Selective Embedding and Forbidden Zone Data Hiding for Strong Video Data Thrashing

R. Ravi Kumar[#1], V. Kesav Kumar[#2]

[#1]*II Year M.Tech Student,* [#2]*Associate Professor*

[#1,#2] *CSE Department, Sri Mittapalli College of Engineering,*

Guntur, Andhra Pradesh, India.

***Abstract*— A new method for high capacity data hiding in H.264 streams takes advantage of the different block sizes used by the H.264 encoder during the inter prediction stage in order to hide the desirable data. This fragile data hiding approach can be mainly used for content-based authentication and covert communication. Information-theoretic analyses for data hiding prescribe embedding the hidden data in the choice of quantizer for the host data. The hidden data can be recovered reliably under attacks, such as compression and limited amounts of image tampering and image resizing. The three main findings are as follows. 1) In order to limit perceivable distortion while hiding large amounts of data. 2) The use of local criteria to choose where to hide data can potentially cause desynchronization of the encoder and decoder. 3) For simplicity, scalar quantization-based hiding is employed, even though information-theoretic guidelines prescribe vector quantization-based methods. We begin with a review of two major types of embedding, based on which we propose a new multilevel embedding framework to allow the amount of extractable data to be adaptive according to the actual noise condition. We propose a new video data hiding method that makes use of erasure correction capability of repeat accumulate codes and superiority of forbidden zone data hiding. Selective embedding is utilized in the proposed method to determine host signal samples suitable for data hiding. The decoding error values are reported for typical system parameters. The simulation results indicate that the framework can be successfully utilized in video data hiding applications.**

***Keywords*— Data Hiding, covert communication, authentication, Quantization, Selective Embedding, Forbidden Zone Data Hiding.**

I. INTRODUCTION

The widespread of the Internet and World Wide Web has changed the way digital data is handled. Data hiding deals with the ability of embedding data into a digital cover with a minimum amount of perceivable degradation, i.e., the embedded data is invisible or inaudible to a human observer. Data hiding consists of two sets of data, namely the cover medium and the embedding data, which is called the message. Only few data hiding algorithms considering the properties of H.264 standard [1] have recently appeared in the open literature.

Transform domain is generally preferred for hiding data since, for the same robustness as for the spatial domain; the result is more pleasant to the Human Visual System (HVS). For this purpose the DFT (Discrete Fourier Transform), the DCT (Discrete Cosine Transform), and the DWT (Discrete Wavelet Transform) domains are usually employed.

We seek to embed much larger volumes of data than required for watermarking, targeting applications such as steganography and seamless upgrade of communication and storage systems, rather than digital rights management. Second, because of our target applications, we aim for robustness not against malicious attacks such as Stirmark's geometric attacks, but against "natural" attacks, such as compression (e.g., a digital image with hidden content may be compressed as it changes hands or as it goes over a low bandwidth link in a wireless network).

The gap between the theoretical embedding capacity in data hiding and what is achievable in practice can be bridged by investigation of such issues as basic embedding mechanisms for embedding one bit and modulation/ multiplexing techniques for embedding multiple bits. The following problems require particular attention:

- *Distortion*: The distortion introduced by watermarking must be imperceptibly small for commercial or artistic reasons. However, an adversary intending to obliterate the watermark may be willing to tolerate certain degree of visible artifacts.
- *Actual noise conditions*: An embedding system is generally designed to survive certain noise conditions. The watermarked data may encounter a variety of legitimate processing and malicious attacks, so the actual noise can vary significantly. Targeting conservatively at surviving severe noise would lead to the waste of actual payload, while targeting aggressively at light noise could result in the corruption of embedded bits.
- *Uneven distribution of embedding capability*: The amounts of data that can be embedded often vary widely from region to region in image and video. This uneven embedding capacity causes serious difficulty to high-rate embedding.

Uncompressed video data has been utilized by most of the video data hiding methods. A large quantity change domain data hiding in MPEG-2 videos is proposed by Sarkar *et al.* [2]. They applied QIM to low frequency DCT coefficients and adapted the quantization parameter based on MPEG-2 parameters. In order to survive erasures, they used Repeat Accumulate (RA) codes. These codes are already used in image data hiding. Adaptive block selection [3] outcomes in de-synchronization and they used RA codes to switch

erasures. These operations such as insertions and erasures can also handled by convolution codes as in [4]. To accurate de-synchronization faults, numerous parallel Viterbi decoders are used. When the amount of chosen host signal samples is much less than the total number of host signals samples, then only it is observed that such a scheme is successful.

