio of both MHT, and JPEG techniques is combined by the

appearance of the blockings effect (visual degradation) in the reconstructed image at low bitrate, this enables the usage of lossy OMHT technique not only in applications require high security but also in many applications need a compression ratio variability with nearly stable PSNR. Whereas, the proposed joint compression-encryption (OMHT) technique employ a Source Symbols Reduction using one dimensional DCT, quantization, and Huffman Coding enhance the performance of the compression. Further, the proposed new compression-encryption technique could be applied on any source data, not only images, which uses Huffman coding. OMHT is a general technique; it is suitable for compression and encryption of text, image, and video files.

Since images have different statistics, using the same fixed JPEG standard predefined coding tables as suggested in MHT technique will not be effective in encoding all image and video types. The OMHT method obtains better performance in terms of storage space use and more stable peak signal to noise ratio than that of JPEG at low bitrates

From the viewpoint of security, the experiments' results reveal that the proposed OMHT technique achieves better security than MHT technique since it is more resistible to ciphertext-only attack, known-plaintext attack, and even chosen-plaintext attack. OMHT satisfies the conditions of the good encryptor:

- High security: resistance against various types of attacks, including the ciphertext-only attack and the known/chosen plaintext attack.
- Low encryption cost: the encryption cost not exceed very small portion of the total computation cost of compression
- No harm to the compression ratio: The increase of the final bit stream size due to encryption is not higher than 0.5% of the original coded bitstream.
- Simple design: Joint compression-encryption OMHT technique achieves both high security and compression performance in one single step, which simplifies the system design and reduces time required to perform compression followed by encryption.

References

[1 W. Lee, T. Chen and C. Chieh Lee, "1mprovement of an encryption scheme for binary images," Pakistan Journal of 1nformation and Technology. Vol. 2, no. 2, 2003, pp. 191-200, http://www.ansinet.org/

[2 Mitra, Y. V. Subba Rao, and S. R. M. Prasnna, "A new image encryption approach using combinational permutation techniques," Journal of computer Science, vol. 1, no. 1, 2006, p.127, http://www.enformatika.org

[3 H. El-din H. Ahmed, M. K Hamdy, and O. S. Farag Allah, "Encryption quality analysis of the RC5 block cipher. algorithm for digital images," Optical Engineering, Vol. 45, 1ssue 10107003, 2006, (7 pages).

[4 Y. Mohammad Ali and J. Aman," 1mage Encryption Using Block-Based Transformation Algorithm," 1AENG 1nternational Journal of Computer Science, Vol. 35, 1ssue. 1, 2008, pp. 15-23.

[5 M. V. Droogenbroech, R. Benedett, "Techniques for a selective encryption of uncompressed and compressed images," 1n AC1VS'02, Ghent, Belgium. Proceedings of Advanced Concepts for 1ntelligent Vision Systems, 2002.

[6 S. Fong, P.B. Ray, and S. Singh, "1mproving the lightweight video encryption algorithm," proceeding of iasted international conference, signal processing, pattern recognition and application, 2002, pp. 25-28.

[7 L.Wei-Bin, C. Tzung-her and L. Chen-Chieh, "1mprovement of an Encryption Scheme for Binary 1mages," Pakistan Journal of 1nformation and Technology. Vol. 2, no. 2, 2003, PP. 191-200, http://www.ansinet.org/

[8 R. m. Syed, "A new encryption algorithm for high throughput multimedia," 1N: 1nteractive Multimedia Systems, 2002, p. 269.

[9 Y. Xun, T. H. Chik, K. S. Chee, and R. S. Mahbubur, "Fast encryption for multimedia," Consumer Electronics, 1EEE Transactions on Publication Date: Feb 2001 Vol. 47, 1ssue: 1,2001, pp.101-107

[10 S. Changgui, B. K Bharat, "An efficient MPEG video encryption algorithm," Proceedings of the symposium on reliable distributed systems, 1EEE computer society Press, 1998, pp. 381-386.

[11 L. Qiao, K. Nahrstedt, and M.-C. Tam, "1s MPEG encryption by using random list instead of zigzag order secure?," in Proceedings of the 1EEE 1nternational Symposium on Consumer Electronics (1SCE '97), pp. 226-229, Singapore, December 1997.

[12 T. Uehara and R. Safavi-Naini, "Chosen DCT coefficients attack on MPEG encryption scheme," in Proceedings of 1EEE Pacific Rim Conference on Multimedia, pp. 316-319, Sydney, Australia, December 2000.

[13 H. Cheng and X. Li, "Partial encryption of compressed images and videos," 1EEE Transactions on Signal Processing, vol. 48, no. 8, pp. 2439-2451, 2000.

[14 C.-P. Wu and C.-C. J. K. Kuo, "Design of integrated multimedia compression and encryption systems," 1EEE Transactions in Multimedia, vol. 7, no. 5, pp. 828-839, 2005.

[15 C.-P. Wu and C.-C. Kuo, "Efficient multimedia encryption via entropy codec design," Proc. SP1E, vol. 4314, Jan. 2001.

[16 Xie and C. J. Kuo, "Enhanced Multiple Huffman Table (MHT) Encryption Scheme Using Key Hoping," 1n Proceedings of 1EEE 1nternational Symposium on Circuits and Systems, pp.568-571, May2004.

[17 Xie and C. J. Kuo ,"Multimedia Data Encryption via Random Rotation in Partitioned Bit Stream," 1n Proceedings of 1EEE 1nternational Symposium on Circuits and Systems, pp.568-571, May2004.

[18 W. Gillman and R. L. Rivest, "On breaking a Huffman code," 1EEE Transactions on 1nformation Theory, vol. 42, no. 3, pp. 972-976, 1996.

[19 S. P. Nana'Vati and K. P. Prasanta, "Wavelets: Applications to 1mage Compression-1," Joined of the Scientific and Engineering Computing. Vol. 9 , No.3: 2004, PP. 4-10 http://www.ias.ac.in/

[20 Ratael gonzales, e. Richard, and woods, "Digital image processing," 2nd ed, Prentice hall, 2002.

[21 AL. Vitali, A. Borneo, M. Fumagalli and R. Rinaldo, "Video over 1P using standard-compatible multiple description coding," Journal of Zhejiang University- Science A, vol. 7, no. 5 ,2006, pp. 668- 676.

[22 Shaimaa A. El-said, Khalid F. A. Hussein, and Mohamed M. Fouad, "Securing Multimedia Transmission Using Multiple Huffman Tables Technique," Electrical and Computer Systems Engineering Conference (ECSE'10), 2010.

[23 Shaimaa A. El-said, Khalid F. A. Hussein, and Mohamed M. Fouad, "Adaptive Lossy 1mage Compression Technique," Electrical and Computer Systems Engineering Conference (ECSE'10), 2010.

[24 C.-P. Wu and C.-C.J. Kuo. "Efficient multimedia encryption via entropy codec design". 1n Proc. SP1E 1nt. Symp. Electronic 1maging 2001, vol. 4314, Jan. 2001, p.128.

[25 J. Nechvatal et al. "Report on the Development of the Advanced Encryption Standard". National 1nstitute of Standards and Technology, U.S. Dept. Commerce, Tech. Rep., Oct. 2000.

[26 E. Filiol and C. Fontain. "A new ultra fast stream cipher design: COS ciphers". 1n Proc. 8th 1MA Conf. Cryptography and Coding, Dec. 2001.

[27 Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli "1mage Quality Assessment: From Error Visibility to Structural Similarity" 1EEE Transactions on 1mage Processing, Vol.13, No.4, April 2004.