Efficient Watermarking for High Efficiency Video Coding

A Thesis

submitted in fulfillment of the requirements for the award of the degree of

Doctor of Philosophy

by

Gagandeep Kaur
(901311008)

Under the guidance of

Dr. Singara Singh Kasana
Associate Professor

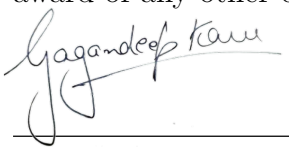
Dr. M. K. Sharma
Professor

Computer Science and Engineering Department
Thapar Institute of Engineering and Technology
Patiala–147004, India
October, 2020

Certificate

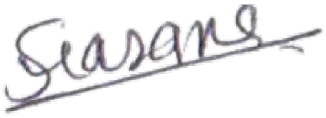
This is to certify that the work presented in this thesis, entitled **Efficient Watermarking for High Efficiency Video Coding**, being submitted to Department of Computer Science and Engineering, Thapar Institute of Engineering and Technology, in fulfilment of the requirements for the award of the degree of **Doctor of Philosophy**, is a bonafide record of my work carried under the supervision of **Dr. Singara Singh Kasana** and **Dr. M. K. Sharma**. I have given the due credit to the sources of the text/figures/tables and theoretical analysis used, by citing them in the thesis and details of the sources are given in the references.

The work presented in this thesis has not been submitted to any other institution for the award of any other degree.

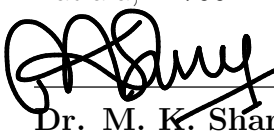


Gagandeep Kaur

This is to certify that the work done in this thesis is worthy of consideration for the award of the degree of **Doctor of Philosophy** in accordance to the norms of the Institution. To the best of our knowledge, the above statement made by the candidate is right.



Dr. Singara Singh Kasana
Associate Professor
Computer Science and Engineering Department
Thapar Institute of Engineering and Technology
Patiala, 147004



Dr. M. K. Sharma
Professor
School of Mathematics
Thapar Institute of Engineering and Technology
Patiala, 147004

Dedicated to

Maa, Papa and Ranjit

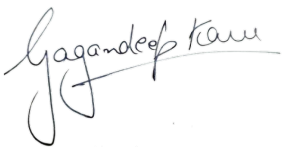
Whose blessings, motivation and love made my path of success

Acknowledgements

First of all, I would like to verbalise my profoundest appreciation to supervisors, **Dr. Singara Singh Kasana**, and **Dr. M. K. Sharma** for their valuable guidance and inspiration. This thesis has been possible only with their constant encouragement and support. Their support in this journey also includes advice on a personal level, which goes above and beyond as a supervisor. I am genuinely thankful for their patience with me in times of ups and downs in my life.

I want to acknowledge the members of the Doctoral committee: **Dr. Maninder Singh**, **Dr. Maninder Kaur**, **Dr. Jhilik Bhattacharya** and **Dr. Mandeep Singh** for monitoring the progress of this research work and providing their insightful suggestions and encouragement. Their suggestions helped me to widen the research from various perspectives. I am thankful to all employees of the Institute for extending their co-operation in terms of technical and official support for the successful completion of this research work. I also thank the University Grants Commission for helping financially the research work in this thesis through the Maulana Azad National Fellowship scheme.

I further want to thank my brothers Ranjit and Gurjot, who constantly stood by me and taught me to never give up during hard times in this journey. In the end, I want to convey my heartfelt gratitude to my parents, relatives and friends for their affection, consolation and care.



(**Gagandeep Kaur**)

List of Publications

Published

1. **Gagandeep Kaur**, Singara Singh Kasana and M. K. Sharma. "An Efficient Watermarking Scheme for Enhanced High Efficiency Video Coding/H.265" *Multimedia Tools and Applications*, Springer, Vol 78(9), pp. 12537–12559 (2019), Impact Factor = 2.313.

2. **Gagandeep Kaur**, Singara Singh Kasana and M. K. Sharma. "An Efficient Authentication Scheme for High Efficiency Video Coding/H.265" *Multimedia Tools and Applications*, Springer, Vol 78(15), pp. 21245–21271 (2019), Impact Factor = 2.313.

Communicated

1. **Gagandeep Kaur**, Singara Singh Kasana and M. K. Sharma. "A Watermarking Scheme for P-frames of Enhanced HEVC/H.265" *The Visual Computer*, Springer [Under Review]

Abstract

According to the report of the Cisco, around 73% of the global traffic of data is in the form of High Definition (HD) and Ultra High Definition (UHD) quality videos. HEVC provides good visual output with 50% more compression accuracy contrasted with that of the previous H.264 / AVC encoding format. Consequently, nowadays, HEVC is hugely used for transmission of HD and UHD videos over the Internet. Therefore, video security concerns such as content protection and authentication of HEVC encoded videos have emerged an important area of research. Over the last few decades, digital watermarking has been used to deter copying, preserve integrity of content, track video transmission and dissemination. The present work is based on developing efficient watermarking schemes for authentication and protection of HEVC encoded videos by addressing the drawbacks of existing literature.

The main features of the authentication scheme are that it should detect the malicious tampering in the contents of the video and must resist the non-malicious contents preserving manipulations that are done for transcoding the encoded videos. The main drawback of the existing authentication schemes is that they cannot resist the transcoding processing, *i.e.* re-compression with different Quantization Parameter (QP). To overcome this problem, in the present thesis, an authentication scheme is designed for HEVC, which can help shield the sanctity of the video during transcoding. In new scheme, invariant features, *i.e.* the relationship between the signs of negative and positive coefficients in blocks, is used to generate the authentication code. These features do not change after re-compression of the video with the same or different QP values. On the other hand, the generated authentication code is fragile to any malicious alterations such as insertion and deletion of objects in the contents of the video. The generated code is inserted into the video by altering the magnitudes of negative and positive coefficients. The integrity of the video is verified at the decoder end in two steps. The integrated authentication code is retrieved from the video during the first process. In the second step, the authentication code is re-generated from the contents of the video. Finally, the extracted and re-generated authentication codes are compared to detect any malicious tampering done in the videos.

There are many schemes in existing literature for the protection of HEVC encoded videos. The compression efficiency of the HEVC is further improved by introducing new functionalities in the transformation process of the standard. For this purpose, a hybrid transformation is applied to the residual error to compact its maximum energy into fewer

non-zero coefficients. In hybrid transformation, multiple transforms, namely Discrete Cosine Transform (DCT) and Discrete Sine Transform (DST) are applied on the single block of residual error.

The existing single transformation based protection schemes are not effective for enhanced HEVC in contexts of resistance to attacks and barely noticeable modification throughout visual quality. Firstly, the existing schemes are not resilient to re-encoding with various QP values. This is due to the coefficients generated using different kernels of DCT and DST transform that have varying features with regards to stability and amount of energy contained in each coefficient. In the case of DCT, the first coefficient is the DC coefficient, which contains the maximum energy, and stability of the coefficients decreases from low to high frequency. On the other hand, the DST generated coefficients do not follow this phenomenon of stability and energy compaction. Moreover, some of the kernels used in hybrid transformation are not symmetric. Therefore any distortion induced by altering the coefficients of these transforms cannot be nullified and are spread to the neighbouring blocks. This mostly results in recognizable artifacts in the video after embedding the watermark. To overcome these problems, hybrid transforms oriented watermarking scheme is developed for the protection of enhanced HEVC video using I-frames. The developed I-frame protection scheme embeds the watermark by altering the residual error blocks. For this, first, the residual error is transformed using a symmetric transform, which is selected based on the hybrid transformation process used by the encoder. Then, the magnitudes of the transform coefficients generated have been adjusted whereas maintaining the same signs of coefficients, as any changes in the signs of the coefficients degrade the quality drastically.

Currently, there is extensive use of Low Delay (LD) configuration for the compression of the video that are employed in the video applications specifically surveillance and conferencing system. The LD configuration is the primary video coding structure, of that every first frame is an I-frame while the subsequent frames being P-frame. So, a P-frame based protection scheme is proposed for video applications with low bit-rate. The above I-frame based protection scheme does not give optimal results for P-frames. In P-frames, dual-mode (intra and inter) is used in the prediction process. The blocks in P-frames are predicted from spatial and temporal neigbours. This results in the synchronization error, which is due to the changes in the prediction mode of the blocks after the re-compression, *i.e.* the prediction mode is changed from intra to inter and vice-versa. The blocks for embedding the watermark are selected using a robustness threshold to fix this issue of synchronization error. The robustness threshold is computed on the fly by exploiting the magnitude of the QuantizedTransform Coefficients (QTC) in the P-frame'sintra-and

inter-luma blocks. Further, the embedding of the watermark in P-frames results in a significant increase in bit-rate of video. For controlling this increase in bit-rate, the pair of coefficients with non-zero magnitudes are altered to embed the watermark.

Table of Contents

List of Figures

List of Tables

List of Notations

α	threshold calculated based on non-zero QDST coefficients
AUC_k	authentication code with k number of bits
AUC'_k	re-constructed authentication code during the decoding process
AUC''_k	authentication code retrieved bythe decoding process
β	threshold calculated based on QDST coefficients with ABGR1
BR	un-watermarked video's bit-rate
$BR^{'}$	watermarked video's bit-rate
$E(TE)$	the expected value of TE
$E(STE)$	the expected value of STE
F	cumulative distribution function
$\overline{F}$	complementary cumulative distribution function
F_D	decryption function
F_E	encryption function
F^o	original frame in a video
F^w	watermarked frame in a video
IF_j	j^{th} I-frames of a video
k	number of CTU's in a frame
μ	CCDF calculated based on QDST coefficients with ABGR1
ν	CCDF calculated based on non-zero QTC's
n_1	width of each frame in a video
n_2	height of each frame in a video
NE	count of elements which are tampered within an AIM
$p(D)$	probability distribution
$P(D)$	complementary probability distribution
PE_k	palette generated by candidate block selection algorithm
PG_k	palette generated by authentication code geneation
$PSNR_Y$	luma component's $PSNR$
$PSNR_U$	chroma Cb component's $PSNR$
$PSNR_V$	chroma Cr component's $PSNR$
$PSNR_{YUV}$	un-watermarked YUV video's $PSNR$
$PSNR^{'}_{YUV}$	watermarked YUV video's $PSNR$
δ_{PSNR}	difference among the un-watermarked and watermarked videos YUV $PSNR$
ρ	CCDF based on non-zero QTC's coefficients with magnitudes greater than one
S_g	watermark generation subset of blocks

S_e	watermark embedding subset of blocks
T_γ	tamper detection threshold
T_λ	malicious tampering threshold
T_1	threshold based on non-zero QTC's
T_2	threshold based on non-zero QTC's with magnitudes greater than one
τ	pseudo random bit
$Var(STE)$	variance of STE
W_m	watermark sequence with m number of bits

List of Abbreviations

ABGR1	Absolute value Greater than 1
BIR	Bit Increase Rate
CABAC	Context-Based Adaptive Binary Arithmetic Coding
CAVLC	Context Adaptive Variable Length Coding
CB	Coding Block
CCDF	Complementary Cumulative Distribution Function
CTB	Coding Tree Block
CTU	Coding Tree Unit
CU	Coding Unit
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DST	Discrete Sine Transform
DWT	Discrete Wavelet Transform
EMT	Enhanced Multiple Transform
HD	High Definition
HEVC	High Efficiency Video Coding
HVS	Human Visual System
JCT-VC	Joint Collaborative Team on Video Coding
LD	Low Delay
LSB	Least Significant Bit
MPEG	Motion Picture Expert Group
NCG	Normed Centre of Gravity
NQDST	Negative Quantized Discrete Sine Transform
PB	Prediction Block
PQDST	Positive Quantized Discrete Sine Transform
PSNR	Peak Signal to Noise Ratio
QDST	Quantized Discrete Sine Transform
QP	Quantization Parameters
QTC	Quantized Transform Coefficients
RQT	Residual Quad Tree
SSIM	Structural Similarity index
STE	Strongly Tampered Elements
STP	Strongly Tampered Percentage
TE	Tampered Elements

TP	Tampered Percentage
TB	Transform Block
UHD	Ultra High Definition
VIFp	Visual Information Fidelity in Pixel Domain

Chapter 1

Introduction

In current digital era, multimedia material specifically videos can be conveniently disseminated at a very high speed via 4G/5G networks. These videos are widely used in many areas such as entertainment, advertising, pattern recognition, recognition of business activities and social media platforms. Through the use of high-end multimedia devices these videos could be cloned, maneuvered and circulated unfairly quite rapidly with infinitesimal changes in visual fidelity. It could contribute chiefly to video infringement, which has an influence on the fiscal benefits as well as the intellectual property rights of the individual proprietors. Digital watermarking has been used for the defense of the intellectual property rights of the user as well as the credibility of video material throughout the past few decades (Cox et al. 2007). In digital watermarking confidential information, specifically, the watermark is ingrained into the video to shield intellectual property rights as well as the video content integrity. Digital watermarking could deter theft, safeguard content integrity, supervise video dissemination, as well as the transmission.

1.1 Digital Video Watermarking

Digital video watermarking is indeed a method of inserting imperceptible digital watermark (information about either the video or its creator) into video frames that is retrieved to authenticate the ownership or legitimacy of the video itself. Figure 1.1 displays a general video watermarking system. There are three inputs to the watermark embedding algorithm *i.e.* watermark, key and original video to be watermarked. The output, watermarked video, is transmitted to the receiver end. During transmission over the Internet, some deliberate or accidental manipulation may happen to the video, which might distort the video. At the receiver end, the watermark extraction algorithm and key are used to get watermark from the video. The watermark will be retrieved from the video and thereafter it would be contrasted with the original ingrained watermark in order to check the authenticity or legitimacy of the watermarked video.

Requirements: The important requirements of an efficient video watermarking scheme are as follows:

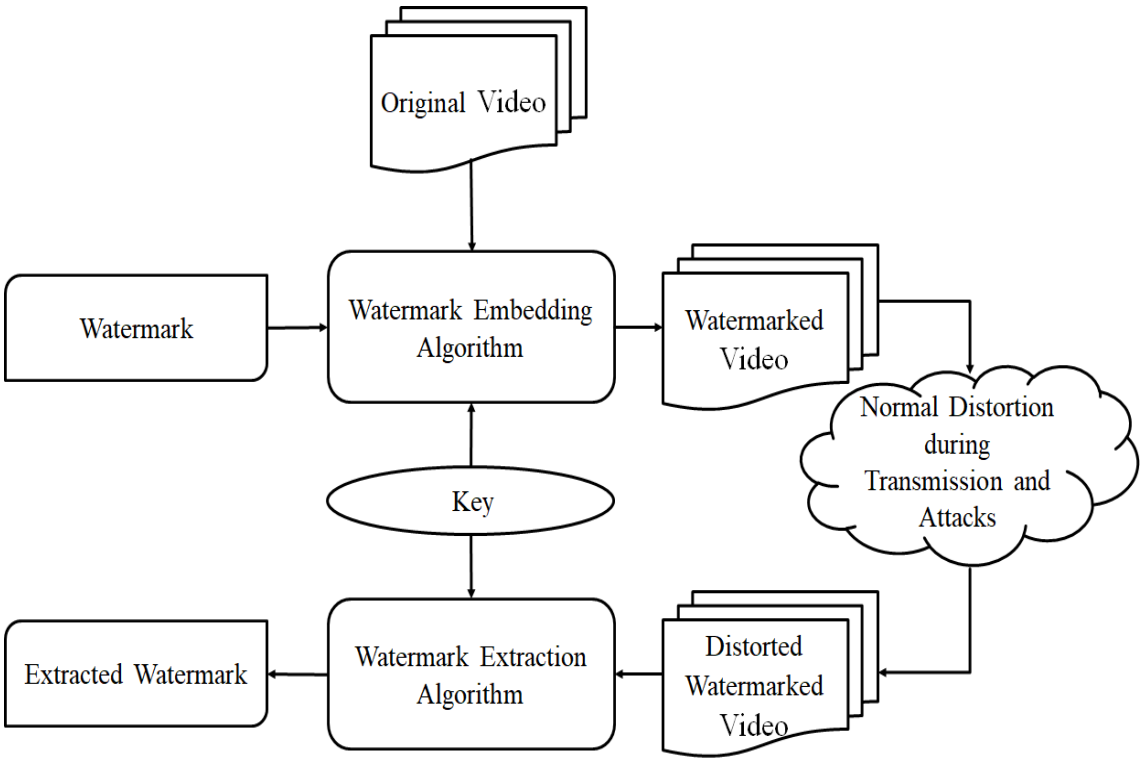


Figure 1.1: General digital video watermarking system

- **Robustness:** The efficiency of the watermarking scheme to withstand the intentional and unintentional attacks is called robustness. A watermarking scheme is called robust, semi-fragile or fragile depending on robustness against different attacks.
- **Imperceptibility:** The embedded watermark should be imperceptible *i.e.*, it should not be visible to the Human Visual System (HVS) while watching the video.
- **Bit Increase Rate:** Video watermarking usually induces an increase in the bit rate of the video. This increase in bit-rate will reckoned through the Bit Increase Rate (BIR), and it should be minimal for an effective watermarking scheme.
- **Blindness:** In a watermarking scheme, if the extraction mechanism does not require the original unwatermarked video, then it will be called as blind.
- **Payload:** It is the size of the watermark specifically the count of the bits ingrained in the video.

1.2 Applications of Video Watermarking

The video watermarking applications are divided into two major categories, video authentication and video protection based on the watermark information embedded into the video.