3-D DWT domain is utilized to conceal data in [5]. LL sub-band coefficients have been used and they do not perform any adaptive selection. To increase error correction capability, they used BCH code. Two methods for applying local criteria are considered. The first is the block-level entropy thresholding (ET) method, which decides whether or not to embed data in each block (typically 8×8) of transform coefficients, depending on the entropy, or energy, within that block. The second is the selectively embedding in coefficients (SEC) method, which decides whether or not to embed data based on the magnitude of the coefficient.

The rest of the paper is arranged as follows. Section II describes the Data hiding schemes which discusses about H.264 video encoder. The study of Forbidden Zone Data Hiding has been discussed in Section III. The proposed video data hiding framework is presented in Section IV. Section V deals with the Experimental consequences what we got. Finally, conclusion is given in Section VI.

II. DATA HIDING SCHEME

The main blocks of the H.264 video encoder are depicted in Fig. 1. The Temporal Prediction block is responsible for the *inter prediction* of each inter frame. Our scheme intervenes in the inter prediction process in order to hide the data.

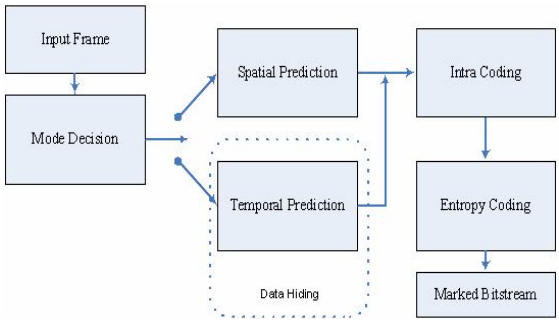

Fig 1: H.264 Video Encoder

The most important part of inter prediction is the motion estimation process, which aims at finding the "closest" macro block (best match) in the previously coded frame for every macro block of the current input frame. Then each macro block, within the current frame, is motion compensated, i.e. its best match is subtracted from it, and the residual [12] macro block is coded. In order to increase the coding efficiency, the H.264 standard has adopted seven (7) different block types (16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4) and the motion estimation is applied on each of these types.

First, assign a binary code to every block type according to Table 1. For simplicity we use only 4 block types. That gives us 2 bits per block. Then we convert the embedding message into a binary number and we separate the bits in pairs. These pairs are mapped into macro blocks, which are going to be motion compensated, using the chosen block types.

Block Type	Binary Code
16×16	00
16×8	01
8×16	10
8×8	11

Table 1: Binary Codes of the Block Types

It is also important to define the data hiding parameters such as:

1. **Starting frame:** It indicates the frame from which the algorithm starts message embedding.
2. **Starting macro block:** It indicates the macro block within the chosen frame from which the algorithm starts message embedding.
3. **Number of macro blocks:** It indicates how many macro blocks within a frame are going to be used for data hiding. These macro blocks may be consecutive or even better; they may be spread within the frame according to a predefined pattern.
4. **Frame period:** It indicates the number of the inter frames, which must pass, before the algorithm repeats the embedding. This parameter is very important since it increases the possibilities of extracting the message even if some parts of the video sequence are missing.

We can view these elements through a layered structure shown in Fig. 2, analogous to that in communications. The lower layers deal with how one or multiple bits are embedded imperceptibly in the host media. Upper layers for achieving additional functionalities can be built on top of these lower layers.

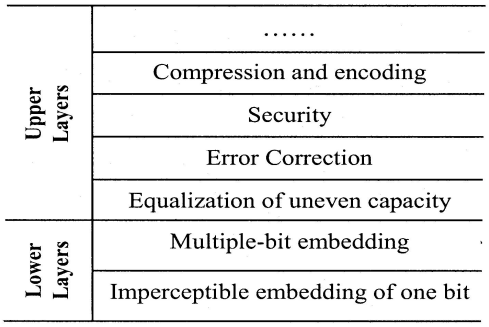

Fig 2: Layered Structure of Data Hiding

III. FORBIDDEN ZONE DATA HIDING

Forbidden zone data hiding (FZDH) has been introduced in [6]. Forbidden zone (FZ) methods are defined as that no change is permissible at the time of data hiding process for a host signal range. FZ has been used by FZDH to regulate the strongness-invisibility tradeoff.

