PARTIAL ENCRYPTION OF COMPRESED VIDEO

Reena R. Ambekar[1], Harsh R. Bhagtani[2], Lavalesh M. Gupta[3],

Priya N. Nichani[4], Tushar P. Borad[5]

[1]A.P, EXTC, [2, 3, 4, 5]B.E. Student

reena.ambekar@yahoo.com, harshbhagtani92@gmail.com, loveleshg100@gmail.com, piyanshi1991@gmail.com, tusharborad123@gmail.com

Abstract

The traffic of digital images and video has grown rapidly in the internet. Security becomes important for several applications like military image database, confidential video conferencing. Real-time secure image and video communication is challenging due to the processing time and computational requirement for encryption and decryption. Also the speed of most internet connections is very limited so the amount of data to be transmitted must be reduced. In this research, we shall introduce partial encryption technique on compressed images and videos. Another algorithm decomposes images into several different parts. We apply encryption algorithm to encrypt only the crucial parts, which are considerably smaller than original image, which result in significant reduction in processing time and computational requirement for encryption and decryption as well as reduce the bit rate and bandwidth requirement.

Keywords-*Compression, MPEG-4, Partial Encryption*

----------***----------

1. INTRODUCTION

Advances in compression, delivery and presentation technologies of digital video in recent years have broadened the share of digital video in (audio) visual communication and entertainment, changing the ways that the end users create, access, store and copy video. In contrast to analog technologies, the digital technology offers

- Computer-aided content creation and manipulation,
- Transmission over computer networks,
- Storage and in computer environment,
- Production of identical copies without any specialized hardware.

However, the listed benefits bring a problem on access control. Video is transmitted over insecure networks, where a malicious party can acquire any packet, including those carrying private communication or commercially valued entertainment data.

[1].This work proposes a method where the video stream is partially encrypted and the distribution of encrypted bits over different syntactical entities of the video stream is optimized constrained to the number of encrypted bits, based on a simple model to assess the time to break the protection so that the average time to break the encryption over a temporal sample is maximized. Therefore, the developed method partial encryption method can be configured in a straightforward way, regarding the value of the data, providing solutions for [2] and [3].A method to estimate the parameters of the model is also proposed. The estimation method produces parameters depending on the video stream to be encrypted and it can be used simultaneously with encoding.

2. BACKGROUND ON VIDEO COMPRESSION

2.1 Video Compression

Video data requires large amount of space for storage in its raw form. For example, a one minute sequence of 352x288 RGB frames at 25fps is approximately 430MB. Fortunately, a large amount of spatial and temporal redundancy resides in such raw sequences, which can be reduced by compression. The human visual system is less sensitive to the chrominance information than the luminance information. Hence, one can down-sample the chrominance information in every individual frame to reduce the amount of data to represent perceptually equivalent frame [5]. A well-known approach for compression is to eliminate the spatial redundancy by transform coding, which involves transforming the image. The image in the transform domain can be approximated with all zero, but a few nonzero pixels.

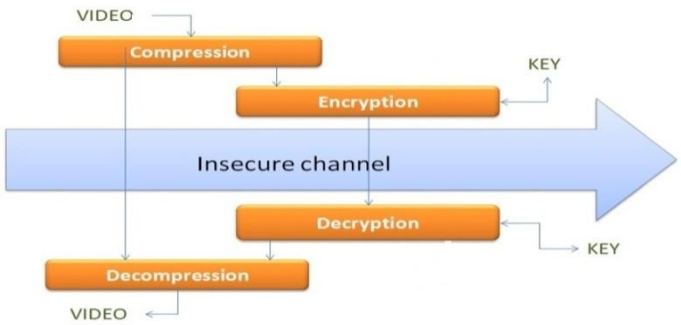

Fig 2.1: Block Diagram

Discrete cosine transform (DCT) and Wavelet transform are the most commonly used transformations [6]. Although inferior in compression, DCT is more commonly used than Wavelet transform since block wise DCT of the image is more suitable for block based motion estimation and it is also more popular than alternative motion estimation methods.