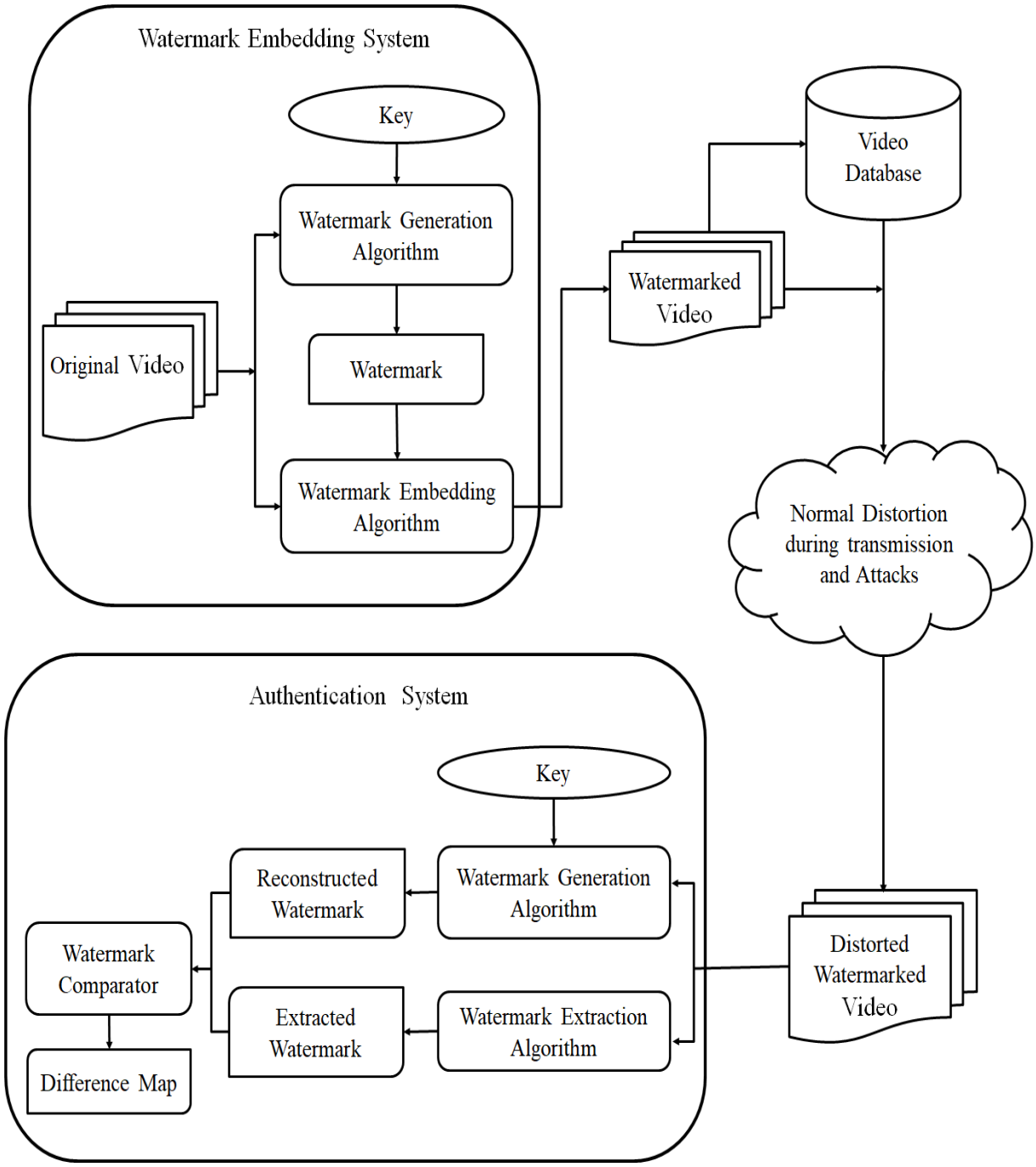


Figure 1.2: Framework of video authentication scheme

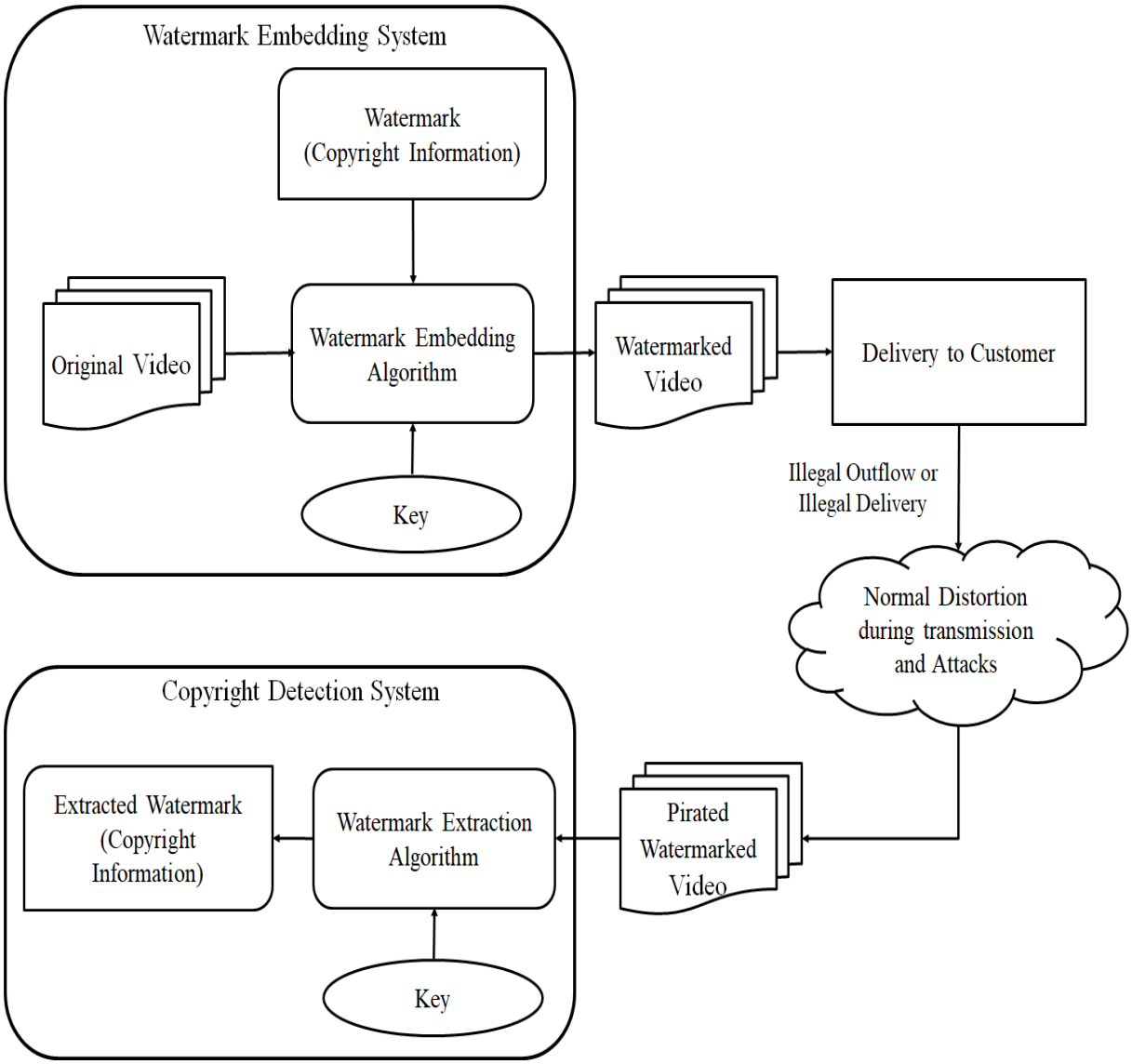


Figure 1.3: Framework of video protection scheme

1.2.1 Video Authentication

The authentication of the video is done to verify its integrity. A general framework appertaining to the video authentication scheme has been revealed through Figure 1.2 (Li 2005). The watermark has been spawned using the key as well as the content of the input video, and then the above watermark is ingrained in the video. Thereafter, the watermarked video is stored in the database or sent to the specific user through the transmission channel. Then the objects in the watermarked videos can be extracted using edge retrieval models (Bayram et al. 2017; Krishnamoorthy and Devi 2013). These objects can be remodeled and can be replaced with the extracted objects. At the specific user side, the watermarked video will be authenticated to detect any changes in the contents of the video during storage or transmission. To authenticate the video, the watermark is reconstructed using the key and watermark generation algorithm. Then the retrieved watermark will be correlated with the reconstructed watermark for creating a difference map. The difference map is used to detect any malicious tampering of video, such as insertion and removal of objects.

1.2.2 Video Protection

The video protection scheme is used to protect the intellectual property rights of the owner and prevent illegal distribution of the video. The framework for the video protection scheme has been illustrated through Figure 1.3. The watermark *i.e.* information about ownership, intellectual property right *etc.*, is embedded in the video using a key and watermark embedding algorithm. This watermarked video is delivered to the customer. Further, to track the customer for any illegal outflow of video, the watermark is extracted using the key and extraction algorithm.

1.3 Classification of Video Watermarking Schemes

Video watermarking schemes can be categorized depending on the embedding domain, extraction, and application specifications. See Figure 1.4 for the various categories of the video watermarking scheme.

1.3.1 Classification Based on Embedding Domain

Watermark can be embedded into the videos, in one of the three distinct domains specifically spatial, transform as well as the compressed domain.

- **Spatial Domain:** In the spatial domain, the watermark is inserted by mutating

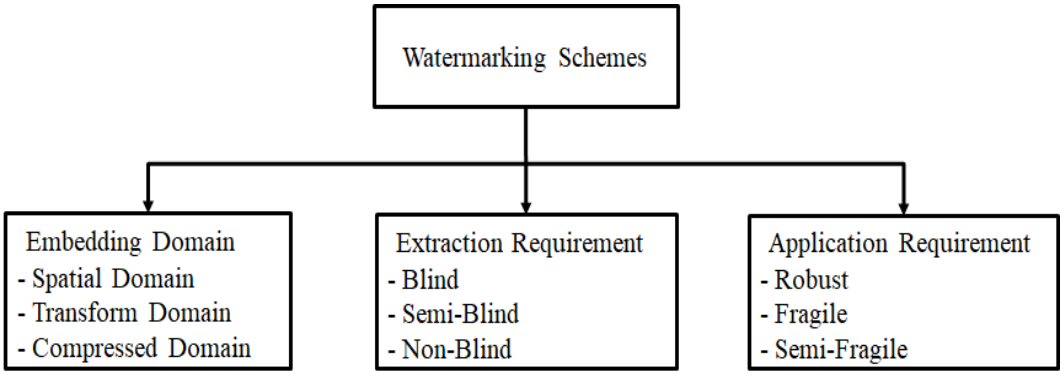


Figure 1.4: Classification of video watermarking schemes

the pixels of a frame. The advantage of using this technique is that it is easy to understand and can be implemented very easily. The significant disadvantage of using this technique is that it cannot handle various deliberate manipulations and data can be dropped amid the video compression mechanism.

- **Transform Domain:** In the transform domain, the pixels of frames are processed using transforms like DCT, Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT) (Peng et al. 2017; Ren et al. 2019; Upadhyay and Dave 2016). The above-mentioned transformed coefficients will be mutated to ingrain the watermark information and then an inverse transform will be employed for generating the watermarked video. This type of scheme can handle various signal processing operations and loss due to multiple processes of encoder during compression than spatial domain schemes.

- **Compressed Domain:** The videos are transmitted and stored in compressed formats like Motion Picture Expert Group (MPEG), H.264, HEVC *etc.* using cloud infrastructure (Niu et al. 2015). Watermarks can be inserted in videos utilising some features of video coding standards, such as block splitting data, intra-prediction modes, transform coefficients, motion vector details etc, *etc.* (Luo et al. 2018).

1.3.2 Classification Based on Extraction Requirement

The watermarking schemes could be segregated within the following relying on requirements during the process of watermark extraction:-

- **Blind:** The watermarking schemes that neither require original video nor the information about the embedded watermark during the extraction process are known as blind schemes.

- **Semi-Blind:** Semi-blind watermarking schemes are those who require information

about the embedded watermark that will be employed while retrieving the watermark at the decoder side. The size of this information will quite small as compared to the original video.

- **Non-Blind:** The schemes that require original video during the extraction of the watermark are known as non-blind watermarking schemes.

1.3.3 Classsification Based on Application Requirement

Depending on the robustness requirement of the application, watermarking schemes could be segregated into three distinct subgroups, namely robust, fragile as well as semi-fragile.

- **Robust:** Robust watermarking schemes can withstand diversified attacks like re-compression, the addition of noise, filtering, *etc.* These schemes are mostly used for the protection of videos.
- **Fragile:** In fragile watermarking schemes, the watermark is undetectable even after a very slight alteration to the watermarked video. The fragile watermarking has been employed for protecting the integrity of content in medical as well as in the legal applications.
- **Semi-Fragile:** Semi-fragile watermarking schemes are robust against content preserving attacks like re-compression, filtering, the addition of noise *etc.* On the other hand, these schemes should be fragile to content changing attacks specifically the addition as well as deletion of objects. These schemes are commonly used for video authentication, where the same representation of the content is not required and the embedded watermark can withstand slight alteration.

1.4 Overview of HEVC Standard

With advances in electronics as well as electrical engineering, more powerful lightweight portable digital systems with multi-core processors and also high-resolution displays are being developed. The above-mentioned gadgets support HD as well as UHD resolution digital videos, however, the existing carriers do not have plentiful bandwidth to transfer enormous volumes of HD multimedia data. This particular deficiency of transmission channels requires a more potent video codec with immense compression and parallel processing potentials. HEVC is the new video coding specification developed in 2013 that embraces parallel processing capabilities and provides 50% greater encoding performance than H.264 / AVC (Sullivan et al. 2012).

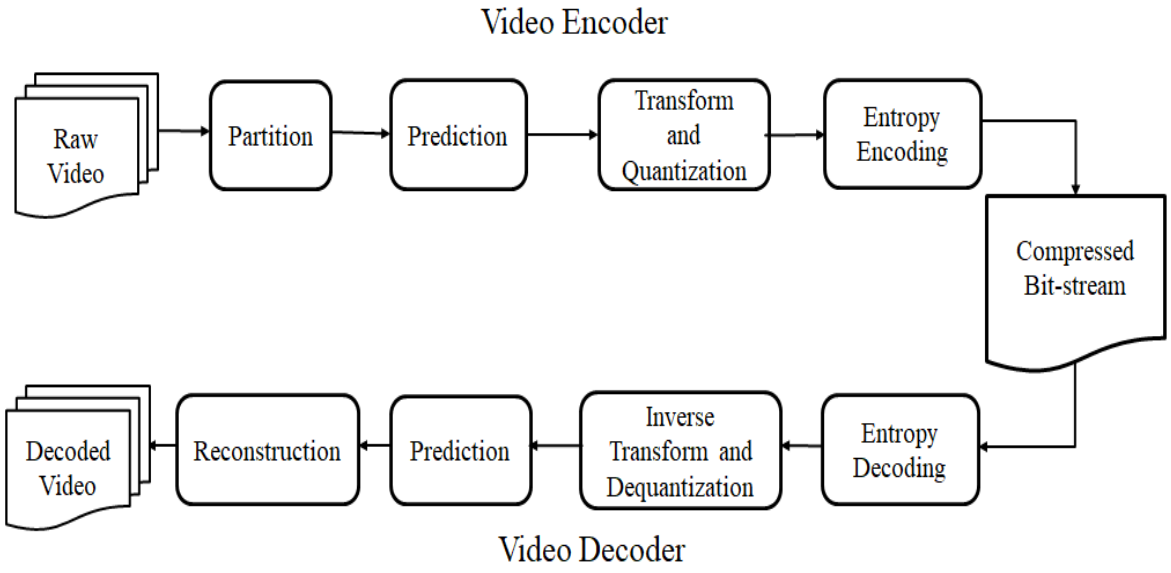


Figure 1.5: Block diagram of HEVC standard

Figure 1.5 shows the block diagram of HEVC standard (Richardson 2014). The raw video, containing a sequence of pictures called frames, is encoded into bit-stream using HEVC encoder. This bit-stream is transmitted or stored. The bit-stream is decoded back into the sequence of frames by a HEVC decoder.

1.4.1 HEVC Encoder

A brief description of each component of the HEVC encoder is given as follows:-

- **Input:** The raw video containing a sequence of frames is given as input to the HEVC encoder.
- **Partition:** Each frame is partitioned into multiple units of simmilar dimension are known as the Coding Tree Units (CTU),has been demonstrated through Figure 1.6. The size of CTU is $M \times M$, where the value M can vary from 64 to 16. The value of M is determined by the external input given in the sequence parameter set, and it is fixed for all the frames in the video. Each CTU contains one luma Coding Tree Block (CTB) of dimension $M \times M$ as well as two Chroma CTB's, one each for Cb and Cr color components of size $M/2 \times M/2$ where the value of M is same as of the size of CTU. In addition to this, each CTU contains information on the partitioning of CTU into Coding Units (CU) called syntax elements. Figure 1.7 shows the structures of CTU and CU, respectively. Similar to CTU, CU also dwells one luma Coding Block (CB) as well as two Chroma CB's. The syntax elements in CU contains the information of prediction and transformation process.

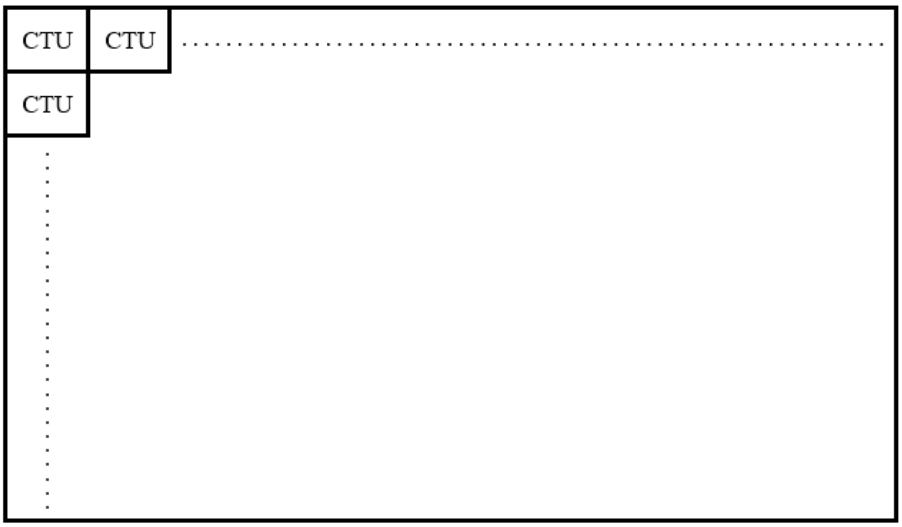


Figure 1.6: Partition of a frame into CTU

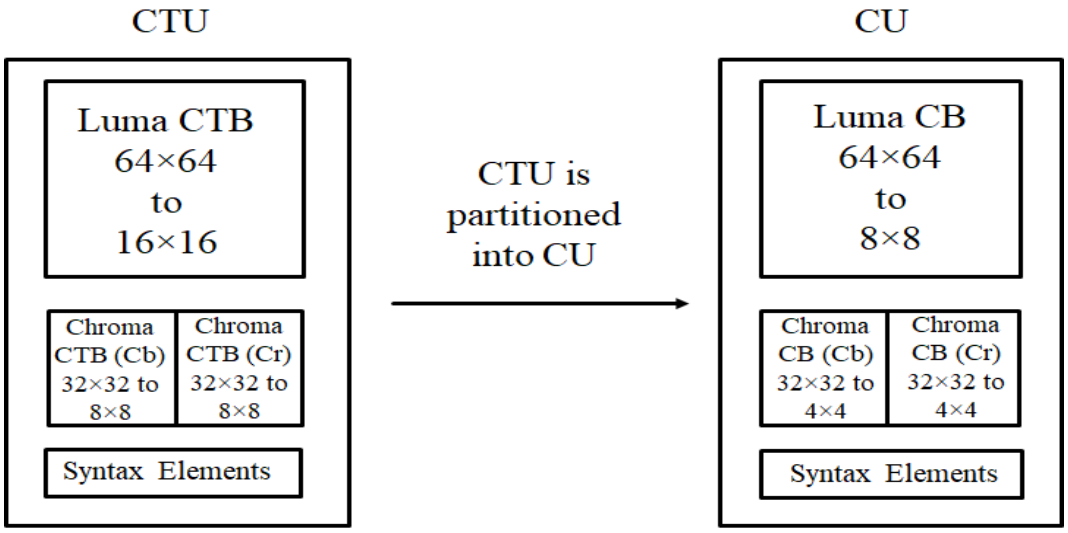
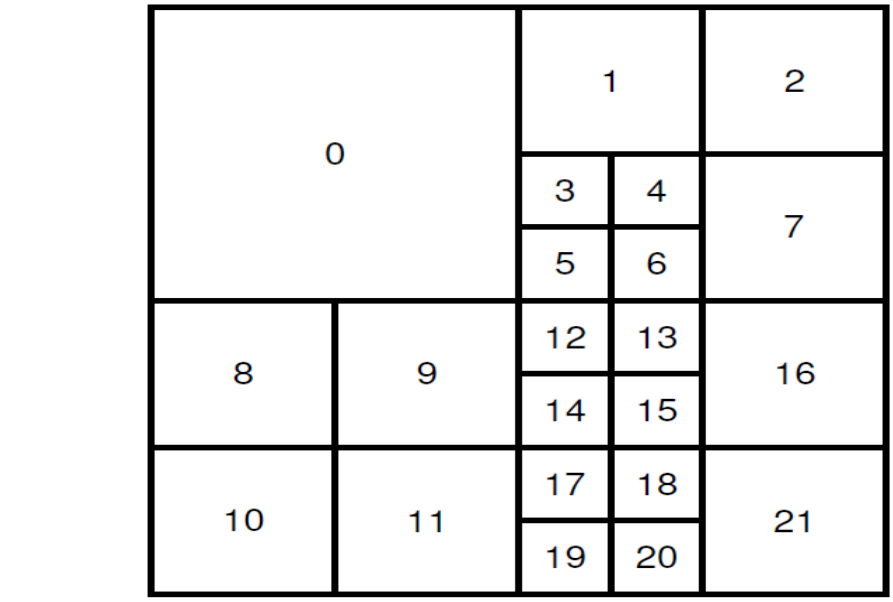
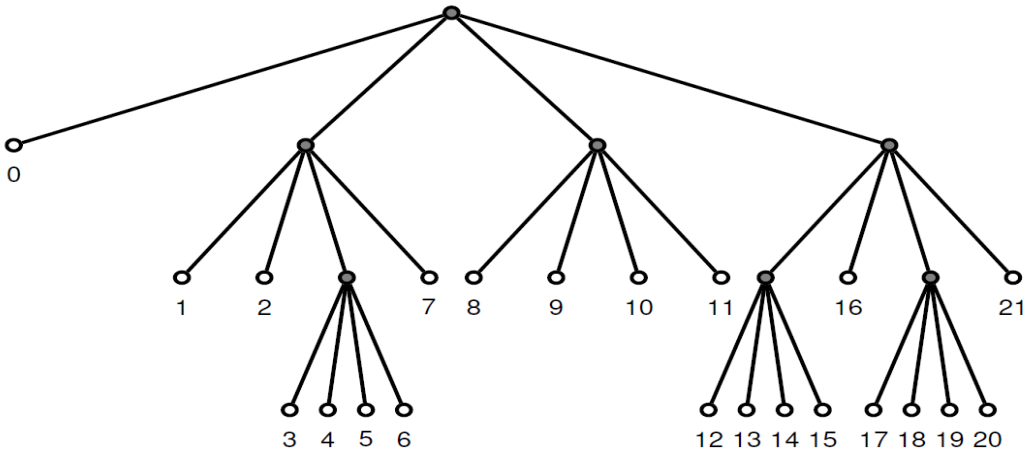


Figure 1.7: Structures of CTU and CU



(a)



(b)

Figure 1.8: Quad-tree partitioning of a CTB

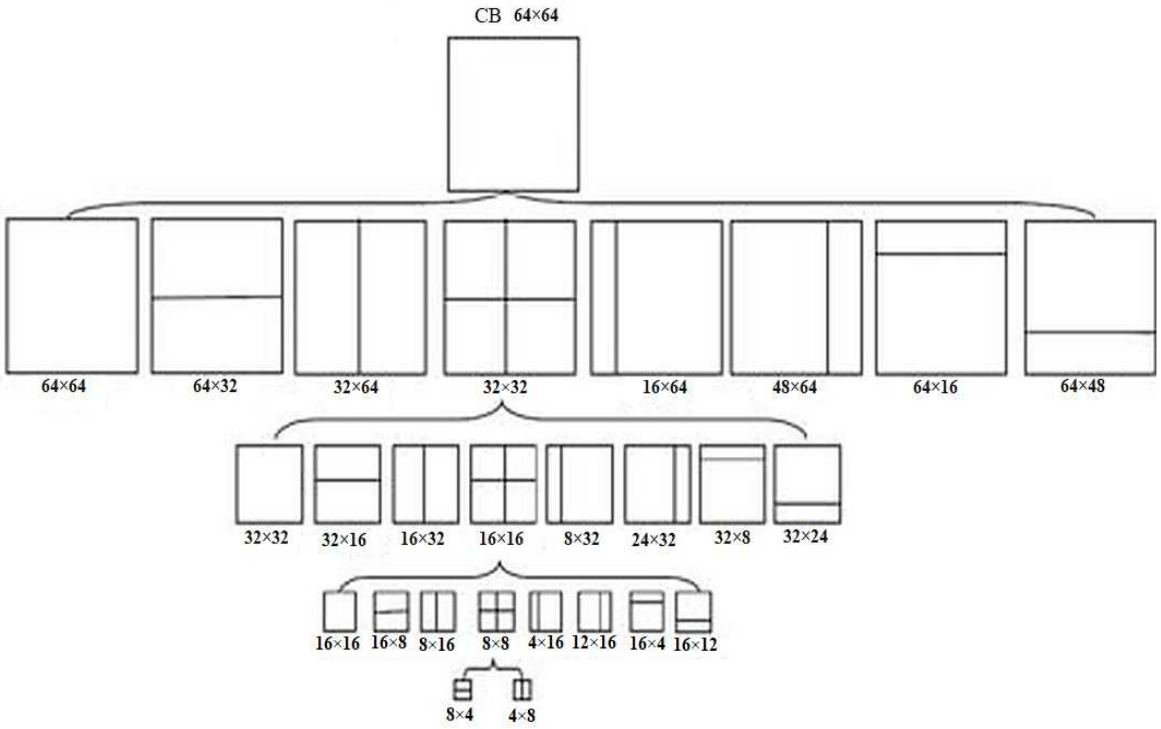


Figure 1.9: Partitioning of CB for inter PB

Figure 1.8a illustrates the partitioning of CTB into a quad-tree structure of multiple CB's (Wien 2015). The root of this quad-tree is CTB, which is recursively divided into CB's, as depicted in Figure 1.8b (Wien 2015). The size of CB's can be the same as CTB, or it can be partitioned recursively into smaller units of size 32×32 to 8×8.