The main concept of FZDH is the identification of zones and the partitions. So many ways are there to attain this; however, by the use of quantizers, practical design can be

performed. How this design can be performed is shown in below equation, where the mapping function is defined as:

$$M_m(s) = \left\{ s + e_m \left(1 - \frac{r}{\| e_m \|} \right) \right\}$$

To regulate the necessity of mutual exclusion, the reconstruction points of the quantizers that are indexed by different *m* should be non-overlapping, which can be achieved by using a base quantizer and shifting its reconstruction points depending on *m*, similar to Dither Modulation [7].

IV. Proposed Video Data Hiding Framework

A block based adaptive video data hiding method has been proposed by us that incorporates FZDH, which is shown to be better-quality to QIM and spirited with DC-QIM [6], and erasure handling through RA Codes. We employ block selection and coefficient selection together rather than like in [3]. RA Codes as in [2] and [3] are used to handle the de-synchronization due to block selection. By using multi-dimensional form of FZDH in varying extents, the de-synchronization due to coefficient selection is handled. The frames are processed independently and it is noticed that [8] intra and inter frames don't give way significant dissimilarities.

The proposed scheme may result in very high capacity proportional to the host video sequence size. Its major advantage is that it does not affect the visual quality of the video sequence and if the hiding parameters are properly controlled it does not affect the coding efficiency. Finally, the message can be extracted directly from the encoded video stream without the need of the original host video sequence.

A. Framework:

A typical data hiding framework is illustrated in Fig. 3. Starting with an original digital media (I_0), which is also known as the *host media* or *cover media*, the embedding module inserts in it a set of secondary data (b), which is referred to as *embedded data* or *watermark*, to obtain the *marked media* (I_1). The insertion or embedding is done such that I_1 is perceptually identical to I_0. The difference between I_1 and I_0 is the distortion introduced by the embedding process.

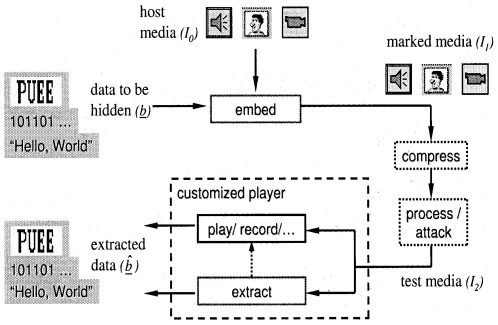

Fig 3: General Framework of Data Hiding Systems

Data embedding process for a solitary frame is shown in Fig. 4. In this, for data embedding Y-channel is utilized. The selected frames are processed block-wise after the frame selection is performed. Simply a single bit is hidden for every block.

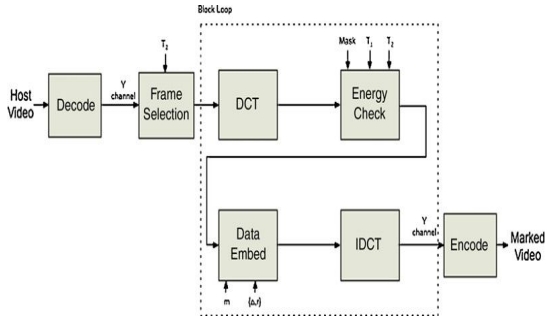

Fig 4: Embedder flowchart of the proposed video data hiding framework for a single frame.

Dual of the embedder is nothing but a Decoder, with the exception that frame selection is not performed. Process flow for a single frame is shown in Fig 5. Frame synchronization markers are used to identify Marked Frames. Decoder employs the same system parameters and determines the marked signal values that will be fed to data extraction step.

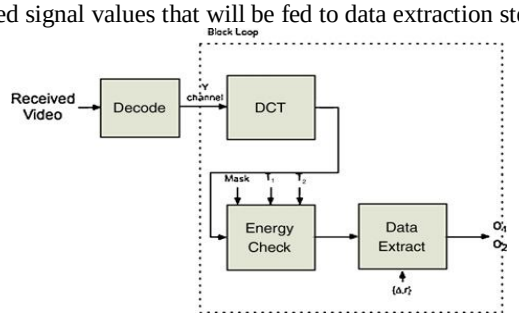

Fig 5: Decoder flowchart of the proposed video data hiding framework for a single frame.