Consecutive frames of a video sequence are usually similar (except for the locations where the scene changes), with slight differences due to motion. The redundancies due to this similarity can be eliminated by modeling the motion. Any source of symbols can be compressed by entropy coding. The symbols are coded in a way that a symbol is mapped to a codeword with the length depending on the frequency of the symbol. Most of the video coding schemes prefer using prefix codes with predefined symbol to codeword mappings to eliminate the overhead due to transmission of the tables. An alternate method is arithmetic coding [7], which maps the string to be encoded to a number in the subinterval using the frequencies of symbols to be encoded. The optimal codeword assignment is achieved with arithmetic coding, but it requires more computational power. The entire process of video encoding can be summarized in Figure 2.1.

2.2 MPEG-4 Natural Video Coding Standard

MPEG-4 is a standard for coding audiovisual objects, enables re-use of audiovisual content, mixtures of natural and synthetic content and spatiotemporal arrangements of objects to form scenes. Thus, natural video coding tools were designed to be used with such compositions as well as ordinary rectangular image sequences. Most of these tools are specialized and practically applicable for a number of configurations. For example, robust and fast segmentation algorithms are required to encode nonrectangular video objects from a nature scene, on the other hand it's much easier with chroma keying in a studio environment. [8] and [9]

2.3 Natural Video Coding Tools Provided by MPEG-4

The audiovisual object is the basic entity in an MPEG-4 scene, which is described in the way specified in ISO/IEC 14496-1, as well as the transmission of the video object to the decoder. Each video object is characterized by spatial and temporal information in the form of texture, motion and shape. Texture has the spatial and motion has the temporal relation between the video samples and the spatiotemporal boundary of the samples is put by the shape information. An MPEG-4 scene may consist of one or more video objects [8]. The visual bit-stream provides a hierarchical description of a visual scene from video objects down to temporal samples of the video objects and the decoder can access any entity in the hierarchy by seeking certain code words called start codes, which are not generated elsewhere in the bit-stream. The hierarchy levels with their commonly used abbreviations are:

- Visual Object Sequence (VS)
- Video Object (VO)
- Video Object Layer (VOL)
- Video Object Plane (VOP)
- Group of Video Object Planes (GOV)

A profile is a subset of MPEG-4 coding tools and a level is the restrictions on the parameters of the encoding tools. Profile and level information is signaled in the bit stream so that a decoder can deduce whether it has the capability of processing the stream. Still textures are supported by scalable still texture profile and mapping of these textures on 2D dynamic meshes is supported by animated 2D mesh profile [9, 10].

2.4 Error Resilience and Concealment Tools

Every unencrypted piece of bit-stream is treated as a bit-stream error by a standard player. Therefore it is desired that the encryption scheme must be robust to any concealment tool which is available due to the nature of the video stream. Bit errors in VLC encoded data results loss of synchronization and the bit-stream till the next synchronization marker or start code cannot be decoded. In this way, error is localized and precise localization results more correct decoding [13]. MPEG-4 markers are placed into the bit-stream so that the macro-blocks between two resynchronization markers are just above a predetermined threshold. In this way, data is packetized so that each packet is equally important since they contain nearly the same amount of compressed bit-stream. The resynchronization marker is followed by the number of the first macro-block in the packet, its absolute quantization scale, optionally redundant header information and the macro-blocks in the packet. The predictive coding used to code the macro-blocks in a packet does not use prediction information from other macro-blocks. In addition to the packet approach, MPEG-4 also adopts a second method called fixed interval resynchronization. This method requires that VOP start codes and resynchronization markers appear at only fixed locations in the bit-stream, which avoids most of the problems due to start code emulation.