- **Prediction:** The process of prediction utilizes correlation to remove the redundancy present in the video frames. A small part of the frame is highly correlated with all its neighbourhood, and the neighbours can be immediate or quite close by neighbours. The prediction is done by using two processes, namely intra and inter prediction. The intra-prediction mechanism eliminates spatial data attrition inside a frame. By comparison, inter-prediction uses data correlation between two frames to remove temporal redundancy. There are two types of frames in an encoded video based on the type of prediction used *i.e.* Intra frame (I-frame), which utilizes intra prediction and inter frames that are predicted using the combination of both intra and inter prediction. To carry out the prediction process, each CB is divided into Prediction Blocks (PB). The size of intra predicted PB could either be the same as the size of CB, or it can be subdivided into four equal parts. Alternatively, for inter prediction more flexible partitioning size are available for subdividing CB into PB as shown in Figure 1.9 (Podder et al. 2014).

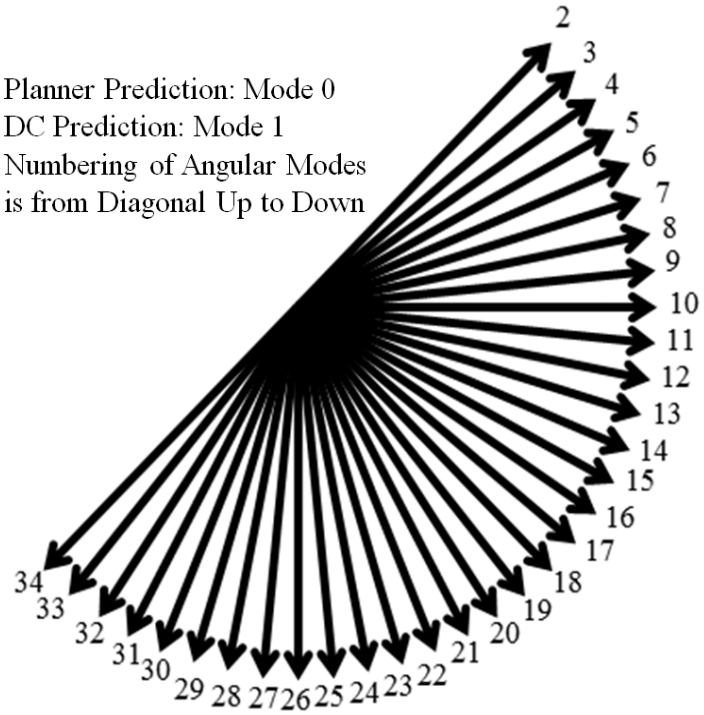


Figure 1.10: Intra prediction modes of HEVC standard

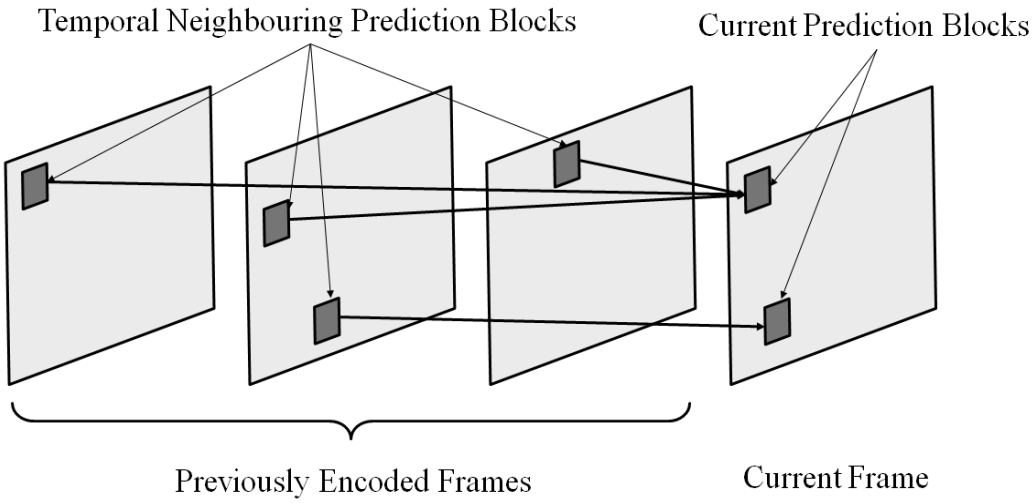


Figure 1.11: Inter prediction in HEVC standard

Intra Prediction: The data of spatial neighbouring PB's are utilized, for generating the data of the current PB using one of the HEVC standard's 35 intra modes. The process of extrapolation is used in angular modes for predicting the current PB from neighbouring PB's. Contrarily, the planar mode uses surface fitting, and the Direct Constant (DC) mode calculates the average value of pixel intensity in the PB. Figure 1.10 shows the intra modes of HEVC standard (Wien 2015).
Inter Prediction: The prediction of the current PB is done from one or two temporal neighbouring PB's. Figure 1.11 serves as an example of the mechanism used in HEVC for inter prediction. The frame which uses only one frame as a reference for prediction is called P-frame, and the frame, which is predicted using two reference frames, is known as B-frame.

- **Transforming and Quantization:** The current PB predicted using intra or inter prediction is subtracted from the original PB, which is given as input to the prediction process. The resulting difference is known as residual error. The residual error further de-correlated by transformed it using either DCT or DST transform. The blocks used for the transformation process are called Transform Blocks (TB). The generated transform coefficients are then quantized. The quantization step removes the information which does not give any visible distortion after reconstructing the frame at the decoder end.
- **Entropy encoding:** The QTC's, prediction modes, *etc.* are coded into the bit-stream by entropy encoding.

1.4.2 HEVC Decoder

HEVC decoder converts the compressed bit-stream into video frames. The decoding mechanism which is exactly opposite to the encoder mechanism is expounded as follows:-

- **Entropy Decoding:** The Syntax elements like QTC's, mode of prediction, direction of intra prediction, *etc.* are parsed back from bit-stream.
- **Dequantization and Inverse Transform:** QTC's are de-quantized, and inverse transformation of the resulting transform coefficients has been performed to re-generate the residual error. The re-generated residual error is not identical to the original residual error due to the lossy process of quantization during the encoding of the video.
- **Prediction:** The prediction each PB is done using the information decoded from the bit-stream. The predicted PB is added to re-generated residual error to get the

CB.

- **Reconstruction:** The merging of the CB's is done to re-generate CTB and reconstruct the decoded frames.

1.4.3 Major Improvements in HEVC Standard

The following are major changes to the HEVC specification considering the H.264 format (Wiegand et al. 2003):-

- HEVC supports very flexible partitioning of CTB's, which ranges from large to small sizes.
- H.264 used only nine modes for intra prediction. This number is increased to 35 in HEVC to enhance the quality of the predicted frame.
- Intra-luma 4×4 TBs use DST rather than DCT.

1.4.4 Enhanced HEVC Standard

HEVC is the current video coding specification that supports the HD as well as the UHD video formats. The compression potency of HEVC is further augmented in order to reduce bandwidth consumption during transmission (Chen et al. 2015). In enhanced HEVC, each frame has been segregated through the CTBs of 256×256. Each CTB is further segregated via employing a quad-tree structure, CB size $64 \times 64$to $8 \times 8$and PB size $64 \times 64$to 4×4. In addition, TB size $32 \times 32$to 4×4uses multiple transformations for the transformation of residual error *i.e.* Enhanced Multiple Transforms (EMT) (Zhao et al. 2016). The inter-prediction process is boosted through using the Overlapped Block Motion Compensation as well as Advanced Temporal Motion Vector Prediction mechanisms.

1.5 Performance Evaluation Metrics

The important metrics used for evaluating and comparing the performance efficiency of the developed video watermarking schemes with existing schemes are as follows:

1. **Correlation coefficient:** The robustness is measured as a coefficient of correlation among the retrieved watermark bits as well as the original ingrained watermark bits as follows (Lin et al. 2001):-

$$\text{Correlation coefficient} = \frac{\sum_{i=1}^{m} W_i \cdot W_i'}{\sqrt{\sum_{i=1}^{m} W_i^2 \cdot \sum_{i=1}^{m} W_i'^2}} \tag{1.1}$$

where, W is the ingrained watermark, $W^{'}$ is the retrieved watermark and m is the number of bits in watermark. The correlation coefficient value lies in [-1,1], and in the case of the retrieved watermark is similar to the ingrained watermark, then its value tends towards 1.

2. ***PSNR:*** Peak Signal to Noise Ratio ($PSNR$) is the ratio of the highest signal power and the power of noise that corrupts the signal affecting the fidelity of its representation. The unit used to express $PSNR$ is logarithmic decibel (dB) because range of signals is very dynamic. The $PSNR$ of watermarked frame is calculated as follows:-

$$PSNR = 10 \times \log_{10} \frac{MAX^2 \times n_1 \times n_2}{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left(F_{ij}^{o} - F_{ij}^{w}\right)^2} \tag{1.2}$$

where MAX is the maximum intensity of a pixel, $n1$ and $n1$ are width and height of each frame in a video, F^o and F^w are original and watermarked video frames respectively.

To calculate the $PSNR_{YUV}$ for YUV video, first the $PSNR$ of luma component ($PSNR_Y$), $PSNR$ of chroma Cb ($PSNR_U$) and $PSNR$ of Chroma Cr ($PSNR_V$) are calculated using (1.2). Then, the $PSNR_{YUV}$ of YUV video has been reckoned through employing the (1.3) provided by Sze et al. (2014). The degradation in $PSNR$ is calculated in terms of the difference in the $PSNR$ of the original and watermarked videos via employing (1.4). The less is the variation among the $PSNR$ of original as well as the watermarked videos, the better will be the quality of the watermarked video.

$$PSNR_{YUV} = \frac{6 \times PSNR_Y + PSNR_U + PSNR_V}{8} \tag{1.3}$$

$$\delta_{PSNR} = PSNR_{YUV} - PSNR^{'}_{YUV} \tag{1.4}$$

where $PSNR_{YUV}$ and $PSNR^{'}_{YUV}$ are $PSNR$ of original and watermarked videos respectively.

3. ***SSIM and VIFp:*** The Structural Similarity Measurement ($SSIM$) and Visual Information Fidelity in pixel domain ($VIFp$) are calculated using VQMT (2015). The original and watermarked videos are given as input to this tool. It compares the spatial and temporal texture of the original and watermarked videos. The out-

put index of this tool lies in [0,1]. The index equal to 1 means the watermarked and original videos are identical, and 0 means the watermarked video is totally changed from the original video.

4. ***BIR:*** The last parameter is the upsurge in video's bit rate after ingraining the watermark. The bit rate of a video is the total number of bit processed in one unit of time. The unit of bit-rate for HEVC standard is Kilobytes per second (Kb/sec). The potency of the new scheme has been assessed specific to the percentage hike in bit-rate of video after embedding a watermark and is denoted by BIR. The BIR of a watermarked video is calculated using (1.5)in which specifically the BR and $BR^{'}$ are bit-rates of original as well as watermarked videos.

$$BIR = \frac{BR' - BR}{BR} \times 100 \tag{1.5}$$

1.6 Summary of The Thesis

The current thesis is divided into six chapters furthermore summaries of all these chapters have been articulated in the following text.

- Chapter 1 introduces the basic concepts of video watermarking and HEVC compression standard. The important metrics used for the evaluation of the efficiency of watermarking schemes and the summary of the thesis are also given in this chapter.
- Chapter 2 briefly outlines the state-of-the-art literature regarding video watermarking schemes. It summarizes the gaps in existing literature for watermarking for HEVC encoded videos and objectives of the thesis. This chapter also explains the methodology used for the watermarking of videos.
- Chapter 3 introduces an authentication scheme for HEVC encoded videos. In this scheme, 4x4 I-frame intra-luma blocks are used to generate and embed authentication code. The spatial exploration of I-frames is used to choose 4×4 blocks most of which are robust to the re-encoding process codec. The spatial exploration has been conducted on the grounds of the non-zero Quantized Discrete Sine Transform (QDST) coefficients available within the block. The chosen blocks are separated into two different subgroups. The first subgroup will be wielded for spawning authentication code, as well as the second one will be utilized for inserting the spawned code. The authentication code has been created by employing the relation among the number of Positive Quantized Discrete Sine Transformed (PQDST) and Negative Quantized Discrete Sine Transformed (NQDST) coefficients within a block.

The created authentication code has been implanted through adjusting the magnitude of the QDST coefficients. At the decoder side, the embedded code is retrieved and correlated to the restored code to validate the authenticity of the video.

- In Chapter 4, a drift free I-frame based protection scheme is proposed for enhanced HEVC. This standard employs several transforms thatcan effectively de-correlate residual error in I-frame blocks. Most of the employed transforms are not symmetric, consequently, theseshould not be employed explicitly while the ingraining watermark. So, the symmetric transforms are used in this scheme to integrate the watermark into intra luma blocks with dimensions 4×4. To ingrain the watermark, the relation among the intensities of pair of transform coefficients in a block has been wielded.Additionally, such a couplet of coefficients has been chosen which do not disseminate the accumulated distortion to the nearby blocks during the standard's intra-prediction process. The ingrained watermark can be removed from the compressed bitstream by decoding only QTC's.
- In chapter 5, a robust P-frame based protection scheme is proposed for enhanced HEVC standard for LD configuration.In LD configuration only the intial one frame is I-frame whereas remaining all are P-frames. Currently, this configuration is the most important video coding structure, which is extensively used for the compression of videos used in applications like a surveillance system, conferencing system, *etc.* The recommended P-frame protection scheme has been grounded on the multiple transforms used by enhanced HEVC to pack residual error into fewer transform coefficients efficiently. Firstly, either intra or inter luma blocks with dimension 4×4 will be culled through employing a robustness threshold. The threshold will be computed wielding the count of non-zero QTC with a magnitude greater than one. The watermark is then inserted in the chosen blocks by altering the magnitude of only non-zero transform coefficient pairs.
- Chapter 6 summarises the key findings of the study and outlines future strategies for research.

Chapter 2

Literature Survey

In this chapter, the efficiency of the existing video watermarking schemes is analyzed for the authentication and protection of HEVC encoded videos. The analysis of the authentication scheme is done based on robustness against the re-compression, detection, and localization of the malicious tampering. Section 2.1 briefly describes the existing authentication schemes in the literature. The Protection schemes are examined for robustness against re-compression, visual imperceptibility and impact on bit-rate with regard to enhanced HEVC. The current protection methods are outlined in section 2.2.

2.1 Authentication Schemes

In this section, various authentication schemes based on different video compression standards, namely MPEG, H.264/AVC, and HEVC, are discussed. Mobasseri et al. (2000) devised an authentication system appertaining to the uncompressed domain where authentication code specifically the watermark matrix has been ingrained in the raw video by manipulating bit plane levels. It can detect the tampering of video and localize the cuts and splices done in length and duration. The main drawback is that the scheme is not complaint to any video coding standard and needs to be reversed before compression of the video. To circumvent this drawback, Cross and Mobasseri (2002) proposed an authentication scheme in the compressed domain for videos encoded using the MPEG-2 standard. The authentication code is embedded by altering the Variable-Length Codes in the MPEG-2 encoded bit-stream. The afore-mentioned scheme can localize the tampered location at the GOP level but is fragile to re-encoding of the bit-stream. Later, He et al. (2004) proposed semi-fragile authentication scheme in un-compressed domain. Firstly, for spawning the authentication code from the content of the video, error correction coding and cryptographic hashing is applied to the selected set of angular radial transformation coefficients. Then the generated authentication code is embedded into the randomly selected DFT coefficients before encoding using MPEG-4. The proposed scheme is resilient towards the encoding using video compression standard MPEG-4 as well as protects the integrity of the video.

Table 2.1: Summary of existing authentication schemes

Existing Schemes	Compression Standard	Syntax Elements	Fragile/Semi-fragile	Localization
Mobasseri et al.(2000)	Un-compressed	Bit Plane	Fragile	Yes
Cross and Mobasseri (2002)	MPEG-2	VLC	Fragile	Yes
He et al. (2004)	Un-compressed	DFT	Semi-fragile	Yes
Profrock et al. (2005)	H.264	Bitstream	Fragile	No
Zhang and Ho (2006)	H.264	Parition modes	Fragile	No
Mobasseri and Raikar (2007)	H.264	CAVLC	Fragile	No
Chen et al. (2008)	H.264	QTC's	Fragile	No
Xu et al. (2011)	H.264	QTC's	Semi-fragile	yes
Wang et al. (2015)	HEVC	Intra Prediction Mode	Fragile	No
Farfoura et al. (2016)	H.264	QTC's	Semi-fragile	yes
Tew et al. (2016)	HEVC	Motion vector/QTC's	Fragile	Yes
Tew et al. (2018)	HEVC	Motion vector/QTC's	Fragile	Yes

Table 2.2: Summary of existing protection schemes

Existing Schemes	Compression Standard	Syntax Elements	Frame Type	Efficiency With respect to Enhanced HEVC		
				Robutness	Imperceptibility	BIR
Hartung and Girod (1996)	Uncompressed/ MPEG	Pixels/ Tranform Coefficients	NA	Low	Low	Low
Cox et al. (1997)	Uncompressed	Pixels	NA	Low	Low	Low
Hsu and Wu (1998)	MPEG	Tranform Coefficients	NA	Low	Low	Low
Langelaar and Lagendijk (2001)	MPEG	Tranform Coefficients	Intra	Low	Low	Low
Alattar et al. (2003)	MPEG	Tranform Coefficients	Intra/Inter	Low	Low	Low
Prfrock et al. (2006)	Uncompressed	NCG	NA	Low	Low	Low
Noorkami and Mersereau (2007)	H.264	QTC's	Intra	Low	Low	Low
Noorkami and Mersereau (2008)	H.264	QTC's	Inter	Low	Low	Low
Zou and Bloom (2009)	H.265	CAVLC	Intra/Inter	Low	Low	Low
Zou and Bloom (2010)	H.264	CABAC	Intra/Inter	Low	Low	Low
Dutta et al. (2013)	H.264	QTC's	Inter	Low	Low	Low
Chang et al. (2014)	HEVC	QTC's	Intra	Moderate	Moderate	Moderate
Swati et al. (2014)	HEVC	QTC's	Intra/Inter	Low	Moderate	Moderate
Gaj et al. (2015)	HEVC	QTC's	Intra	Moderate	Low	Low

Table 2.2: Summary of existing protection schemes

Existing Schemes	Compression Standard	Syntax Elements	Frame Type	Efficiency With respect to Enhanced HEVC		
				Robutness	Imperceptibility	BIR
Ogawa and Ohtake (2015)	HEVC	QTC's	Intra	Low	Moderate	Moderate
Van et al. (2015)	HEVC	Motion Vector /QTC	Inter	Low	Moderate	Moderate
Elrowayati et al. (2016)	HEVC	QTC's	Intra	Low	Moderate	Moderate
Dutta and Gupta (2016)	HEVC	QTC's	Intra	Moderate	Moderate	Moderate
Dutta and Gupta (2017)	HEVC	QTC's	Inter	Moderate	Moderate	Moderate
Gaj et al. (2017)	HEVC	Tranform Coefficients	Intra	Moderate	Moderate	Moderate
Liu et al. (2018)	HEVC	QTC's	Intra	Low	Moderate	Moderate
Shanableh (2018)	HEVC	Split Decision of blocks	Intra/Inter	Low	High	Moderate
Yang and Li (2018)	HEVC	Motion Vector	Inter	Low	Moderate	Moderate
Liu et al. (2019)	HEVC	QTC's	Intra	Moderate	Moderate	Moderate
Li et al. (2019)	HEVC	Split Decision of blocks	Inter	Low	High	Moderate
Shanableh (2019)	HEVC	Split Decision of blocks	Intra/Inter	Low	High	Moderate
Yang et.al. (2019b)	HEVC	Split Decision of blocks	Inter	Low	High	Moderate
Xu (2019)	HEVC	Rice Parameters	Intra/Inter	Low	Moderate	Moderate
Konyar et.al. (2020)	HEVC	QTC's	Intra	Low	Moderate	Moderate

Pröfrock et al. (2005) proposed reversible fragile authentication scheme for H.264 encoded bit-stream. The authentication code is generated using an encryption function, public and private keys. Then the skipped macroblocks are selected from parsed bit-stream to embed the authentication code. The Macroblocks are chosen effectively to reduce deterioration of visual quality and minimum impact on bit-rate. Zhang and Ho (2006) suggested a fragile authentication mechanism centered on the H.264 compression format inter-prediction process. In this scheme, an authentication code is embedded during the motion estimation and compensation process of P and B slices. The scheme has an indistinguishable loss in the optical quality, a moderate increase in bit rate, but it is vulnerable towards transcoding and does not provide localization of the attack. Mobasseri and Raikar (2007) embedded reversible fragile watermark in unused code space of CAVLC codes directly into bit-stream. The scheme has no increase in bit rate and visually imperceptible as only unused code space is used for embedding the watermark. The scheme cannot withstand re-compression with various QP values as CAVLC codes change during the re-compression process.