B. Selective Embedding:

In the selectively embedding scheme, instead of deciding where to embed at the block level, we do a coefficient-by-coefficient selection, with the goal of embedding in those coefficients that cause minimal perceptual distortion. The selection is performed at four stages:

1. **Frame selection:** Selected number of blocks in the whole frame is counted.
2. **Frequency band determination:** Only certain DCT coefficients are utilized.
3. **Block selection:** energy of the coefficients in the mask is computed.
4. **Coefficient selection:** Each coefficient's energy is compared to another threshold.

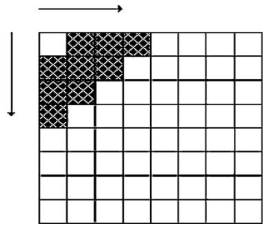

Fig 6: Sample coefficient masks denoting the selected frequency band.

In the SEC scheme, we have more control on *where to hide data* compared to the ET scheme; hence, it achieves better performance in terms of smaller perceptual degradation for a given amount of data. Another key advantage of the scheme is that it automatically determines the right amount of data to be hidden in an image based on its characteristics.

C. Repeat-Accumulate (RA) Coding for SEC Scheme:

Any turbo-like code that operates close to Shannon limit for the erasures channel, while possessing a reasonable error-correcting capability, could be used with the SEC scheme. We used RA codes in our experiments because of their simplicity and near-capacity performance for erasure channels. This codeword is hidden using the local criteria such that if a coefficient does not pass the threshold test, the corresponding code symbol is erased (i.e., not hidden).

D. Block Division:

Two displace data sets are embedded: message bits (m_1) and frame synchronization markers (m_2). With the help of a random key, the block locations of m_2 are determined. The remaining blocks are reserved for m_1. The same partitioning is used for all frames and the characteristic division is shown in Fig. 7; m_1 is dispersed to T consecutive frames and m_2 is embedded frame by frame.

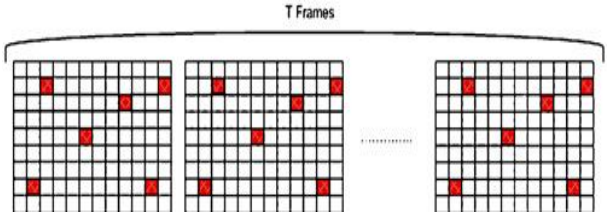

Fig 7: Typical block partitioning for message bits and frame synchronization markers.

V. Experimental Consequences

Experiments are done in three stages:

1. Without any error correction, comparison of QIM and FZDH with the help of their raw decoding error performances.
2. Performance of the proposed framework has been observed against various common video processing attacks.
3. Proposed video data hiding framework has been compared against JAWS [9], [10] and the method in [2] by using MPEG-2 compression attack.

A. Forbidden Zone Data Hiding Versus Quantization Index Modulation:

At the same embedding distortion and data hiding rate, QIM and FZDH are compared. Two unusual embedding distortion values are used: 41 dB and 45 dB average PSNR. The normal PSNR between host and marked frames is calculated as Embedding distortion. We have to identify that dissimilar pairs of (Δ, r) may give in the same embedding distortion. By manually, we make use of distinctive values. T_1 is selected as 3000 and T_2 is set to 2000. A typical host and marked frame pair for FZDH (at 42 dB).

B. General Video Dispensation Assaults:

We affect error correction and assess the performance of FZDH against some common video processing attacks in the second stage. We use a distinctive TV transmit material of 15min. From that, we choose a slighter period, which is still precise to draw terminations, due to the computational burden of RA decoding.

The effect of the parameters on the number of selected block rate has been observed first. The number of the selected blocks depends on the content and varies slowly with time. Corresponding to shot boundaries, the abrupt will change. What we identify is that embedder and decoder select unusual number of blocks. Higher number of blocks has been selected by the decoder for low rates.

The decoding error performance against compression attack has been observed next. We use different bitrates for this resolution and those results indicate that we need repetition number higher than the erasure rate. The reason for this observation is due to the fact that decoding errors occur as a result of compression as well as the erasures due to the block selection.

The performance of the method against another common video processing: frame-rate conversion has been tested. The original video frame-rate is 20 f/s. To measure the decoding error rate, we vary this frame rate to a superior as well as a inferior value. We should note that frame-rate conversion could be achieved in various ways, some of which could be quite complex [13].