3. VIDEO ENCRYPTION

3.1 Cryptography

Cryptography is the subset of science concerned in encoding data, also called encryption, so that it can only be decoded, also called as decryption, by specific individuals. A system for encrypting and decrypting data is a cryptosystem. Encryption usually involves an algorithm for combining the original data ("plaintext") with one or more "keys" — numbers or strings of characters known only to the sender and/or recipient. The resulting output of encryption is known as "cipher text". There are two main classes of cryptosystems, with different practical application areas in today's technology. Public key methods use two different keys for encryption and decryption. On the other hand, secret key encryption methods use the same key for encryption and decryption.

3.2 Cryptosystems

Secret key methods can be classified in two groups, namely block and stream ciphers. Block ciphers encrypt and decrypt in multiples of blocks and stream ciphers encrypt and decrypt at arbitrary data sizes. Block ciphers are mostly based on the idea by Shannon that sequential application of confusion and diffusion will obscure redundancies in the plaintext, where confusion involves substitutions to conceal redundancies and statistical patterns in the plaintext and diffusion involves transformations (or permutations) to dissipate the redundancy of the plaintext by spreading it out over the cipher text. DES and Rijndael are examples of algorithms based on this idea, which allows simple hardware implementations or fast computer implementations by use of simple arithmetic, however they are not fast enough to encrypt large volumes of data in real time. Most of the stream ciphers rely on the fact that XORing the plaintext with a string only known to the sender and receiver provides strong encryption. In order to generate the string one can use a block cipher to encrypt a sequence known to both, as suggested in Rijndael specification. A stream can also be encrypted by block ciphers after being aligned to block boundaries, in cipher block chaining mode, where the encryption process of a block depends on the previous block due to XORing of previous cipher text with the plaintext of the block. The most popular public key method is RSA, which uses large prime numbers and modular arithmetic to encrypt a given text. RSA is slower and more complicated to be implemented in hardware since the primes are usually greater than 512-bits in size and the algorithm requires computation of powers and remainders with those large primes, the benchmark in Slagell's thesis [13] concludes that RSA is at least three times slower than secret-key methods and processing time increases cubically with key size on x86 architecture whereas secret-key methods cause slight increases. However, private key is not predictable given the public key and vice versa, therefore a sender-receiver pair can establish a one-way secure channel with the transfer of the encryption key from receiver to the sender. A common application of public key methods is to transfer a secret key to encrypt a larger amount of data.

3.3 Cryptanalysis

Cryptanalysis is the science concerned in breaking cryptosystems. Cryptanalysis generally involves the following main methods:

• A cryptanalyst can inspect a number of particular cipher texts for certain patterns and correlations. This method of attempting to break a cryptosystem is called a cipher text-only attack. An MMX implementation for inverse DCT requires not less than a thousand processor cycles per 8×8 block and iDCT counts one third of decoding effort

• The cryptanalyst may have the plaintexts besides the cipher texts. In this case, it may be possible to investigate the relation between the plaintexts and the corresponding Cipher texts. This type of attack is called a known-plaintext attack.

• As a last method, one can exhaustively try a set of keys until a decryption decided to be valid is achieved, which is impractical for large amounts of data or large key spaces.

3.4 Encryption Algorithm

In this algorithm we take MPEG-4 compressed video as data for starting the algorithm process. It will first read the number of frames and then each frame it will divide into type of frames i.e. I, P or B frames. Depending on the type then it will start actual encryption process. I-frame is an independent frame which requires full encoding process. P-frame is a Predicted frame which will be predicted from previous frame and requires partly encoding process.

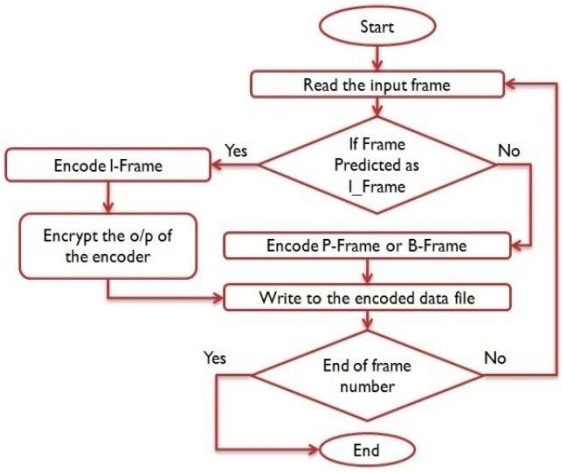

Fig 3.4: Encryption Algorithm

B-frame is a Bi-directional predicted frame which will depend on previous and future frames and requires partly encoding process as algorithm explained in figure 3.4.