Further, Chen et al. (2008) developed a semi-fragile scheme based on sub-band indices of 4×4 blocks. The sub-band index is calculated based on the non-zero AC coefficients after quantization. Then based on the sub-band index, high-frequency AC coefficients are adjusted to add the authentication code. The scheme is robustness against transcoding processes and can localize the tampered area but has high degradation in PSNR. Later Xu et al. (2011), designed a semi-fragile system in which initially authentication code is produced out of the characteristics of the video frame and the generated authentication code is incorporated in the 4x4 blocks diagonal AC coefficients.This system can be used to identify and locate the affected area, but the localization algorithm has low precision. Additionally, this scheme suffers from security issues as diagonal Ac coefficients are used embedding authentication code. To combat these problems, Farfoura et al. (2016) proposed a secure semi-fragile authentication scheme in which the generated authentication code is encrypted through employing a cryptographic function before embedding in the video. Further, to improve the visual imperceptibility and increase the robustness against transcoding, blocks selected for embedding the authenticate code are selected using spatial analysis. Moreover, to increase the precision of the tamper detection and localization 3×3 median filter is used to remove random noise elements. The authentication scheme of Farfoura et al. (2016) is based on the H.264 compression standard and does not provide adequate results for the HEVC standard owing to differences in the structure of video coding standards.

There are a few authentication methods for the HEVC standard in current literature.

The limitations of the existing schemes are explained in the following text. In Wang et al. (2015), the intra-prediction modes employed in the luma category blocks with the dimension of 4×4 are considered to injecting the authentication code during the video encoding procedure. The angle values are allocated to all the corresponding 33 angular modes involved in the intra-prediction mechanism. The gap in the angle of two consecutive blocks has been altered to ingrain the watermark bit via swapping the best optimum mode with the sub-optimal mode through employing the mapping of angle ratios. This system is not resilient to the re-encoding of the video since, during this process, the sub-optimum mode is substituted with the best optimum mode mostly through the encoder (Kim et al. 2010). Tew et al. (2016) suggested a scheme in which the authentication code has been created exploiting the syntax components such as the CU category as well as the CU prediction method. The created code is repetitively integrated into the video through the motion vectors, the QP of each CTU together with the least significant bit (LSB's) of the quantized DCT transformed coefficients. These codes are recovered at the decoder side as well as analysed to detect the locations that have been tampered while the transmission.

Tew et al. (2018) has designed an authentication mechanism for HEVC encoded videos. In this system, the syntax elements are separated into two distinctive collections, one collection is wielded for the substantiation process, at the same time the other one is wielded for encryption. The size of the CU within a slice has been adopted to spawn the authentication code, that could later be repetitively inserted in the LSBs of the quantized DCT transformed coefficients, the QP of individual CTU, as well as the motion trajectories of the individual slice. This code is retrieved by the decoder, to recognize the tampered areas. The schemes of Tew et al. (2016, 2018) were based on the LSB's quantized DCT transformed coefficients, QP, as well as motion trajectories. The amounts of such syntax components are tweaked through the encoder as the video is re-compressed by the codec. Table 2.1 summarises the characteristics of the existing authentication schemes.

2.2 Protection Schemes

The following text provides a brief survey of video protection schemes. Hartung and Girod (1996) developed an algorithm for insertion of watermark *i.e.*, modulated signal into the uncompressed video by using a threshold value so that it cannot be identified *i.e.*, it is visually undetectable. In the aforementioned scheme, the second watermark has been ingrained in the compressed domain in which the first DCT is applied to modulated signal *i.e.*, watermark, and then it is inserted into transform coefficients frames of video. In this

addition of the arithmetic mean of the modulated signal and the arithmetic mean of coefficients is carried out. Cox et al. (1997) inserted a watermark signal by incorporating spread spectrum technique into spectral portions, which have the highest repercussion on the optical quality of multimedia. The implementation of schemes of Hartung and Girod (1996) and Cox et al. (1997) is straightforward, but the watermark inserted with in uncompressed video is mostly dropped during video compression operation.

Hsu and Wu (1998) inserted watermark bits into MPEG compressed intra and inter frames of video, by altering middle-frequency coefficients spawned through the DCT with indistinguishable degradation in the optical quality of the video. Langelaar and Lagendijk (2001) developed a watermarking scheme that utilizes n blocks for inserting one bit of watermark. The watermark has been inserted within the JPEG / MPEG bit-streams through eliminating the high-frequency coefficients spawned via DCT in specific image areas. In (Alattar et al. 2003), the spatial spread spectrum watermark is inserted withing the MPEG-4 bit-stream via adjusting the DCT coefficients. This system has enhanced visual transparency and tolerance to attacks relative to previous systems.

Pröfrock et al. (2006) introduced an uncompressed domain watermarking system for H.264. The watermark has been integrated within the relevant part of the video frames in this scheme, which are not removed by video compressed standard while eliminating redundant data. For this new Normed Centre of Gravity (NCG) technique is recommended for describing the spatial borders of the objects. The NCG is affected by the lossy process of the H.264 compression standard, so to overcome this problem, a robustness threshold is used for embedding the watermark. Still, owing to the restriction of communication channels, most videos are processed and distributed in the compressed domain, and for watermarking in uncompressed domain first complete decoding of the video is required, which is to be watermarked. To overcome this problem, Noorkami and Mersereau (2007) had already introduced a watermarking mechanism considering the H.264 compression format. This scheme integrates the watermark within the quantized DCT coefficients of blocks with the dimension 4×4 available in the I-frames, which have been chosen by HVS masking criteria. The scheme was efficient for H.264 standard in terms of imperceptible deterioration in visual quality, marginal increase in bit rate and attack resilience. Later, Noorkami and Mersereau (2008) explored the watermarking capacity of inter frames of H.264. In the aforementioned system, the watermark was implanted into P-frames through modulating the quantized DCT coefficients of inter and intra block. To overcome the problem of an increase in bit-rate only non-zero AC coefficients were used for embedding watermark. Zou and Bloom (2009, 2010) proposed new schemes with low computational complexity, in which watermarks are incorporated in entropy-encoded bit-

stream syntax components such as CAVLC and CABAC. These schemes are not resilient to re-compression, and there is a significant increase in bit-rate. Besides, Dutta et al. (2013) recommended a watermarking technique wherein the watermark has been infused by altering the quantized AC coefficients of the block present within the inter frames precisely within the H.264 encoded bit-stream. The appropriate blocks are selected using a robustness threshold, and the only magnitude of the coefficients have been altered for ingraining the watermark, so this scheme achieves desirable robustness against various attacks, minimal decrease in ocular quality, as well as the increase in bit-rate.

The watermarking schemes appertaining to MPEG and H.264 could not be extended to HEVC owing to encoding variations in specifications. The efficacy of the latest literature has been investigated in the following paragraphs considering the improved HEVC. In Chang et al. (2014), the watermark is inserted through adjusting DCT together with DST coefficients on the principle of their properties without error spread into adjoining blocks. The aforementioned one has a significant embedding potential for the watermark but is not resilient towards re-compression. Swati et al. (2014) recommended a strategy that would insert a watermark via QTC's LSB's mutation. This scheme has a nominal deterioration in the visual output of the video and a marginal upsurge in the bit rate since particularly the LSB's of the QTC's are adjusted. The key drawback concerning the scheme is that this is vulnerable against re-compression due to the change in the value of the LSBs after re-compression of the video. Gaj et al. (2015) introduced a scheme in which watermarks are inserted by adjusting non-zero coefficients count within the 4x4 intra-luma block. To accomplish this the coefficients acquiring zero intensity are shifted to non-zero, resulting in a substantial upsurge in bit rate together with a drift error. This drift error is perpetuated to bordering blocks and contributes to visual artifacts.

Van et al. (2015) inserted watermark in the inter frames by altering motion vector components and LSB's of last non-zero quantized AC coefficient. The scheme has a moderate increase in bit rate as selected motion components are modified, and only LBS's are altered for embedding watermark. Further, this scheme has average imperceptibility but is not robust against transcoding as the re-encoding process changes the motion vectors and LSB's of coefficients. In Elrowayati et al. (2016), the watermark has been implanted on the grounds appertaining to non-zero QTC's parity. However, the previously mentioned system is resilient to noise attacks. Moreover, it allows the HEVC codec to detect as well as correct errors induced by transmission over noisy networks. Still, it is vulnerable to the re-encoding attack since the parities of QTC's alter subsequently the re-encoding the video. Ogawa and Ohtake (2015) inserted watermark by modifying the specific component of the QTC. This method seems to have very low visual quality degradation however

it is not resilient toward re-compression considering the value of the QTC shifts subsequently after the re-compression. The framework of Dutta and Gupta (2016) inserted watermark in the very first two non-zero DCT coefficients of luma category blocks with a dimension of 4×4 within the I-frame. In Dutta and Gupta (2017), the embedding of the watermark is accomplished by altering the very first two non-zero DCT coefficients of luma category blocks with a dimension of 4×4 both intra together with inter-luma blocks within P frames. The above two approaches are resilient against a re-compression attack, however the main drawback of both these methods is that these employs only the properties of the DCT coefficients.

Gaj et al. (2017) presented an approach in which the watermark is inserted in the DST spawned coefficients through varying the chosen group of coefficients in order to prevent the issue of the distortion dissemination towards the nearby block. The previously mentioned method has a low bit rate increment, an unnoticeable deterioration in visual fidelity, and is resilient toward re-compression.Liu et al. (2018) suggested an I-frame watermarking framework relying on the DST coefficients of the 4×4 luma blocks. This scheme has an imperceptible decrease in ocular quality and a retrenched upsurge in bit-rate. The key drawback in the Gaj et al. (2017) and Liu et al. (2018) is that both approaches are centered on the features of the DST transformed coefficients. Shanableh (2018) embedded watermark through modifying the splitting decisions of $32 \times 32$and 16×16 blocks. This method does have a large capacity for inserting watermark with good visual quality but also is fragile for re-compression. Yang and Li (2018) utilized the inter frames for embedding watermark by altering motion vector components of selected blocks. This scheme has a slight increase in bit-rate but has an average visual quality owing to modifications in motion vectors of blocks. The above-mentioned scheme is vulnerable towards the re-encoding attacks since the motion trajectories are changed throughout the re-encoding with different parameters. The information is embedded by adjusting the parity of the specific subset of DST coefficients largely dependent upon intra prediction mode of the 4×4 intra-luma block in Liu et al. (2019). The method has a high embedding capacity with a somewhat acceptable degradation in visual quality, but this method focuses primarily on the characteristics of the DST coefficients.

Li et al. (2019) inserted watermark by modifying the size of either the 8×8 or 16×16 blocks. There seems to be an unnoticeable alteration in the visual fidelity of the video, and it contributes to a slight increase in bit-rate. Shanableh (2019) improved the scheme in Shanableh (2018), by using only 16×16 for embedding watermark. This scheme has better visual quality with minimum changes in the splits decisions of the blocks. This results in a retrenched upsurge in the bit rate of the video. In Yang et al. (2019), watermark

embedding is done by altering the size of 64×64 to 8×8 blocks. The block size has been selected on the basis of the QP used by the encoder. The suggested technique has very high embedding potential and very low visual quality distortions. The key drawback with these systems is that they are not resilient to re-compression, since the split decision of blocks varies during re-compression. Xu (2019) has suggested a protection scheme in which watermark embedding is paired with the encryption process. The watermark is inserted by changing the QTC 's rice parameters. This scheme has a high visual quality and relatively small bit-rate increase, but is vulnerable to re-compression. Konyar et al. (2020) suggested a matrix-based scheme for the embedding of a watermark in DST coefficients within the blocks with dimension 4×4 throughout the I-frames. This new scheme has an indistinguishable decline in optical quality as well as a retrenched upsurge in the bit-rate. The imperceptibility of the scheme is obtained by choosing the DST coefficients such that the error is not transmitted to the neigbhouring blocks. The main limitations of the scheme are that its masking parameter is calculated on the basis of DST coefficients only. The summary of the protection schemes is given in Table 2.2 with respect to compression format, embedding syntax elements, type of video frame, and efficiency of the scheme for enhanced HEVC.

2.3 Gaps in Literature

The gaps in the existing literature are as follows:-

Gaps in Video Authentication Schemes: The main limitation of the existing Wang et al. (2015) and Tew et al. (2016, 2018) schemes is that the syntax elements used to embed the watermark are fragile to the HEVC encoder's re-compression process. In the current literature, therefore, there is no semi-fragile technique that is robust against information preserving modification such as re-compression, but fragile to information changing alterations.

Gaps in Video Protection Schemes: There are several watermarking schemes for protection of HEVC encoded videos in state-of-art literature. There are many drawbacks in the existing schemes, and based on these drawbacks, these schemes are categorized into two sets. In the first set, the inadequacy of the schemes for enhanced HEVC is due to the syntax elements used for the embedding of the watermark. The schemes in the first set are (Chang et al. 2014; Elrowayati et al. 2016; Li et al. 2019; Ogawa and Ohtake 2015; Shanableh 2018, 2019; Swati et al. 2014; Van et al. 2015; Xu 2019; Yang and Li 2018; Yang et al. 2019). The syntax elements, namely LSB's and rice parameters of the QTC's, motion vector information, and size of the blocks, are altered during the re-encoding of

the videos with different QP values. As a result, the embedded watermark cannot be retrieved after re-encoding the video, *i.e.* these schemes are not resilient against the re-encoding process of the codec.

In the second set, the limitations are due to the competence of the existing schemes for enhanced HEVC. The main reason for the inadequacy of the existing schemes is the introduction of new features in the main modules of the compression standard. The key feature of the enhance HEVC is EMT is being used to effectively de-correct residual block error in all the frames specifically I as well as P. The process of EMT implements different kernels of DCT and DST transforms to compact the maximum energy in the minimum quantity of coefficients per block. The coefficients of these transforms have different characteristics in terms of stability and amount of energy contained in each coefficient. The main limitation of the schemes of Dutta and Gupta (2016, 2017) is that these are grounded on the characteristics of the DCT transformed coefficients. The stability of the first two non-zero AC coefficients of DCT is used for embedding watermark. The principal problem of these schemes is that DST transformed coefficients do not have the same stability property. Hence the embedded watermark does not survive during the re-compression attacks. The schemes in (Gaj et al. 2017, 2015; Konyar et al. 2020; Liu et al. 2019, 2018) used the features of the DST transformed coefficients for embedding watermark in HEVC encoded videos. The main feature of the DST transform is that it can efficiently compact energy of the residual errors generated at the boundary of the block. The concentration of these errors at a specific edge of the block would be attributed mostly to angular intra-prediction mode that has beenused for this block. Subsequently, schemes based on this property of DST are not optimal for embedding watermark in the DCT transform coefficients. It can therefore be inferred that the performance of the protection scheme is dependent on the properties of the transforms being used by block. Therefore, new schemes should be designed for enhanced HEVC depending on the characteristics of the both DCT and DST.

2.4 Objectives

Focusing over the gaps revealed through the precursorysection, the main research objectives are summarised as follows:

- To develop efficient watermarking scheme(s) for authentication of HEVC video coding standard in compressed domain.
- To develop efficient watermarking scheme(s) for protection of HEVC video coding standard in compressed domain.

- To verify and validate developed schemes.

2.5 Methodology

In context to the objectives presented in Section 2.4, the following methodology has been adopted for developing, verifying and validating the authentication and protection schemes.

- **Phase I:** In this phase, the existing authentication and protection schemes are studied and analyzed.
- **Phase II:** The efficiency of the authentication and protection schemes depends on the selection of the syntax components as well as the stage of the encoding process that have been employed for ingraining the watermark.

  Syntax Elements: Watermark is embedded in the syntax elements which are not altered by the re-compression of the video. So, the syntax elements like magnitudes of the transform coefficients and QTC's, which are not changed by the re-compression process of the encoder would be utilized for integration of the watermark.

  Stages: The imperceptibility and BIR appertaining to a scheme are dependent on the stage of the encoder used for embedding the watermark. There is a trade-off between the stage of the encoder used for watermark ingraining as well as the computational intricacy concerning to the scheme. Higher the stage of encoder lesser is the intricacy particularly, embedding watermark in the entropy coding stage takes less computational time. In contrast, the partitioning stage takes more computational time for embedding watermark. The entropy coding stage has a higher increase in BIR and degradation in visual quality. So, this stage is rarely used for watermark embedding. The output of the transformation and quantization steps are used for embedding the watermark in the protection and authentication schemes, respectively, to obtain the desired results.
- **Phase III:** The authentication and protection schemes should have invisible degradation in visual quality and retrenched increase in bit-rate. Protection as well as the authentication schemes should be resilient to transcoding techniques such as re-encoding through varied QP values and video processing attacks. In addition to this, the authentication scheme must also be fragile against malicious activities, in other words, this should detect the removal and addition of objects. The effective authentication scheme should always be semi-fragile, specifically resilient towards the feature maintaining manipulations, on the other hand, susceptible to content modulating manipulations. The performance evaluation and verification of the pro-

posed schemes are done with the existing schemes by conducting experiments on videos with different textures and resolutions (Derf,Digiturk).

Chapter 3

An Authentication Scheme for HEVC

This chapter introduces a new, semi-fragile authentication scheme for HEVC encoded videos to address the shortcomings of current literature discussed in the previous chapter[1]. The scheme is designed by taking into account the newly introduced features of HEVC standard. In HEVC, the 4×4 intra-luma blocks used DST for effectively compacting residuals into transforming coefficients. The proposed scheme exploits properties of DST coefficients for embedding semi-fragile watermark into HEVC encoded videos. Firstly, the main limitations of the existing literature are discussed in Section 3.1. The scheme was further illustrated in Section 3.2 and the experimental findings are compiled in section 3.3. The inference of the chapter is ultimately set out in Section 3.4.

3.1 Introduction

As discussed through precursory chapter, the research issue premeditated in this work is based on the difficulties posed by the use of existing authentication schemes for HEVC encoded videos in terms of robustness against attacks that preserve information, namely re-encoding by diverse QP parameter. The performance of current schemes is assessed in the following paragraph. In Chang et al. (2014), the watermark bits have been ingrained in the video through the use of QDST coefficient in the intra-luma blocks of dimension 4×4 using a drift rectification method for avoiding the error spread across blocks. Although the aforementioned scheme provides a wide potential for embedding along with satisfactory ocular quality, it is not effective for non-malicious video processing modification since the choice of 4×4 blocks for implanting watermarks is rendered with no robustness study. Wang et al. (2015) inserted watermark through adjusting the 4×4 luma blocks angular intra-prediction directions. To do this, predefined angles are assigned to each angular mode. Then, the difference in the angles between the two successive blocks is calculated, and this difference is altered according to the mapping relationships of angle values. To modify the difference, the prediction mode chosen by the encoder is replaced by the sub-optimal mode. The drawback of the scheme is that this alteration in the mode is

[1]The contents of this chapter are published in Gagandeep Kaur, Singara Singh Kasana and M. K. Sharma. "An efficient authentication scheme for high efficiency video coding/H.265" *Multimedia Tools and Applications*, Springer, Vol 78(15), pp. 21245–21271 (2019)

reverted to the original mode by the encoder during re-compression. So, the embedded information gets lost after the re-encoding process.

Dutta and Gupta (2016) proposed an appropriate framework established on the 4×4 intra-luma block coefficients transformed employing DCT. This scheme is resilient towards the re-compression as well as signal processing attacks such as noise additions and filtering. In the case of Liu et al. (2019) the watermark has been integrated through revamping the culled subgroup of the QDST coefficients established on the intra-mode of prediction. The above framework has an immense propensity for ingraining alongside a fair deterioration in ocular quality. Notwithstanding, the framework is not successful with regards to the re-compression employing various QP values, as the choice of 4×4 blocks has been rendered without any rigorous spatial exploration.

In Tew et al. (2016), an authentication code is created utilizing elements of syntax such as CU category and CU prediction mode. The created code has been ingrained through adjusting LSB's appertaining to quantized transformed DCT coefficients, each CTU's QP, as well as the motion vectors, through all these components the tinkered positions are recognized at decoder side. The encryption as well as an authentication scenario considering HEVC encoded videos has been recommended by Tew et al. (2018). The syntax components have been rived in a couplet of distinct groups within this framework, one group have been wielded during the authentication process, together with the second subgroup which has been employed during encryption. The size of the coding component has been employed to generate the authentication code, that has been incorporated within the LSB of the quantized DCT transformed coefficients, every CTU's QP, as well as the motion vectors in order to classify the sectors that have been impaired. These methods are capable of detecting malicious attacks but are vulnerable to manipulations that keep information intact. When the video is re-compressed with different QP values, the encoder changes the syntax elements that are used to embed the authentication code.

According to the aforementioned analysis, it should be remembered that the schemes in (Tew et al. 2016, 2018; Wang et al. 2015) are vulnerable to re-compression of HEVC encoders. In addition, the watermarking scheme (Dutta and Gupta 2016) has employed the DCT transformed coefficients within intra-luma blocks with dimensions 4×4, which are resilient towards re-compression, but the mode-driven option between DCT and DST is omitted from the HM reference framework (HM16) and hence solely DST has been deployed for 4x4 intra-luma blocks (Saxena and Fernandes 2013). There have been divergences among the peculiarities of DST and DCT transforms, such that the strategies centred on DCT transformed coefficients do not work optimally with DST-transformed coefficients. The Chang et al. (2014) as well as the Liu et al. (2019) approaches are

established on the coefficients of the DST transform available in intra-luma blocks with dimension 4×4. However, these schemes are not effective for re-compression and that's because 4×4 blocks for embedding watermark bits are not chosen using any spatial exploration.