C. Proposed Framework:

According to the type of the frame, the quantization parameter is adaptively adjusted. By the means of RA codes, the resultant desynchronization due to coefficient selection is handled. The decoding error decreases with decreasing embedding distortion with the different utilization of I/P/B frames results in the unexpected situation. However, we base our comparison with the best result obtained in [2]. We utilize the same host video as in [2]. We adjust method parameters according to the host video i.e. QVGA size. First, we reduce the repetition number, R, to 5 and obtain embedding rate of 280 bits per frame, whereas the best result is obtained for 273.6 bits per frame in [2].

VI. Conclusion

The image-in-image hiding presented here uses the fact that we can send a high volume of data with robustness against JPEG compression using the uncoded SEC scheme. The signature image is compressed into a sequence of bits and these bits are hidden into the host (disregarding the actual meaning of the bits). The system is designed for the worst anticipated attack. In practice, the attack level is seldom known *a priori*, and if the actual attack is less severe than the design attack, we are still stuck with the design signature image quality.

FZDH and QIM are compared as the data hiding method of the proposed framework. Especially for low embedding distortion levels, we found that QIM is inferior to FZDH. The framework was tested with MPEG-2, H.264 compression, scaling and frame-rate conversion attacks. The experimental consequence gives that this framework can be productively used in video data hiding applications. Tardos fingerprinting [11], which is a randomized construction of binary fingerprint codes that are optimal against collusion attack, can be employed within the proposed framework. The proposed method has been compared with the canonical watermarking method, JAWS, and a more recent quantization based method [2]. The results point out a significant advantage over JAWS and a similar presentation with [2].

ACKNOWLEDGEMENT

We, authors express gratitude to all the anonymous reviewers for their affirmative annotations among our paper and also would like to thank the anonymous reviewers for their valuable comments.

VII. REFERENCES

[1] S. K. Kapotas, E. E. Varsaki, and A. N. Skodras, "Data hiding in H- 264 encoded video sequences," in *Proc. IEEE 9th Workshop Multimedia Signal Process.*, Oct. 2007, pp. 373–376.

[2] A. Sarkar, U. Madhow, S. Chandrasekaran, and B. S. Manjunath, "Adaptive MPEG-2 video data hiding scheme," in *Proc. 9th SPIE Security Steganography Watermarking Multimedia Contents*, 2007, pp. 373–376.

[3] K. Solanki, N. Jacobsen, U. Madhow, B. S. Manjunath, and S. Chandrasekaran, "Robust image-adaptive data hiding using erasure and error correction," *IEEE Trans. Image Process.*, vol. 13, no. 12, pp. 1627–1639, Dec. 2004.

[4] M. Schlauweg, D. Profrock, and E. Muller, "Correction of insertions and deletions in selective watermarking," in *Proc. IEEE Int. Conf. SITIS*, Nov.–Dec. 2008, pp. 277–284.

[5] H. Liu, J. Huang, and Y. Q. Shi, "DWT-based video data hiding robust to MPEG compression and frame loss," *Int. J. Image Graph.*, vol. 5, no. 1, pp. 111–134, Jan. 2005.

[6] E. Esen and A. A. Alatan, "Forbidden zone data hiding," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2006, pp. 1393–1396.

[7] B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[8] E. Esen, Z. Do˘gan, T. K. Ates, and A. A. Alatan, "Comparison of quantization index modulation and forbidden zone data hiding for compressed domain video data hiding," in *Proc. IEEE 17th Signal Process. Commun. Applicat. Conf.*, Apr. 2009, pp. 404–407.

[9] M. Maes, T. Kalker, J. Haitsma, and G. Depovere, "Exploiting shift invariance to obtain a high payload in digital image watermarking," in *Proc. IEEE ICMCS*, vol. 1. Jul. 1999, pp. 7–12.

[10] T. Kalker, G. Depovere, J. Haitsma, and M. J. Maes, "Video watermarking system for broadcast monitoring," in *Proc. SPIE Security Watermarking Multimedia Contents Conf.*, vol. 3657. 1999, pp. 103– 112.

[11] G. Tardos, "Optimal probabilistic fingerprint codes," in *Proc. 35th Annu. ACM STOC*, 2003, pp. 116–125.

[12] Y. Bodo, N. Laurent, J.-L. Dugelay, "Watermarking Video, Hierarchical Embedding in Motion Vectors", Proc. Int. Conference on Image Processing, Sept. 2003.

[13] M. Wu, H. Yu, and B. Liu, "Data hiding in image and video: I. Fundamental issues and solutions," *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 685–695, Jun. 2003.