4. TEST ENVIRONMENT

We have implemented encryption process on images. The following figures show the encrypted as well as the decrypted image.

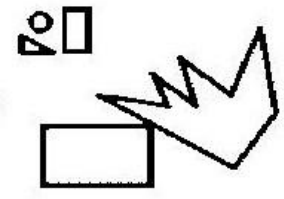

Fig 4(a): Original Image

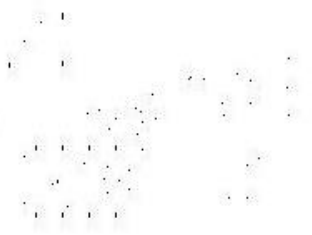

Fig 4(b): Encrypted Image

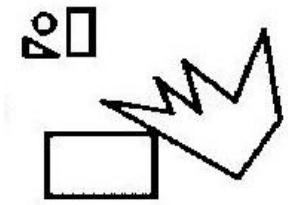

Fig 4(c): Decrypted Image

5. PROBLEM DEFINATION

However, the listed benefits bring a problem on access control. Video is transmitted over insecure networks, where a malicious party can acquire any packet, including those carrying private communication or commercially valued entertainment data.

The network, in particular the Internet, also allows peers to share their files, resulting exponentially increasing number of copies; a phenomenon called super-distribution. The path between the content creator and the viewer must be secured, so that the only viewers that are authorized by the content creator (or presenter) can access the video, which corresponds to preservation of privacy and prevention of piracy in one-to-one communication and broadcasting cases, respectively. It is also desirable that the viewer must be able to produce copies as long as a policy established by the content creator permits. Encryption of video, combined with access control logic implemented in the player is essential to prevent unwanted content acquisition. There are a number of issues to be considered while designing an access control mechanism, as pointed out by previous works:

1. Encryption (and decryption) of a video stream entirely takes considerable amount of time, which can be comparable to the decoding time. Therefore, only a carefully selected portion of video can be encrypted, to limit the cost of the operation.
2. The protection level for the content must be identified. Considering the business of copyrighted items trade, in particular entertainment, the increase in piracy boosts the demand for legitimate items. Therefore, paranoid protection may offend the end user and reduce the demand; on the other hand a loose protection mechanism may harm the business setup, reducing the revenues.
3. The protected video may have a limited lifetime, in the sense that it is of no value after some time on. For example, piracy makes sense if a protected live soccer broadcast can be broken until excerpts from the match are broadcast publicly in the succeeding sports programs. Therefore a protection scheme that needs just more time than the lifetime of the content is robust.

6. PROPOSED ENCRYPTION TECHNIQUE

As pointed by [13], data can be encrypted in any stage of the encoding process. However, every point is not equally advantageous in terms of format compliance, encryption overhead, compression efficiency, process ability, syntax awareness and transmission friendliness, which form a set of important criteria for many applications. Having attempted to encrypt every syntactical entity in the encoded video, the recent concerns of the study of video encryption were syntax compliance and process ability of the unencrypted bit stream by third parties to manipulate transmission rates and to allow searches. However, the limitation of the bit rate of the encrypted portion of the video stream while keeping security maximized remains as an open problem, which requires distribution of the budget of encrypted bits over the syntactical entities of video. Another unattacked problem is encoding of the encryption side information compactly and error resilient. We propose a novel solution called partial encryption, in which a secure encryption algorithm is used to encrypt only part of the compressed data for compressed images and videos. We shall implement various methods of compression and decide the best that suites us. Another algorithm decomposes images into several different parts. We will encryption algorithm to encrypt only the crucial parts, which are considerably smaller than the original image, which result in significant reduction in processing time and computational requirement for encryption and decryption as well as reduce the bit rate and bandwidth requirement.