In an attempt for circumventing the constraints in the current literature, an appropriate authentication scheme for HEVC coded videos is introduced, centered around QDST coefficients in intra-luma blocks with dimension $\times 4$. The proposed scheme is stable against re-compression attacks as well as it could identify deceitful maneuvering within a video. In this scheme, I-frame's intra-luma blocks of dimension 4×4 have been employed to authenticate the videos, in contrast to inter-frames, I-frames provide much more crucial data. An effective spatial exploration of the frames is conducted to identify the required 4×4 blocks. The spatial exploration has been implemented by exploiting the the count of non-zero QDST coefficients in a block of dimesnison 4×4 for augmenting the ocular quality of the proffered scheme. In addition, spatial analysis is carried out appertaining to obtain the robustness for the proffered scheme depending on the percentage of non-zero QDST coefficients with an absolute value greater than 1 (ABGR1). The blocks chosen through spatial scrutiny have been ripped in couplet of subgroups, first one being employed in the authentication code procreation as well as the another being utilized in authentication code ingraining. It derives each authentication codeutilizingthe signs of the QTC'senclosed within the chosen block. To invigorating the safety concerns of the proffered scheme, the produced authentication code will be encrypted employing a content dependent public key as well as a pseudo-randomly spawned private key. In addition to this, the offered scheme precipitated inconsequential upsurge in bit-rate because the encrypted code has been concealed through modifying the magnitudes appertaining to the QDST coefficients within intra-luma block having dimension 4×4, through retaining the signs of the coefficients.

3.2 Proposed Authentication Scheme

A protracted sketch of the envisaged authentication scheme is provided in this section. First, spatial exploration is defined that is employed to choose blocks. And thereafter, each strategy for generating and embedding authentication codes are discussed. The extraction and validation algorithms are finally construed.

3.2.1 Spatial Exploration of the I-Frames

Each I-Frame can indeed be segmented via TB of dimensions from 32×32 upto 4×4 mostly on grounds of its texture. Uniform segments are split into either 32×32 or 16×16, particularly sections with far less detail are split into larger dimensions and dense textured sections are split into 8×8 or 4×4, specifically the sections with more texture are broken into smaller sizes. TB size 4x4 is chosen for watermark insertion since HVS is less susceptible toward more textured regions (Mansouri et al. 2010).

Spatial exploration for imperceptibility and robustness is depends upon the count of non-zero QDST coefficients contained inside the 4×4 intra-luma block. Spatial evaluation implemented in (Mansouri et al. 2010) was directly related to the number of non-zero quantized DCT coefficients, but perhaps the HEVC codec employs DST rather than DCT to 4×4 intra-luma blocks. Therefore, the study suggested in (Mansouri et al. 2010) cannot be extended for HEVC coded videos owing to their variations in the energy compaction mechanisms among these two transforms. Such distinctions in the energy compaction mechanisms for DST and DCT as described by Salomon (2007) are: (1) DCT yields the DC coefficient, which is the mean value of all the coefficients contained in a block, however DST does not produce a DC coefficient, and Throughout DCT transform, those coefficients which are produced beside the DC coefficient are named as AC coefficients, the consistency of these decreases as they switch towards high frequency from low frequency coefficients, however the DST coefficients don't really obey a certain pattern. An instance of such QDST coefficients within 4×4 intra-luma block can be seen in Table 3.1.

Table 3.1: An instance of QDST coefficients within a 4×4 block

0	1	1	0
4	1	0	0
-4	-3	-1	0
-1	1	0	0

The imperceptibility of the method has been accomplished through picking the right 4×4 intra-luma blocks with quite a hightexture. The block texture has been evaluated through considering countof non-zero QDST coefficients available with in 4×4 intra-luma block. Blocks with a heavily textured area do have a higher count of non-zero QDST coefficients. Quality threshold α, used in (Mansouri et al. 2010), is measured by counting non-zero quantized DCT's AC coefficients within4×4 intra-luma block, but in HEVC, DST has been employed in 4×4blocks rather than DCT. There are no DC and AC coefficients in the DST, so the quality threshold α has been computed relying upon the count of all non-zero QDST coefficients within 4×4 block.

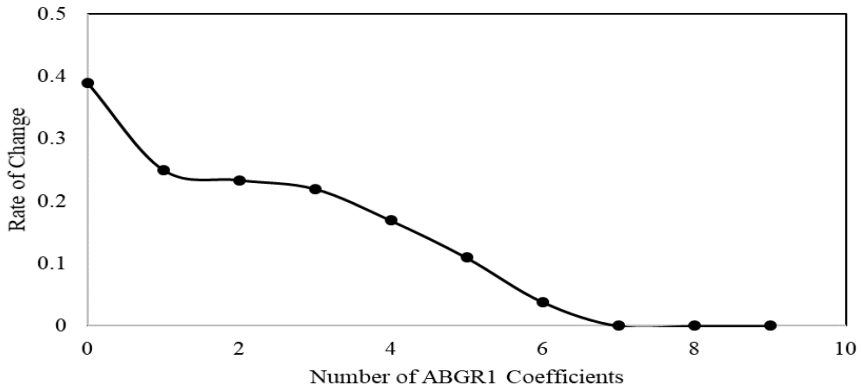
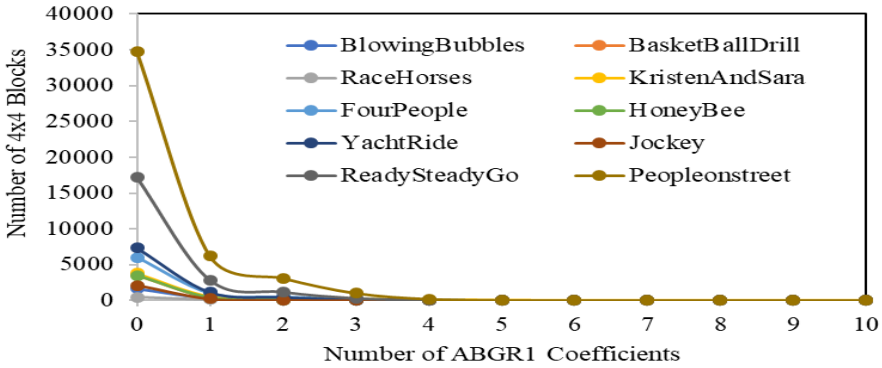
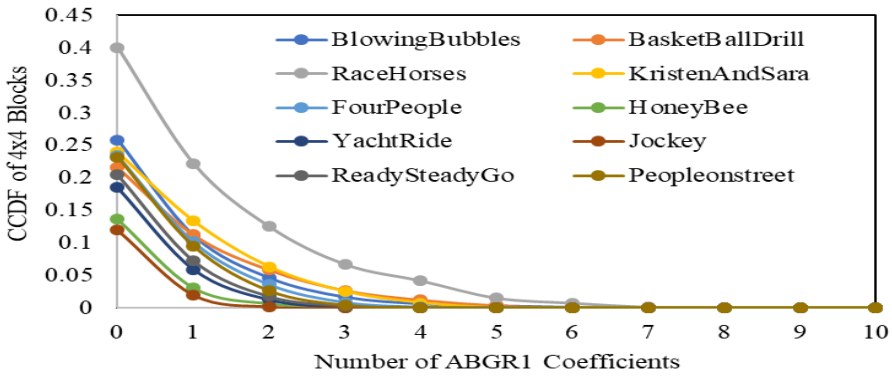


Figure 3.1: Rate of change in 4×4 blocks while re-encoding is conducted through QP=32 to QP=34



(a) 4×4 blocks variation corresponding to ABGR1



(b) CCDF of 4×4 blocks corresponding to ABGR1

Figure 3.2: Variations and CCDF of the first I-frame blocks at QP=32

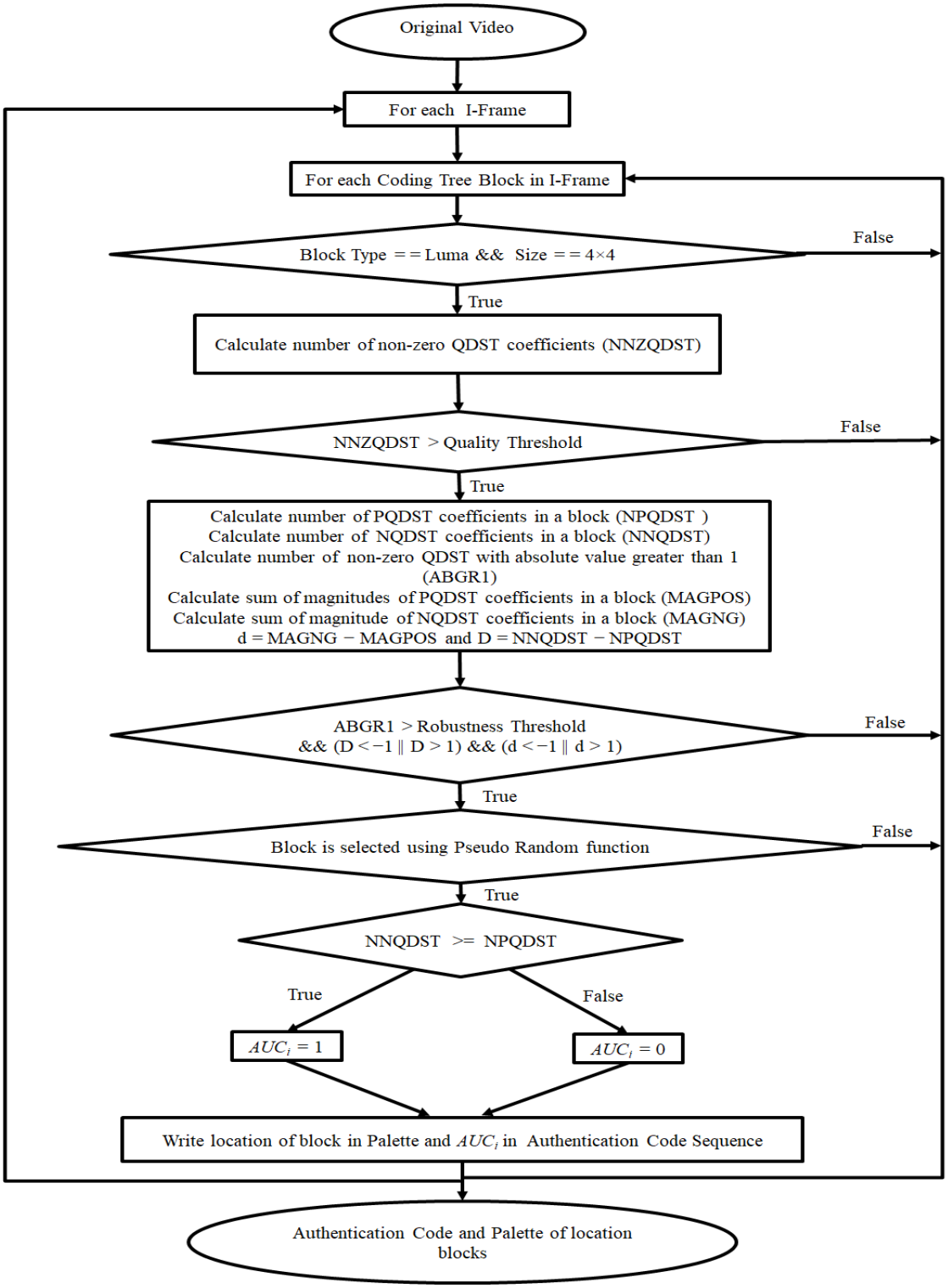


Figure 3.3: Flowchart of authentication code generation

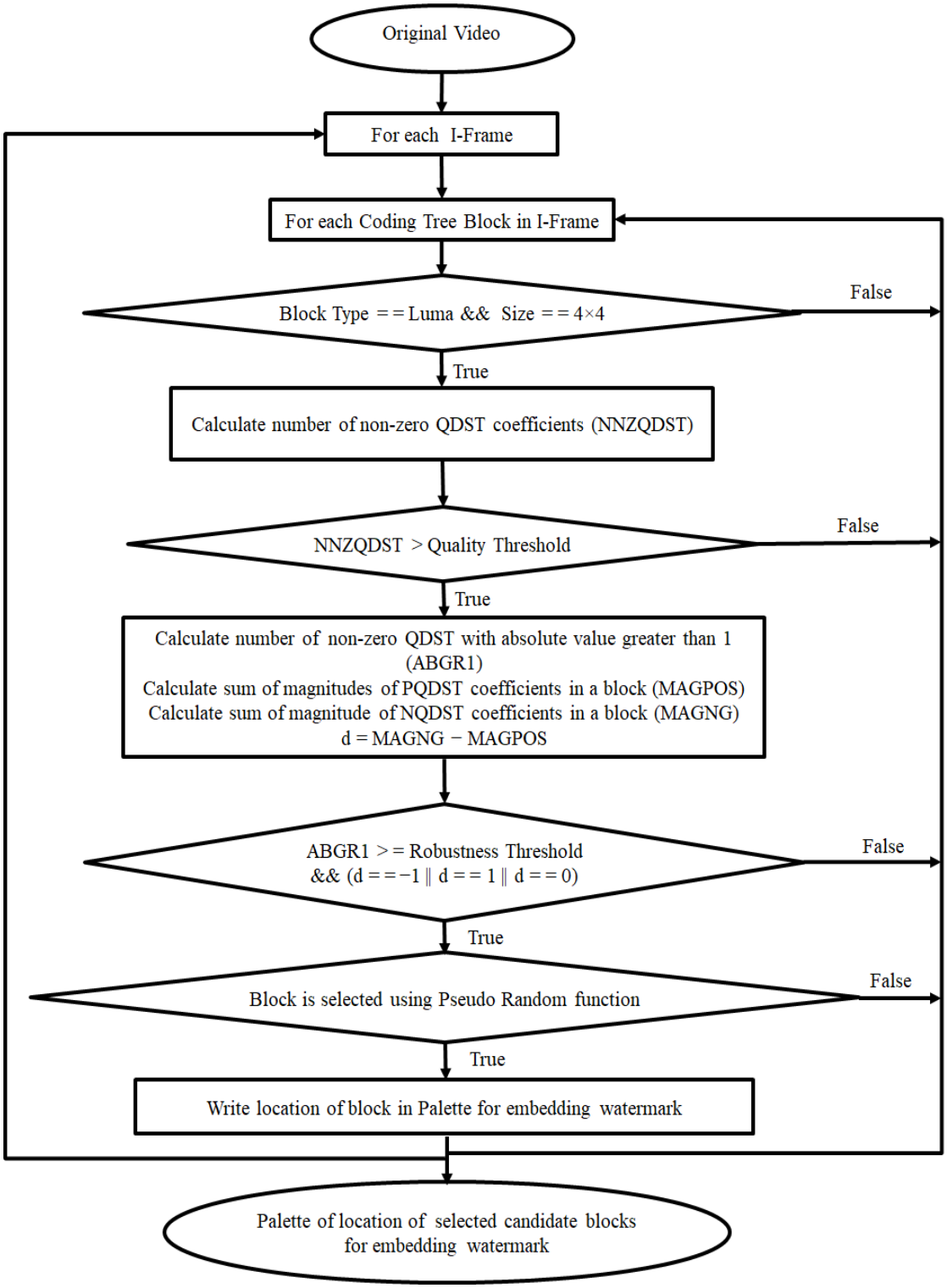


Figure 3.4: Flowchart for choosing the candidate blocks for embedding watermark

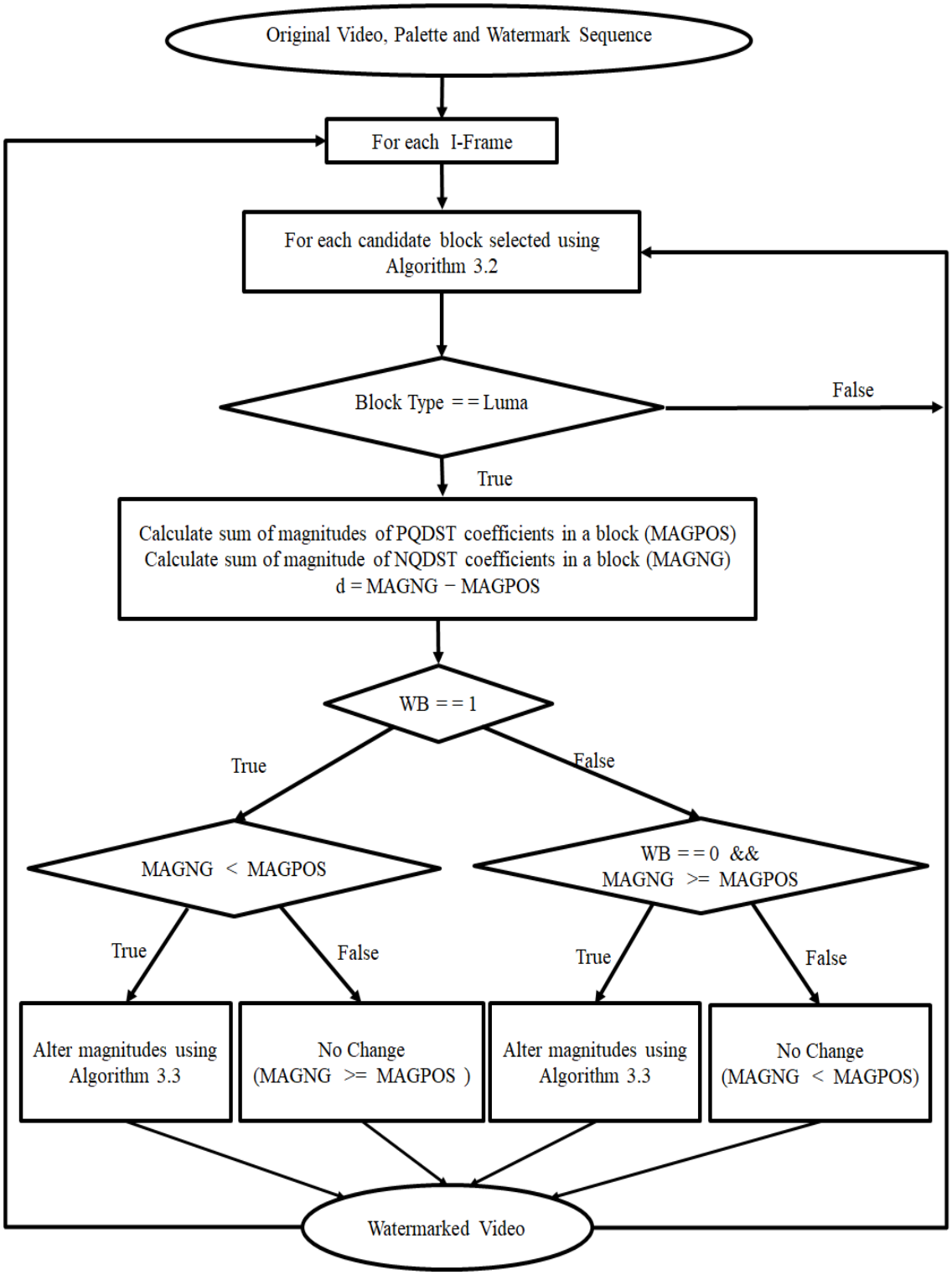


Figure 3.5: Flowchart of embedding of watermark

The efficacy of the suggested solution is dependent mostly on β threshold, specifically established on the non-zero QDST coefficients throughout the 4×4 intra-luma block with ABGR1. The 4×4 blocks with more ABGR1 coefficients indicate that there is a more inhomogeneous texture with fine details. The minutiae information can be reliably estimated through adjoining blocks using the angular intra-prediction method especially whenever such categories of blocks are subdivided into smaller ones. Thus, there are far less risks that somehow the size of such block would be modified to a larger size. Besides that, modifications to angular intra-prediction mode to another mode are often less likely to occur in areas of minute information. This contributes in the consistency of the association in both the PQDST and the NQDST coefficients existing with in block. To illustrate such correlation, we deployed re-compression to a number of standard unmarked videos and thereafter measured the variation for the relevant ABGR1. The Instances of change typically involve (a) an alteration in the relation among the sum of magnitudes of PQDST and the sum of magnitudes of NQDST coefficients, and (b) a modificationin 4×4 block size into a larger one.

The level for β is estimated on the criterion particularly dependent on count of ABGR1 in a 4×4 intra-luma block. Figure 3.1 clearly shows the transition rate with reference to the ABGR1 at the diverse videos are re-compressed through QP 32 to 34. Figure 3.1 indicates that the incidences of modification declines through a growth in the proportion of ABGR1 coefficients present inside a block. The count for ABGR1 coefficients differs through one block to the other depending mostly on I-frame texture. Furthermore, the variation among blocks associated with a particular amount of non-zero QDST coefficients with ABGR1 fluctuates as per the texture of every I-frame. The optimum value for β would be computed during running time relying on both the texture and quantity of blocks needed for the insertion of watermark. A persistence function defined with Complementary Cumulative Distribution Function (CCDF) of the count of blocks containing various ABGR1 quantities has been computed employing (3.1) and (3.2) together.

$$\overline{F}(\beta) > \mu \tag{3.1}$$

The $\overline{F}$ is estimated from the probability distribution of a count of 4×4 blocks with varied amounts of ABGR1 coefficients, employing the following formula of the Mansouri et al. (2010):

$$\overline{F}(\beta) = P(D > a) = 1 - F(\beta) = \sum_{D>\beta} p(D) \tag{3.2}$$

Figure 3.2a demonstrates the variation in 4×4 blocks for ABGR1 throughout the first I-frame from videos containing various textures. Figure 3.2b reveals $\overline{F}$ of blocks in various

videos encoded with HEVC at QP = 32. As can be shown, various values for β will be produced for the identical value of μ completely reliant upon textures within various video. As an instance, for the $\mu = 0.02$, the β for RaceHorses is 3 and perhaps for the PeopleonStreet video is 2. Therefore, every optimum valuation of β is specified on the fly through μ dependent entirely upon video's textures.

3.2.2 Authentication Code Generation

Authentication code has been created from the unfluctuating features contained inside I-frames throughout video coding. The 4×4 intra-luma blocks of I-frame have been subdivided via two disarticulate subgroups, namely the the watermark production subgroup S_g and the watermark embedding subgroup S_e. Both subgroups are created on the principle of the variation among the sum of magnitudes of PQDST coefficients and the sum of magnitudes of NQDST coefficients. This is required to prevent intervention among the S_g and S_e subgroups, ensuring that the identical authentication code could be recovered by the decoding process. The block corresponds to S_e when the difference is in interval of -1 and 1, or else it corresponds to the S_g.

Figure 3.3 displays the process flow for authentication code generator and also the full description of authentication code generator is provided in algorithm 3.1. The association among the PQDST coefficients and the NQDST coefficients has been employed to produce authentication code in this algorithm. The outcomes for this algorithm are PG_k and AUC_k, wherein k is count of CTUs within an I-frame of a video. The decoder requires PG_k to replicate AUC_k throughout validation. The PG_k palette is a document that includes the position of blocks intended for generation of authenticating code. It is required to prevent de-synchronization of code retrieval and confirmation at decoder side (Mansouri et al. 2010). So, the strategy is semi-blind, because merely palette is requisite for the decoder to retrieve and validate the authentication code rather than the entire original video. A protected medium will conveniently relay the palette.

For intensifying the invulnerability of the prototype solution, choice of blocks from the S_g subgroup is accomplished utilizing some pseudo random mechanism to produce authentication code. Then derived authentication code AUC_k is encrypted with the nominal complexity encryption mechanism F_E as well as the content-based key (Key) to obtain the watermark W_k. The key would be required to encrypt the spawned authentication code to counteract the intra collusion threat (Furht and Marques 2003). Initially, each public key would be created by randomly choosing a 16×16 intra-luma block from every I-frame. The DC coefficient and the first two non-zero AC coefficients of each of 8 bits

have been retrieved via the chosen 16×16 block. The public key, a tuple containing 24 bits, has been created by the concatenation of all these coefficients (Farfoura et al. 2016). This same resulting key has been employed to encrypt A_k utilizing the minimal complexity encryption scheme FE which creates watermark, aka W_k The watermark Wk is created by using (3.3). This palette has been created by the encoding process at the time of generation and embedding phase of the authentication code and is relayed to the approved user via a reliable medium. The decoding process retrieves the positions of the blocks employed for obtaining both the public and private key from the palette. The public key would then be derived via DC and the first two non-zero AC coefficients. Lastly, the derived Then the public key would be encrypted with the secret key. This resulting key has been used to decode the authentication code for retrieval and confirmation.

$$W_k = F_E(AUC_k, Key) = \{W_1, W_2, .., W_k\} \tag{3.3}$$

3.2.3 Watermark Embedding

The spawned watermark W_k would be inserted within intra-luma 4×4 blocks identified by the spatial exploration mechanism provided via Section 3.2.1. The Watermark security is provided through pseudo randomly choosing candidate blocks in the S_e subgroup . In the recommended strategy, the watermark is incorporated on the rationale of the correlation within the sum of the magnitude of the PQDST coefficients and the sum of the magnitude of the NQDST coefficients available within that block.

All blocks that could be considered in embedding process have been chosen from the S_e subgroup, dependent on the quality criterion α denotes the count of non-zero QDST coefficients occurring within block as well as the robustness level β that is computed by counting the ABGR1 coefficients occurring throughout the block. The above threshold values are identified on the grounds of the spatial exploration of the Section 3.2.1. Furthermore, a subgroup of eligible blocks from the S_e subgroup are chosen pseudo randomly. The steps in the candidate block selecting procedure as can be seen through Figure 3.4 and the full explanation for each candidate block selecting procedure has been described through the Algorithm 3.2. The whole method creates the PE_k palette consisting the positions each candidate blocks chosen to incorporating the derived W_k watermark. The PE_k, would be employed with Algorithm 3.3 to add the W_k then with Algorithm 3.4 to retrieve and validate this inserted watermark by the decoding process.

Algorithm 3.1: Authentication Code Generation

Data: Original Video
Result: Authentication Code AUC_k, Palette of location of blocks PG_k
foreach *Frame in Original Video* **do**
  if *I Frame* **then** `/* If Frame Type is Intra */`
    foreach *Coding Tree Block in an I Frame* **do**
      if $BlockType == Luma$ && $Size == 4 \times 4$ && $Block \in S_g$ **then**
        Calculate NZQDST `/* Count QDST coefficients with non-zero value */`
        if $NZQDST > \alpha$ **then** `/* α is quality threshold */`
          Calculate NPQDST `/* Count PQDST coefficients */`
          Calculate NNQDST `/* Count NQDST coefficients */`
          Calculate $ABGR1$ `/* Count QDST with absolute value greater than 1 */`
          Calculate MAGPOS `/* sum of magnitudes of PQDST coefficients in a block */`
          Calculate MAGNG `/* sum of magnitude of NQDST coefficients in a block */`
          $d = MAGNG - MAGPOS$
          $D = NNQDST - NPQDST$
          if $(d < -1 \parallel d > 1)$&& $ABGR1 > \beta$ && $(D < -1 \parallel D > 1)$ **then** `/* β is robustness threshold */`
            if *Block is selected using Pseudo random function* **then**
              if $NNQDST \geq NPQDST$ **then**
                AUC_i=1
              else
                AUC_i=0
              Add location of block in Palette PG_k
            else
              Go to next block if block is not selected by pseudo random function

Algorithm 3.2: Selection of Candidate blocks

Data: Original Video
Result: Palette of location of chosen candidate blocks for the embedding process PE_k

```
foreach Frame in Original Video do
    if I Frame then                                   /* If Frame Type is Intra */
        foreach Coding Tree Blocks In Frame do
            if BlockType == Luma && Size == 4 × 4 && Block ∈ S_e then
                Calculate NZQDST                      /* Count QDST coefficients with
                 non-zero value */
                if NZQDST ≥ α then                    /* α represents quality threshold
                 */
                    Calculate MAGPOS                  /* sum of magnitudes of PQDST
                     coefficients in a block */
                    Calculate MAGNG                   /* sum of magnitude of NQDST
                     coefficients in a block */
                    Calculate ABGR1                   /* Number of non-zero QDST with
                     magnitude greater than 1 */
                    d = MAGNG − MAGPOS
                    if ABGR1 ≥ β && (d == −1 || d == 1 || d == 0) then  /* β
                     is robustness threshold */
                        if Block is selected using Pseudo random function then
                            Add location of block in Palette PE_k for watermark
                             embedding
                        else
                            Go to next block if block is not selected by pseudo
                             random function
```

Algorithm 3.3: Watermark Embedding Algorithm

Data: PE_k is the location of blocks chosen by the process of watermark embedding and W_k(Sequence of Watermark Bits(WB)

Result: Watermarked Video

foreach *Frame in Original Video* **do**
  if *I Frame* **then** `/* If Frame Type is Intra */`
    foreach *Candidate block selected using Algorithm 3.2 in I Frame* **do**
      if *BlockType* == *Luma* **then**
        Calculate MAGPOS `/* sum of magnitudes of PQDST coefficients in a block */`
        Calculate MAGNG `/* sum of magnitude of NQDST coefficients in a block */`
        $d = MAGNG - MAGPOS$
        if $WB == 1 \&\& MAGNG >= MAGPOS \| WB == 0 \&\& MAGNG < MAGPOS$ **then**
          No change
        else
          if $WB == 1 \&\& MAGNG < MAGPOS$ **then**
            while $MAGNG < MAGPOS$ **do**
              for $i <= 3$ **do**
                for $j <= 3$ **do**
                  if $C_{ij} < -1 \&\& C_{ij}$ *does not belong to protected set of pixels* **then**
                    Increase magnitude of C_{ij} by one
          else if $WB == 0 \&\& MAGNG >= MAGPOS$ **then**
            while $MAGNG >= MAGPOS$ **do**
              for $i <= 3$ **do**
                for $j <= 3$ **do**
                  if $C_{ij} > 1 \&\& C_{ij}$ *does not belong to protected set of pixels* **then**
                    Increase magnitude of C_{ij} by one

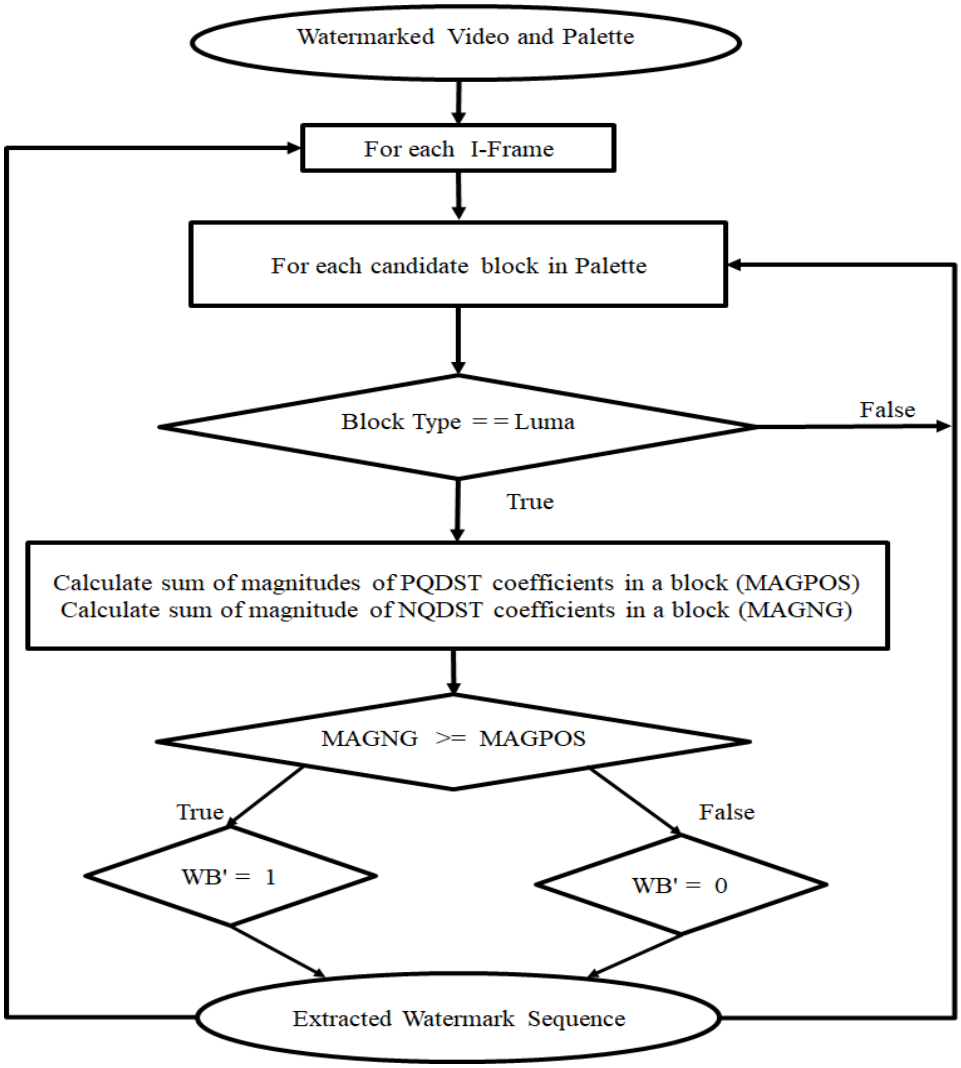


Figure 3.6: Flowchart for extraction of watermark

At the very last level in I-frame coding, watermark chunks have been incorporated through QDST coefficients. Such inclusion introduces the issue of drift aggregation, as the adjustments rendered to QDST coefficients in the current block will spread to the further nearby blocks at the decoder side's prediction mechanism. For fixing this concern, 4×4 QDST coefficients were separated among two groups, particularly the first group of coefficients required to accurately incorporate the watermark but the other group would not be employed for embedding watermark. The indexes of QDST coefficients contained in 4×4 block are displayed in Table 3.2. The seven indexes $\{C_{i,3}\}_{i=0,1,2,3} \bigcup \{C_{3,j}\}_{j=0,1,2,3}$ specifically correspond to a reserved set out of sixteen indexes of a 4×4 block that will never be utilized to insert watermark bits to eliminate I-frame error drift glitches (Chang et al. 2014).

Table 3.2: QDST coefficients indexes in 4×4 block

C_{00}	C_{01}	C_{02}	C_{03}
C_{10}	C_{11}	C_{12}	C_{13}
C_{20}	C_{21}	C_{22}	C_{23}
C_{30}	C_{31}	C_{32}	C_{33}

Figure 3.5 displays the flow diagram of the watermark integration mechanism and is clarified in depth through Algorithm 3.3. The parameters for the algorithm are the palette with specified candidate blocks PE_k as well as the watermark W_k, particularly developed through Algorithm 3.1 as defined previously. The watermark bit has been integrated on criterion of the magnitudes of PQDST and NQDST, such that only non-zero QDST coefficients have been updated for incorporating the watermark bit for limiting the spike in the bitrate of the watermarked video. In additament, the watermark has always been inserted only in certain blocks where the margin among MAGPOS and MAGNG is -1, 0 or 1. This is just to eliminate noticeable distortion, since the maximum of two QDST have been affected in such a block by limiting this disparity.

3.2.4 Watermark Extraction and Verification

The retrieval and validation of watermarks by the decoder component is successfully accomplished through entropy decoding of watermarked video's QDST coefficients. This watermark retrieval strategy works precisely the opposite of the watermark addition, and the watermark retrieval sequence is shown through the Figure 3.6. The QDST coefficients are parsed for two primary reasons, to either recreate the authentication code AUC'_k, or to recover the watermark W'_k engrained through the encoding of the video. Content dependent authentication code AUC'_k would be resurrected through Algorithm 3.1 and palette PG_k. Each bit of W'_k sequence is derived utilising Algorithm 3.4 and the PE_k palette created through Algorithm 3.2. The retrieved W'_k would be decrypted employing a decryption operation F_D employing (3.4) that further overturns the encryption component for encoding and the key created by the encoder, for the purpose of extraction of the authentication code AUC''_k.

$$AUC''_k = F_D(W'_k, Key) \tag{3.4}$$

Each I-frame throughout the video would be screened via matching the authentication code AUC''_k retrieved just at decoder point with the authentication Code AUC'_k recon-

structed then at decoder point by employing (3.5).

$$\begin{aligned} &if\ AUC'_k \oplus AUC''_k == 1,\ IF_j\ is\ Unauthentic \\ &if\ AUC'_k \oplus AUC''_k == 0,\ IF_j\ is\ Authentic \end{aligned} \tag{3.5}$$

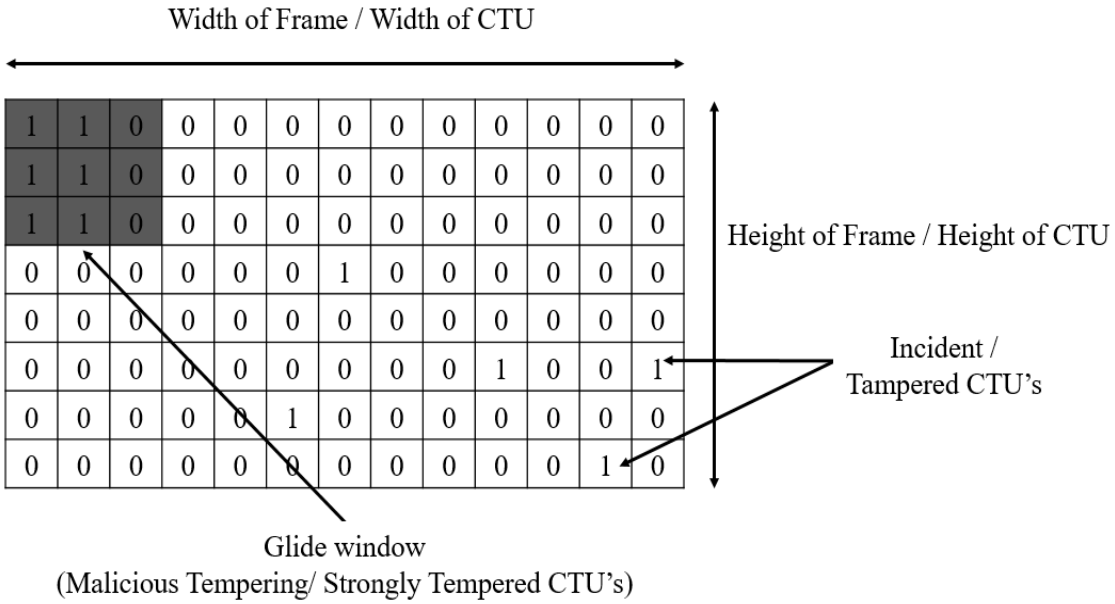


Figure 3.7: AIM with the glide window of dimension 3×3

The Equation (3.5) is employed to produce the Attack Identity Matrix (AIM) for every IF_j (the value of the variable j ranges between 1 to f as well as f is therefore the maximum count for the I-frames throughout the video). If $\oplus$ *i.e.* xor of AUC'_k and AUC''_k is 1, it denotes that the CTU is authentic and 0 denotes CTU is unauthentic. For further tamper extrapolation, the methodologies of Xu et al. (2011) as well as Farfoura et al. (2016) are exploited. For diagnosing the maliciously fiddled location in an $I - Frame_i$, the gliding window measuring 3×3 and a tamper mitigation criterion T_γ (total value of 1's in a gliding window) have been specified in (Xu et al. 2011). The T_γ is set to invariable value specifically 5 dependent upon a probability theorem which utilises a binomial distribution. The media filtering in yet another Farfoura et al. (2016) technique has been deployed for mostly two primary reasons, first for the elimination of dispersed incidental 1fls, as well as for dealing with non-error items 0, particularly if the 0 is ringed with five and sometimes more 1's, in order to preserve clustering to discover malicious interference. As observed through both precursive techniques, these fiddled positions throughout AIM have been scattered in two ways: in the event through re-compression as well as noise attack, the 1's were scattered in the pattern of random noise whereas in the incident of deliberate

interference, the 1's were mostly acquired at specific points. An instance of the AIM is illustrated via Figure 3.7 containing a gliding window of dimension 3×3.

In the AIM in Figure 3.7, perhaps the manipulated CTU's (1's) within the area of the glide window will be considered as Strongly Tampered Elements (STE) because they have been concentrated in a specific area of the frame, while the 1's outside the glide window are considered to be Incidental / Tampered Elements (TE) as these are almost dispersed. The width and height of the rolling window has been set to 3×3 as well as the $T\gamma$, particularly the count of the 1's inside the rolling window has been set to 5 (Xu et al. 2011).

The method of the Farfoura et al. (2016) for the inspection of the malicious alterations is as under:

1. Compute the Tampered Percentage (TP) using (3.6)

$$\text{TP} = \frac{\text{Count Tampered CTU of the frame}}{\text{Count all CTU of the frame}} \tag{3.6}$$

2. Scattered TE are eliminated using 3×3 median filter.

3. Compute Strongly Tampered Percentage (STP) as follows:-

$$\text{STP} = \frac{\sum_{\forall} \text{STE}}{\sum_{\forall} \text{TE}} \tag{3.7}$$

4. Evaluate the STP with respect to the tamper detection threshold T_m using (3.8):

$$\begin{cases} IF_i \text{ is unauthnetic, if STP} \geq T_\lambda \\ IF_i \text{ is authnetic, if STP} < T_\lambda \end{cases} \tag{3.8}$$

$$where\ P(T_\lambda \geq \Omega) = 1 - P(T_\lambda < \Omega) \approx 1 - \left(\left(\frac{E(\text{STE})}{\sqrt{Var(\text{STE})}} \right) / E(\text{TE}) \right) \tag{3.9}$$

It is obvious that the every component of the AIM could either be 1 or otherwise 0, and that each component has a chance of 0.5. It may then be regarded as a random variable thatpursuethe distribution of the binomial probability. Assuming that the likelihood that an error will occur with every random variable of AIM is 0.5. Equivalently, the likelihood of any 5 and sometimes more spontaneous errors in the $n \times n$ slide window can

be determined using (3.10) that is particularly the cumulative likelihood.

$$P(i \geq 5) = \sum_{i=5}^{n \times n} \binom{n \times n}{i} (0.5)^i (0.5)^{n \times n - 1} \tag{3.10}$$

The output of the Equation (3.10) by setting $n = 3$ is 0.5. Formely, the expected output of the STE ($E(STE)$) and also the variance of STE ($Var(STE)$) both will be $0.5 \times NE$ and $0.5 \times (1 - 0.5) \times NE$. The NE is variable to represent the AIM's elements which are tampered. The likelihood of any possible aspect being identified as tampered within the gliding window is $0.5(1 - (0.5)^8)$=0.498. Therefore $0.498 \times NE$ is the expected value of just the tampered elements ($E(TE)$)(Farfoura et al. 2016).