7. RESULT

Access control on the media is essential in both commercial broadcasting and peer-to-peer communication. An access control mechanism must be supported by encryption in order to ensure that only authorized accesses are possible. Partial encryption takes relatively small amount of time, compared to the decoding process; the time is not negligible, however. A configurable yet maximally secure encryption method is required, as not all video streams are of equal value. In order to accomplish this task, a solution to encrypt video is proposed, which consists of a simple model of the average time required to break a portion of the encrypted video.

CONCLUSIONS

Unlike previously developed video encryption methods, the proposed method is capable of controlling the rate of the encrypted stream at a level that can be specified by the content

creator/provider while keeping the stream as robust as possible. The advantage of this control can be used in two ways:

• The content provider can assess the level of encryption to protect a video stream with known value

•The player designer/implementer can estimate the computational power required to play a video with certain security requirements, which will lead to a more efficient design. The model is generalizable to other video coding schemes, including codecs with temporal scalability; the procedure involves identification of levels with dependency relations between one another.

FUTURE SCOPE

Future research will focus on adding support for encryption framework, thereby further increasing the flexibility of the encryption process. The method can be extended to video access control implementations that en-crypt the indexes of codeword instead of code words themselves, instead of the direct encryption method implemented here. One can set up a series of experiments to establish a number of encryption profiles. Selection between preset profiles might be helpful with a lightweight encoder. Experiments show that the encryption in slow motion video is more uniform, thus this method can be implemented for video communication applications.

BIOGRAPHIES

[1] M. Eskicioglu, J. Town, and E. J. Delp, "Security of digital entertainment content from creation to consumption," Signal Processing: Image Communication, vol. 18, pp. 237–262, and 2003.

[2] J. Wen, M. Severa, W. Zeng, M. Luttrell, and W. Jin, "A format compliant configurable Encryption framework for access control of multimedia," in Proc. IEEE Workshop on Multimedia Signal Processing, pp. 435–440, 2001.

[3] A. M. Eskicioglu and E. J. Delp, "An overview of multimedia content protection in consumer electronics devices," Signal Processing: Image Communication, vol. 16, pp. 681–699, 2001.

[4] A. Gayer and O. Shy, "Copyright protection and hardware taxation," Information Economics and Policy, vol. 0 (in print), pp. 0–0, 2003

[5] A. M. Tekalp, Digital Video Processing, pp. 432–500. Prentice Hall, 1995.

[6] S. Mallat, A Wavelet Tour of Signal Processing. Academic Press, 1998.

[7] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Numerical Recipes in C: The Art of Scientific Computing, pp. 903–926. Cambridge University Press, 2 ed., 1992.

[8] T. Ebrahimi and C. Horne, "MPEG-4 natural video coding – an overview," Signal Processing: Image Communication, vol. 15, pp. 365–385, 2000.

[9] ISO/IEC JTC1/SC29 WG11, ISO/IEC 14496-2 FCD N2202, March 1998.

[10] R. Koenen, "Profiles and levels in MPEG-4: approach and overview," Signal Processing: Image Communication, vol. 15, pp. 463–478, 2000.

[11] H. Cheng, "Partial encryption for image and video communication," Master's thesis, University of Alberta, 1998.

[12] T. Zhang, U. Jennehag, and Y. Xu, "Numerical modeling of transmission errors and video quality of MPEG-2," Signal Processing: Image Communication, vol. 16, pp. 817–825, 2001.

[13] A. J. Slagell, "A simple, portable and expandable cryptographic application pro-gram interface," Master's thesis, University of Illinois at Urbana-Champaign, 2003.

[14] M. Wu and Y. Mao, "Communication-friendly encryption of multimedia," in Proc. of IEEE Multimedia System Processing Workshop, 2002.