Algorithm 3.4: Watermark Extraction Algorithm

Data: PE_k (Palette of location of selected candidate blocks for embedding the watermark), Watermarked Video
Result: W'_k (Watermark Bits (WB) sequence inserted while video encoding)
foreach *Frame in Watermarked video* **do**
  if *I Frame* **then** `/* If Frame Type is Intra */`
    foreach *Candidate block in palette* PE_k **do**
      if *BlockType == Luma* **then**
        Calculate MAGPOS `/* sum of magnitudes of PQDST coefficients in a block */`
        Calculate MAGNG `/* sum of magnitude of NQDST coefficients in a block */`
        if $MAGNG >= MAGPOS$ **then**
          $WB' = 1$
        else
          $WB' = 0$

3.3 Experimental Results

The proffered authentication scheme has been enforced and investigated via HM model (HM16) excecution on a workstation with Windows 10 software platform, 8 G RAM and Intel i7-8550U Processor@1.80 GHz. Simulation findings were acquired through employing the default configuration file of intra-main profile via setting closed GOP layout, QP=32, IntraPeriod= 8, GOPSize= 4 and DecodingRefereshType = 2. The different videos of varying resolutions and textures were inspected to demonstrate overall reliability of the proffered authentication scheme. The Table 3.3, unveils the particulars of the various videos used during screening.

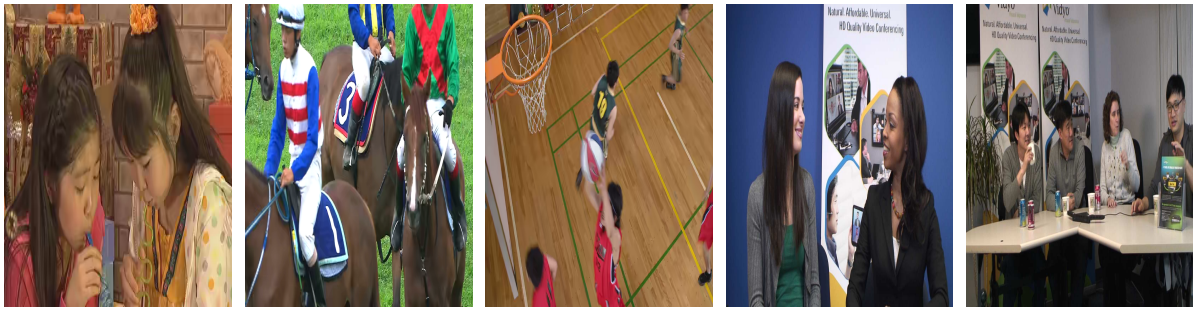
The potency of the proffered scheme would be scrutinized in terms of computational intricacy, BIR, visual quality, tolerance to re-compression, inclusion of noise, dropping of frame and fragility towards deceitful assaults. The association between robustness, visual efficiency, embedding capability and escalation in bit rate has been intersecting. The higher embedding capability leads to greater visual distortion, BIR as well as decreased robustness towards attacks, also vice versa. Transparency, tolerance towards re-compression as well as fragility towards deceitful assaults are preeminent requirements of the authentication scheme.

The appropriate transparency as well as robustness can be procured employing α quality threshold and even β robustness threshold. These control parameters are primarily dependent on the texture of various videos and estimated on the fly with CCDF, as demonstrated in the Section 3.2.1. There is sharp inflation in robustness and reduction in visual distortion from the higher value of both α and β, however this will cutback the number of candidate blocks both for producing and as well as integrating watermark information as well as vice versa. Thus, relying on preliminary findings from the diverse videos, the optimal value for the CCDF is set to 0.1 for α while 0.01 for β to achieve an optimum value for these thresholds during execution.

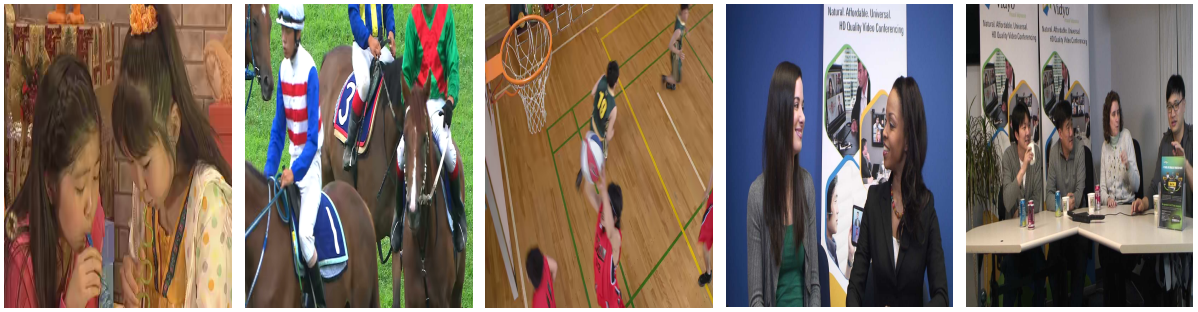
The proffered method exploits the QTC's in the blocks with dimension 4×4 to implement the procedure ofauthentication. The competency of the proffered scheme has been contrasted with the other schemes particularly, Chang et al. (2014), Dutta and Gupta (2016) and Liu et al. (2019) since these are all focused around QTC's in blocks of dimension 4×4 and that of the intra-luma category.

The ocular quality of the recommended authentication system is assessed in terms of subjective aspect along with objective metrics. Appertain to subjective assessments, original and watermarked frames of video under experimentation are revealed in the Figure 3.8. The Figures 3.8a, 3.8c, 3.8b, 3.8d, 3.8e, 3.8k, 3.8l, 3.8m, 3.8n and 3.8o show the original frames of tested videos. The Figure. 3.8f, 3.8g, 3.8h, 3.8i, 3.8j, 3.8p, 3.8q, 3.8r, 3.8s and 3.8t display the watermarked frames of the videos under investigation. Thus, the Figure 3.8, provides the ocular proof that there are absolutely no perceptible artifacts in the watermarked video frames.

The results for objective visual quality metrics for proffered authentication scheme as well as of the (Chang et al. 2014; Dutta and Gupta 2016; Liu et al. 2019) has been provided in Table 3.4. In Table 3.4 the average is the average of the results for all the videos in the data set. The δ_{PSNR}, $SSIM$ and $VIFp$ for the proffered scheme are 0.1253 dB, 0.99274 and 0.96603 respectively. The δ_{PSNR}, $SSIM$ and $VIFp$ for Chang et al. (2014) scheme are 0.1874dB, 0.99236 and 0.95145 respectively. The δ_{PSNR}, $SSIM$ and $VIFp$



(a) Original (b) Original (c) Original (d) Original (e) Original



(f) Watermarked (g) Watermarked (h) Watermarked (i) Watermarked (j) Watermarked

(k) Original (l) Original (m) Original (n) Original (o) Original

(p) Watermarked (q) Watermarked (r) Watermarked (s) Watermarked (t) Watermarked

Figure 3.8: Original as well as the watermarked I-frames appertaining to various videos at QP equivalent to 32

Table 3.3: Data set of videos used for experimentation of the authentication scheme

Serial Number of Video	Name of Video	Resolution	Number of Frames	Frame Rate
1	BlowingBubbles	416×240	500	50
2	BasketBallDrill	832×480	500	50
3	RaceHorses	832×480	500	30
4	KristenAndSara	1280× 720	600	60
5	FourPeople	1280×720	600	60
6	HoneyBee	1920×1080	600	120
7	YachtRide	1920×1080	600	120
8	Jockey	1920×1080	600	120
9	ReadySteadyGo	1920× 1080	600	120
10	PeopleOnStreet	2560× 1600	150	30

for Dutta and Gupta (2016) are 0.2298dB, 0.99035 and 0.94294 respectively. The δ_{PSNR}, $SSIM$ and $VIFp$ for the Liu et al. (2019) are 0.1982dB, 0.99158 and 0.95079 respectively. Gleaning from the above-mentioned outcomes, it can be wisecracked that the proffered authentication scheme is more potent in contrast to the Chang et al. (2014), Dutta and Gupta (2016) and Liu et al. (2019) schemes in aspects such as imperceptibility, since the 4×4 intra-luma blocks which have been wielded during watermark insertion are culled using a precised quality threshold α specifically reckoned utilizing the count of non-zero QDST coefficients within a blocks. Therefore, through the ocular interpretation from Figure 3.8 as well as the objective assessment in terms of $PSNR$, $SSIM$ and $VIFp$, the proffered authentication scheme is optically indistinguishable moreover satisfies the demands appertaining to the a practical video authentication strategy.

The time intricacy concerning the proffered authentication scheme is remarkably low since algorithms wielded to derive the authentication code, choosing the candidate blocks, implant the derived authentication code, retrieve, and verify employ some primitive arithmetic operations. The proffered authentication system has been realized through changing the HM software. Consequently, the time intricacy of the proffered scheme has been determined with regard to the basic HM software through the computations of the overhead processing time appertaining to amended HM software.

Figure 3.9a together with Figure 3.9b exhibit the mean overhead computing time in tweaked encoder as well as the decoder of HM software. Initially, intra-prediction modes

for the nearby blocks are extracted by (Chang et al. 2014; Liu et al. 2019) before picking the QTC's for watermark ingraining to circumvent the dissemination of intra-drift errors. Alternatively, the proffered authentication scheme directly prohibits the use of coefficients for incorporating the watermark, which will be utilized for the prediction the neighbouring blocks avoid the retrieval of intra-prediction mode. Further, the supplementary time overhead in (Dutta and Gupta 2016) is attributed to pseudo-motion-vector reckoning for avoiding the error accumulation to further blocks. The average encoding time overheads of the proffered scheme, Chang et al. (2014), Dutta and Gupta (2016) and Liu et al. (2019) schemes are 1.64, 1.75, 1.85 and 1.83 sec respectively. The average decoding time overheads of the proffered scheme, Chang et al. (2014), Dutta and Gupta (2016) and Liu et al. (2019) schemes are 0.57 sec, 0.6 sec, 0.62 sec and 0.61 sec respectively. Therefore, it could be reckoned from the above scrutiny that the proffered authentication scheme has minor time intricacy in contrast to the current schemes.

The proffered authentication scheme ingrains the watermark bits through tweaking amplitudes of the non-zero QDST coefficients whereas retaining the signs, which then in effect retrench the increase in bit-rate of the videos after the implantation of the watermark.The BIR of the different videos for the proffered authentication as well as the schemes of (Chang et al. 2014; Dutta and Gupta 2016; Liu et al. 2019) has been shown through the Figure 3.10 with QP=32. The BIR of the proffered scheme and for the Chang et al. (2014), Dutta and Gupta (2016), Liu et al. (2019) is 1.54%, 1.65%, 1.63% and 1.69% respectively. From all these results of the BIR, it could be inferred that the proffered scheme is superior to any of the schemes of (Chang et al. 2014; Dutta and Gupta 2016; Liu et al. 2019).

In the first instance, the potency of the proffered authentication scheme has been assessed specifically for robustness towards content protecting assaults, particularly resilience against re-compression through QP's, noise intrusions as well as frame dropping. Furthermore, the productivity of the scheme has been measured for its vulnerability to malicious maneuverings. The re-compression tolerance has been accomplished by re-encoding watermarked videos including differing QP values of 30, 32 as well as 34. The outcomes of the re-compression intrusion of the proffered authentication scheme and that for schemes of (Chang et al. 2014; Dutta and Gupta 2016; Liu et al. 2019) has been revealed in Table 3.5.

The means correlation coefficient for re-encoding with QP=30, with respect to the proffered scheme and that for the Chang et al. (2014), Dutta and Gupta (2016) and Liu et al. (2019) 0.8026, 0.4694, 0.4265 and 0.4763. The mean correlation coefficient for re-encoding again with QP=32, for the proffered scheme and that of schemes in Chang et al. (2014),

Dutta and Gupta (2016) and Liu et al. (2019) are 0.8775, 0.8221, 0.7568 and 0.8277. The mean correlation coefficient for encoding through a larger, specifically QP=34, with respect to proffered scheme as well as for the scheme in Chang et al. (2014), Dutta and Gupta (2016) and Liu et al. (2019) is 0.7438, 0.4471, 0.399 and 0.4587. The cumulative output of the proffered scheme in regards to re-compression through all the three QP values is 0.8079 as well as for the Chang et al. (2014), Dutta and Gupta (2016) and Liu et al. (2019) are 0.5795, 0.5274 and 0.5875 respectively.

In consideration to numerous video manipulation assaults, including the frame dropping, Gaussian noise having average zero as well as variance = 0.001, Gaussian filtering of radius 3×3 as well as with the sigma 0.3, the proffered authentication scheme has been tested. The frame dropping assault is achieved through dropping 25 percent arbitrary frames employing the method of (Li et al. 2015). Li et al. (2015) integrated a fifteen-bit binary frame reference before the real watermark. In retrieval process, initially the binary reference is retrieved at the decoder end for defining the frame address, and then the integrated watermark would be retrieved.

In the Table 3.6, the video manipulation attacks outcomes are summarized for the both proffered authentication scheme as well as the current schemes. The mean correlation coefficients for the Gaussian filter, noise and frame dropping for the proffered scheme are 0.7246, 0.7085 and 0.7238. Further, for the Chang et al. (2014) scheme are 0.6917, 0.6766 and 0.3307 respectively. The results for Dutta and Gupta (2016) scheme are 0.4886, 0.4843 and 0.2973 respectively. Lastly, for Liu et al. (2019) scheme are 0.7054, 0.6796 and 0.3289 respectively. Thus, the aforementioned outcomes point out that the robustness of the proffered authentication system is superior to that of the schemes in (Chang et al. 2014; Dutta and Gupta 2016; Liu et al. 2019).

The robustness for this proffered authentication scheme has been accomplished via choosing blocks for ingraining the watermark bits on the criterion of the robustness threshold β computed during running time employing CCDF relying on a spatial exploration of every I-frame expounded through Section 3.2.1. The strategies of (Chang et al. 2014; Liu et al. 2019) have not wielded any standards while assigning blocks for ingraining of the watermark as well as the selection criteria of Dutta and Gupta (2016) has been reliant upon DCT coefficients, however HEVC DST employs in the 4×4 intra-luma blocks. Thereupon, above mentioned current methods are not potent for the HEVC format.

Table 3.4: Contrasting the ocular quality of the proffered authentication scheme with the current schemes

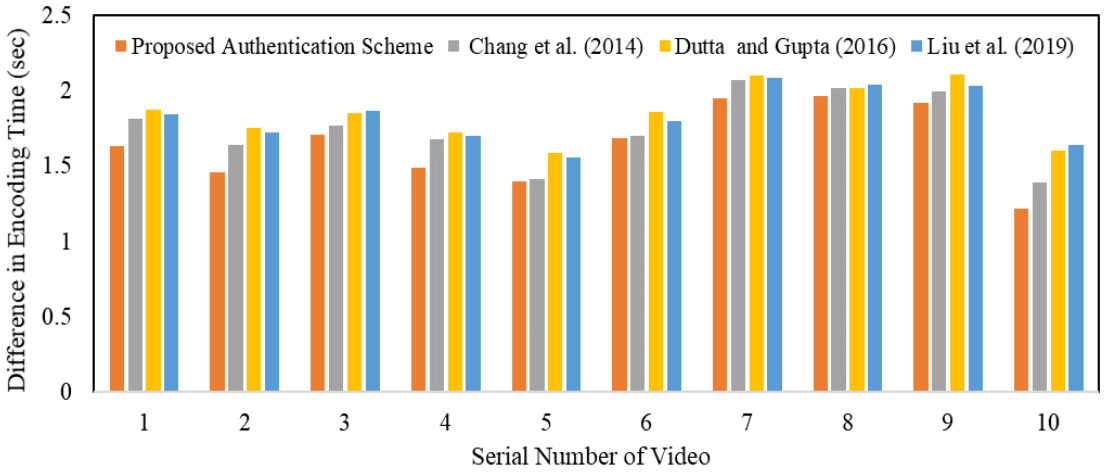
Name of Video	Proposed Authentication Scheme			Chang et al. (2014)			Dutta and Gupta (2016)			Liu et al. (2019)		
	δ_{PSNR}(dB)	$SSIM$	$VIFp$	δ_{PSNR}(dB)	$SSIM$	$VIFp$	δ_{PSNR}(dB)	$SSIM$	$VIFp$	δ_{PSNR}(dB)	$SSIM$	$VIFp$
BlowingBubbles	0.1054	0.99252	0.93885	0.1542	0.99203	0.91068	0.1941	0.99002	0.90217	0.1697	0.99173	0.91022
BasketBallDrill	0.1179	0.99259	0.96648	0.1749	0.99213	0.95212	0.2141	0.99012	0.94361	0.1876	0.99159	0.95182
RaceHorses	0.1329	0.99319	0.96791	0.1966	0.99303	0.95425	0.2397	0.99102	0.94574	0.2024	0.99217	0.95335
KristenAndSara	0.1186	0.99344	0.98518	0.1886	0.99341	0.98017	0.2317	0.99142	0.97166	0.1971	0.99197	0.97964
FourPeople	0.1391	0.99284	0.97356	0.1951	0.99252	0.96274	0.2382	0.99051	0.95423	0.1983	0.99086	0.96209
HoneyBee	0.1474	0.99231	0.95806	0.2136	0.99172	0.93949	0.2567	0.98971	0.93098	0.2253	0.99133	0.93852
YachtRide	0.1284	0.99237	0.96695	0.2027	0.99181	0.95282	0.2458	0.98981	0.94431	0.2218	0.99125	0.95215
Jockey	0.1414	0.99194	0.96293	0.2041	0.99116	0.94679	0.2472	0.98915	0.93828	0.2189	0.99017	0.94629
ReadySteadyGo	0.1301	0.99295	0.96884	0.1969	0.99268	0.95566	0.2402	0.99067	0.94715	0.2066	0.99219	0.95493
PeopleOnStreet	0.0919	0.99326	0.97157	0.1471	0.99314	0.95976	0.1902	0.99113	0.95125	0.1546	0.99255	0.95894
Average	0.1253	0.99274	0.96603	0.1874	0.99236	0.95145	0.2298	0.99035	0.94294	0.1982	0.99158	0.95079

Table 3.5: The Correlation coefficient comparison of proffered authentication scheme with current schemes appertaining to re-compression attack

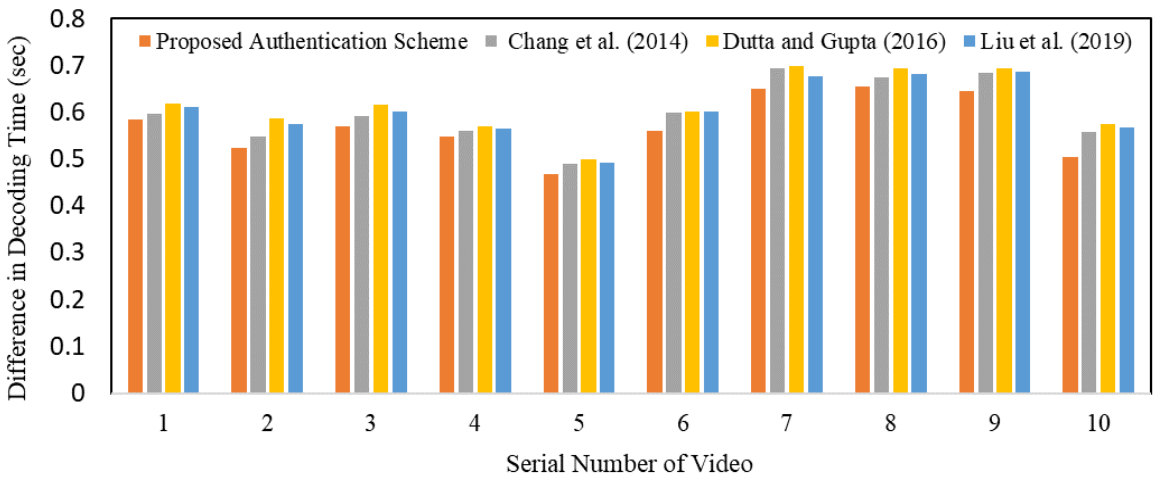
Name of Video	Proposed Authentication Scheme			Chang et al. (2014)			Dutta and Gupta (2016)			Liu et al. (2019)		
	Re-compression			Re-compression			Re-compression			Re-compression		
	QP=30	QP=32	QP=34	QP=30	QP=32	QP=34	QP=30	QP=32	QP=34	QP=30	QP=32	QP=34
BlowingBubbles	0.8194	0.8805	0.7786	0.5124	0.814	0.4214	0.4695	0.7487	0.3733	0.5267	0.8281	0.4287
BasketBallDrill	0.8018	0.8705	0.7622	0.5608	0.89	0.4612	0.5179	0.8247	0.4131	0.5747	0.8806	0.4716
RaceHorses	0.7975	0.8606	0.6503	0.4353	0.8857	0.3979	0.3924	0.8204	0.3498	0.4379	0.8928	0.4061
KristenAndSara	0.7534	0.8738	0.7255	0.4018	0.8825	0.391	0.3589	0.8172	0.3429	0.4076	0.8982	0.4172
FourPeople	0.7864	0.8761	0.7172	0.4279	0.8274	0.5334	0.385	0.7621	0.4853	0.4328	0.8395	0.5385
HoneyBee	0.8524	0.8763	0.7519	0.4379	0.539	0.4261	0.395	0.4737	0.378	0.4415	0.5441	0.4276
YachtRide	0.8231	0.8875	0.8134	0.5033	0.8324	0.4896	0.4604	0.7671	0.4415	0.5088	0.8247	0.4967
Jockey	0.7962	0.8857	0.7323	0.4986	0.8122	0.4644	0.4557	0.7469	0.4163	0.5021	0.8147	0.4689
ReadySteadyGo	0.7928	0.8829	0.7823	0.469	0.8561	0.4342	0.4261	0.7908	0.3861	0.4642	0.8621	0.4676
PeopleOnStreet	0.8031	0.8813	0.7244	0.4471	0.882	0.4521	0.4042	0.8167	0.404	0.4662	0.8918	0.4636
Average	0.8026	0.8775	0.7438	0.4694	0.8221	0.4471	0.4265	0.7568	0.399	0.4763	0.8277	0.4587

Table 3.6: The Correlation coefficient comparison of proffered authentication scheme with current schemes appertaining to frame dropping as well as noise attacks

Name of Video	Proposed Authentication Scheme			Chang et al. (2014)			Dutta and Gupta (2016)			Liu et al. (2019)		
	Gaussian Filter	Gaussian Noise	Frame Dropping	Gaussian Filter	Gaussian Noise	Frame Dropping	Gaussian Filter	Gaussian Noise	Frame Dropping	Gaussian Filter	Gaussian Noise	Frame Dropping
BlowingBubbles	0.6543	0.7044	0.7229	0.6247	0.6748	0.3542	0.4434	0.4854	0.3197	0.6383	0.6607	0.3615
BasketBallDrill	0.6153	0.7024	0.6892	0.5994	0.6691	0.3011	0.4246	0.4644	0.257	0.6133	0.6797	0.2907
RaceHorses	0.7044	0.7132	0.7736	0.6609	0.6807	0.3633	0.4464	0.5762	0.3441	0.6635	0.6978	0.3814
KristenAndSara	0.7005	0.7108	0.7069	0.6386	0.672	0.3619	0.4497	0.4798	0.3288	0.6438	0.6563	0.3981
FourPeople	0.7618	0.7174	0.7077	0.7426	0.6923	0.3563	0.4788	0.4484	0.3222	0.7475	0.6802	0.3512
HoneyBee	0.8245	0.7171	0.7526	0.7846	0.6848	0.3414	0.5885	0.5581	0.3334	0.8082	0.6998	0.3429
YachtRide	0.7141	0.6943	0.7217	0.6957	0.6693	0.2973	0.4771	0.4083	0.2396	0.7112	0.6816	0.2809
Jockey	0.8002	0.7119	0.7692	0.7296	0.6792	0.3138	0.5362	0.5019	0.2896	0.7431	0.6817	0.3093
ReadySteadyGo	0.7581	0.7189	0.6545	0.7445	0.6905	0.3017	0.5381	0.4359	0.2432	0.7695	0.6845	0.2683
PeopleOnStreet	0.7126	0.6941	0.7393	0.6966	0.6535	0.3158	0.5036	0.4841	0.2954	0.7157	0.6733	0.3043
Average	0.7246	0.7085	0.7238	0.6917	0.6766	0.3307	0.4886	0.4843	0.2973	0.7054	0.6796	0.3289



(a) Watermark Embedding



(b) Watermark extraction and verification

Figure 3.9: Results of average time overhead

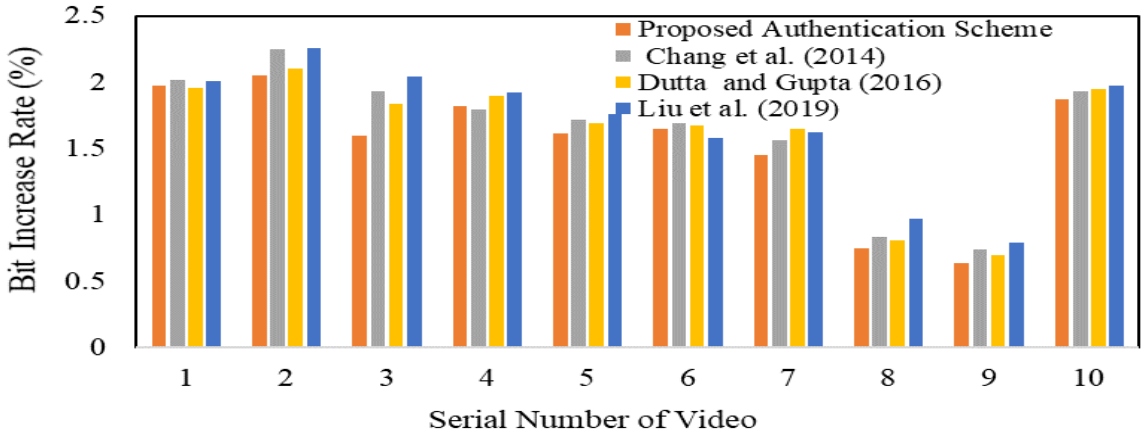


Figure 3.10: Observation appertaining to BIR of the proffered authentication scheme as well as the current schemes.

(a) Watermarked Frame

(b) Malicious Frame

1	1	1	0	0	1	0	0	0	0	1	1	1
1	1	1	0	1	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	1	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1	0	0	0

(c) Extracted AIM of malicious Frame (TP=0.201)

1	1	1	0	0	0	0	0	0	0	0	1	1
1	1	1	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0

(d) Processed AIM of malicious Frame (TP=0.115)(STP=0.571)

Figure 3.11: Adequacy of the proffered scheme appertaining to malicious maneuvering revelation for RaceHorses video

The re-compression through differing QP's, noise assaults as well as frame reduction will never affect the composition of the frame, in other words the entities in the frame remain the identical. In attempt to investigate the efficacy of the recommended authentication scheme for distinguishing between deliberate rigging and inadvertent attacks, the deliberate tweaking of the watermarked frame has been conducted. For this, one object has been purposefully omitted from the watermarked RaceHorses video's I-frame. Figure 3.11a reveals a watermarked I-frame whereas Figure 3.11b displays this same purposefully maneuvered I-frame of the RaceHorses video. Through Figure 3.11b, manoeuvred frame is illustrated whose content is altered through deleting the horse's face from its upper left. Figure 3.11c demonstrates AIM of that same watermarked Race Horse frame spawning TP value 0.201. Figure 3.11d exemplifies the tainted part with the 3×3 sliding window wherein the count of tainted items are above 5 that is determined through refining the AIM through a 3×3 median filter. Yet in the refined AIM, the amount for STP is 0.571, which again is greater than 0.553, this value of T_λ determined through (3.8) and (3.9) collectively. It substantiates, the fact that this frame has been maliciously maneuvered.

3.4 Conclusion of the Chapter

This chapter proposes a semi-fragile method of authentication for HEVC encoded videos. The proffered method focuses mostly on the QDST coefficients of 4×4 intra-luma block. In order to accomplish the admissible robustness towards re-compression, the amplitudes of the QDST coefficients are used for spatial exploration. Besides this, the proffered authentication scheme will be fragile for any malicious manoeuvrings in the video, particularly it will recognise the removal and addition of objects. Experiments are conducted on videos of varying textures and resolutions to test the efficiency of the scheme. The average cutback appertaining to $PSNR$ for the proffered scheme is 0.1253 dB and for Chang et al. (2014), Dutta and Gupta (2016) and Liu et al. (2019) schemes is 0.1874dB, 0.2298 dB and 0.1982 dB respectively. The mean BIR of the proffered scheme, Chang et al. (2014), Dutta and Gupta (2016), Liu et al. (2019) is 1.54%, 1.65%, 1.63% and 1.69% respectively. Moreover, the acquirement by the proffered authentication scheme for re-compression by divergent QP values is 0.8079 and then for Chang et al. (2014), Dutta and Gupta (2016) and Liu et al. (2019) techniques being 0.5795, 0.5274 and 0.5875 respectively. The above experimental findings indicate further that proffered authentication method is far more powerful than any of the state-of-the-art schemes in context to imperceptible deterioration in video quality and nominal escalation in bitrate and robustness towards re-encoding.

Chapter 4

An I-Frame Based Protection Scheme for Enhanced HEVC

The compression proficiency of HEVC is further improved by amending new features in the standard to encounter the problem of bandwidth required for transmission of the huge amount of video content over the Internet. In enhanced HEVC, EMT *i.e.* different variations of DCT, as well as the DST transforms, has been employed for energy compaction of residual coefficients into transform coefficients depending on the prediction mode used by the block. There are substantial divergences between the peculiarities of DST and DCT transformed coefficients in terms of stability and the amount of energy contained in each coefficient. In the case of DCT, the first coefficient is the DC coefficient, which contains the maximum energy, and stability of the coefficients decreases from low to high frequency. On the other hand, the DST generated coefficients do not follow this phenomenon of stability and energy compaction. Consequently, the DST transform based scheme proposed in Chapter 3 does not give optimal results for enhanced HEVC. So, in this chapter, an I-frame based protection scheme is proposed for enhanced HEVC[2] which employs the properties of both DST and DCT transforms used in EMT.

The schemes in the existing literature are investigated in Section 4.1. Then the overview of the EMT used in I-frames of the enhanced HEVC standard is presented in Section 4.2. Section 4.3 will provide a comprehensive description for the proposed I-frame protection scheme. The Section 4.4 articulates the experimental findings and the Section 4.5, summarises this chapter.

4.1 Introduction

The following text analyses the applicability of the existing schemes for HEVC format which have been augmented with innovative features. The watermarking method of Chang et al. (2014) is implemented by disrupting DCT / DST coefficients on the basis

[2]The contents of this chapter are published in Gagandeep Kaur, Singara Singh Kasana and M. K. Sharma. "An efficient watermarking scheme for enhanced high efficiency video coding/H.265." *Multimedia Tools and Applications*, Springer, Vol 78(9), pp. 12537–12559 (2019)

of their characteristics, without error circulation to blocks bordering the current block. Albeit the aforementioned method delivers a significant watermark ingraining ability, however, is not resilient towards diverse attacks. The watermark in the I-frame based method Dutta and Gupta (2016) has been inserted in the luma category block with dimension 4×4 via revamping the DCT coefficients with non-zero intensities. It is resilient to diverse signal manipulation as well as encoding attacks. The major drawback pertaining to the above approach is just that it incorporates the peculiarities of the DCT spawned coefficients. The (Dutta and Gupta 2017) framework included the watermark in the very first two DCT coefficients with non-zero intensities in P-frames' blocks with the dimensions 4×4. It is resilient to diverse signal modulation as well as the re-compression attempts. The key drawback of the above-mentioned approach is that it employs the peculiarities of only DCT spawned coefficients.

The watermark in the Gaj et al. (2017) scheme is embedded through adjusting the chosen subgroup of the coefficients spawned via DST to bypass the problem of distortion circulated to the surrounding region. This system is resilient to numerous attacks, have indistinguishable ocular quality as well as the retrenched upsurge in the bit rate. The biggest drawback concerning to the system is that it employs the peculiarities of one transform specifically DST. The watermark ingraining process employed by the Liu et al. (2019) fixes the parity of the specific subgroups of the DST coefficients that have been chosen depending upon the intra-prediction mode wielded in the 4×4 block. Albeit the scheme has a high capacity for embedding with a reasonable deterioration in ocular quality but it is vulnerable to numerous attacks.

According to the aforementioned study, it can be concluded that the existing techniques Chang et al. (2014); Dutta and Gupta (2016, 2017); Gaj et al. (2017); Liu et al. (2019) are not functional for enhanced HEVC standards owing to the introduction of additional features to the standard.There are drawbacks in state-of-the-art watermarking techniques with context with enhanced HEVC codec owing to variation in formats. The above-mentioned variations serve a crucial function for the accuracy concerning the watermarking frameworks. The strategies of Gaj et al. (2017); Liu et al. (2019) employ the coefficients spawned through DST_{VII} on the other hand the framework of the Dutta and Gupta (2016) wields DCT_{II} transform for implanting the watermark. In contrast to this, the augmented HEVC wields the process of EMT specifically the transforms derived from the diverse variants of DCT as well as the DST (DCT_{II}, DCT_{V}, DCT_{VIII}, DST_{I} as well as the DST_{VII}). This has contributed to the inception of a new I-frame protection scheme in which diversified variants of the DCT, as well as DST, transforms, have been employed to produce optimum outcomes with the improved HEVC format.

For resolving the constraints within the prevailing literature, a robust I-frame protection scheme for enhanced HEVC has been introduced in this chapter. The recommended I-frame protection scheme employs the peculiarities of the DCT as well as the DST transforms utilized within the residual block. The recommended method employs the luma category 4×4 blocks present in the I-frames for ingraining the watermarks since I-frames provide critical information in comparison with the P-frames. The ingraining of the watermark embedding via employing the relation among the intensities of the selected subgroup of the transform coefficients within a luma category block of dimension 4×4.The robustness of the proffered I-frame protection scheme has been realized via employing the symmetric transform depending upon the transformation utilized in the residual block. The Indistinguishable ocular quality has been realized through implanting the watermark in a 4×4 block with plenty of detail as well as by maintaining the transform coefficient sign while insertion of the watermark. The accumulation of the error within the current block as well as transmission of the produced error to the nearby block in the current I-frame has been avoided via choosing a subgroup of the transformation coefficients that are not being wielded by the adjoining block for the intra-prediction process. Besides this, the dissemination of the inter-drift error has been avoided, through the ingraining of the watermark before entropy coding throughout the encoding of the video. Furthermore, the confidentiality of the proffered I-frame protection scheme is enforced through a pseudo-random function that has been employed for choosing blocks while ingraining of the watermark. The prevention of the de-synchronization issues is attained through employing a palette in which the positions of the blocks are penned throughout the process of the watermark ingraining and then it has been utilized at the decoder side for retrieving the watermark.

4.2 Overview of EMT for I-Frames

EMT has been used for effectively loading residual coefficient energy into less non-zero transformcoefficients. In I-frames DCT_{II}, DCT_{V}, DCT_{VIII}, DST_{I} and DST_{VII} have been utilized relying upon angular mode employed by the the intra-prediction process. These are employed to capture the extremely complex peculiarities of the realistic video components, such as the edges together with the curves (Saxena and Fernandes 2013; Zhao et al. 2016).

The $R_{n\times n}$ is the matrix of residual coefficients, $T_{n\times n}$ is a matrix of final transformed coefficients, $H_{n\times n}$ is a horizontal transformation kernel, $V_{n\times n}$ is a vertical transformation kernel, and $V^{T}_{n\times n}$ is the transpose of $V_{n\times n}$ where $n \times n$ is the dimension of the block which ranges from 64×64 to 4×4. The DCT as well as the DST kernels are orthogonal,

Table 4.1: Pre-designated horizontal as well as the vertical transform sets with respect to intra prediction mode

Intra Mode	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
H	2	1	0	1	0	1	0	1	0	0	0	0	0	1	0	1	0	1
V	2	1	0	1	0	1	0	1	2	2	2	2	2	1	0	1	0	1

Intra Mode	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
H	0	1	0	1	0	1	2	2	2	2	2	1	0	1	0	1	0
V	0	1	0	1	0	1	0	0	0	0	0	1	0	1	0	1	0

consequently the $H^{-1}_{n\times n}$ will be equivalent to the $H^{T}_{n\times n}$ as well as the $V^{-1}_{n\times n}$ will be equivalent to the $V^{T}_{n\times n}$. The $T_{n\times n}$ as well as the $R_{n\times n}$ matrices could be attained through the following equations:

$$\boldsymbol{H}_{n\times n} \times \boldsymbol{R}_{n\times n} \times \boldsymbol{V}^{T}_{n\times n} = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{21} & T_{22} & \cdots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1} & T_{n2} & \cdots & T_{nn} \end{bmatrix} \tag{4.1}$$

$$\boldsymbol{H}^{T}_{n\times n} \times \boldsymbol{T}_{n\times n} \times \boldsymbol{V}_{n\times n} = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ R_{21} & R_{22} & \cdots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \cdots & R_{nn} \end{bmatrix} \tag{4.2}$$

In (4.1) and (4.2) the kernels $H_{n\times n}$ and $V_{n\times n}$ are generated using (4.3) to (4.7) in which i as well as the j varies starting from *0* to *N-1* together with the value of N is equal to n as given in (Zhao et al. 2016). The $R_{n\times n}$ matrix contains the residual error spawned through the augmented HEVC's angular intra prediction method. The Tables 4.1 as well as 4.2 predefined by Zhao et al. (2016) gives the set of transforms that are chosen depending on the intra prediction mode employed within the block. The choice among the DCT_{II} and transform sets provided in the Table 4.2 done depending on the rate-distortion expense appertaining to each block.

Table 4.2: Pre-designated set of transform

Transform Set	Transform Candidate 1	Transform Candidate 2
0	DST_{VII}	DCT_{VIII}
1	DST_{VII}	DST_{I}
2	DST_{VII}	DCT_{V}

$$DCT_{II} = \Phi_0 \times \sqrt{\frac{2}{N}} \times \cos\left(\frac{\pi \times i \times (2j+1)}{2N}\right)$$
$$where \quad \Phi_0 = \begin{cases} \sqrt{\frac{2}{N}} & \text{if } i = 1 \\ 1 & \text{if } i \neq 1 \end{cases} \tag{4.3}$$

$$DCT_{V} = \Phi_1 \times \Phi_2 \times \sqrt{\frac{2}{2N-1}} \times \cos\left(\frac{2\pi \times i \times j}{2N-1}\right)$$
$$where \quad \Phi_1 = \begin{cases} \sqrt{\frac{2}{N}} & \text{if } i = 1 \\ 1 & \text{if } i \neq 1 \end{cases} \tag{4.4}$$
$$and \quad \Phi_2 = \begin{cases} \sqrt{\frac{2}{N}} & \text{if } j = 1 \\ 1 & \text{if } j \neq 1 \end{cases}$$

$$DCT_{VIII} = \sqrt{\frac{4}{2N+1}} \times \cos\left(\frac{\pi \times (2i+1) \times (2j+1)}{4N+2}\right) \tag{4.5}$$

$$DST_{I} = \sqrt{\frac{2}{N+1}} \times \sin\left(\frac{\pi \times (i+1) \times (j+1)}{N+1}\right) \tag{4.6}$$

$$DST_{VII} = \frac{2}{\sqrt{2N+1}} \times \sin\left(\frac{\pi \times (2i-1) \times (j)}{2N+1}\right) \tag{4.7}$$

4.3 Proposed I-Frame Protection Scheme

An EMT dependent I-frame protection scheme is introduced for the enhanced HEVC standard in the current section. The comprehensive process appertaining to watermark

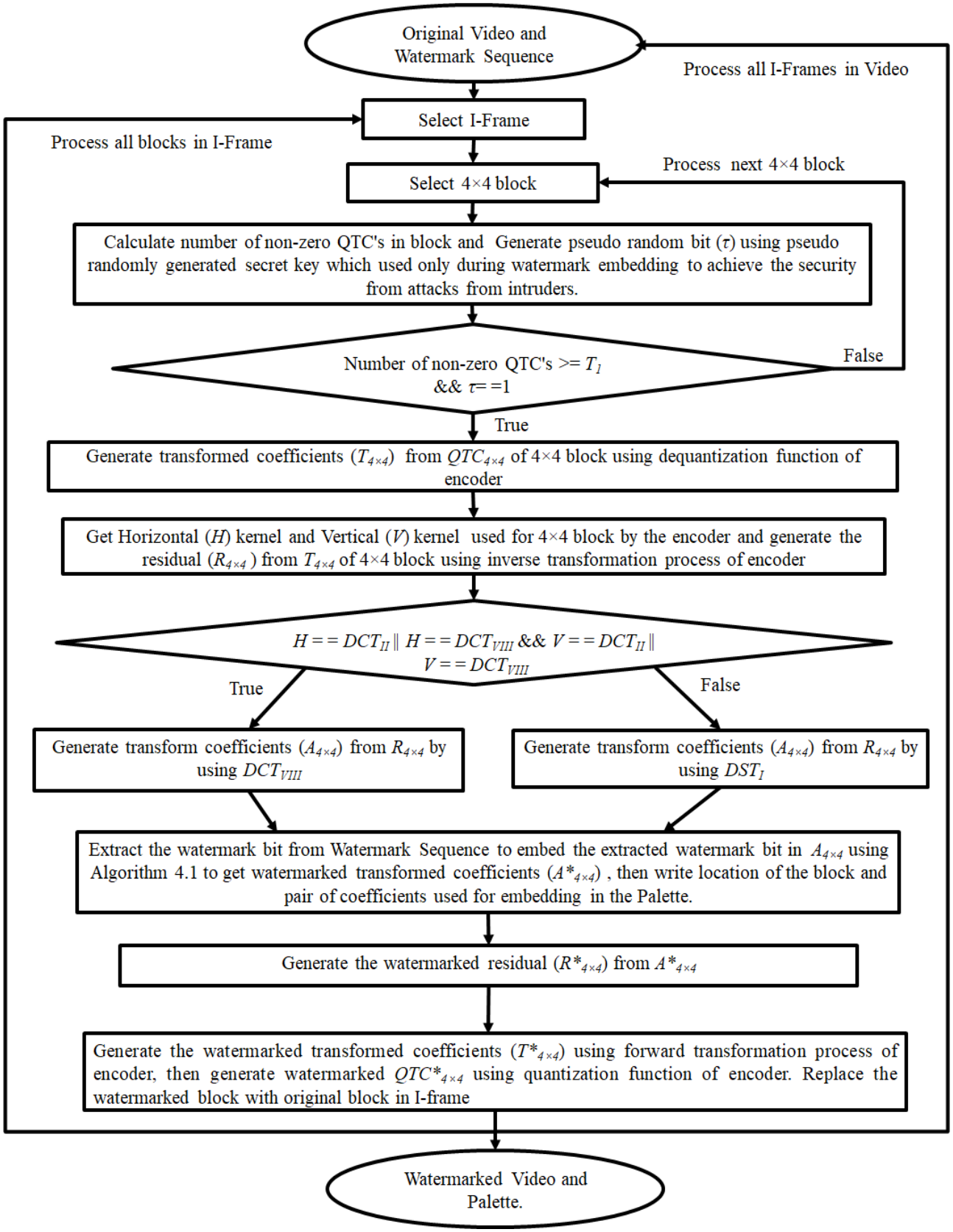


Figure 4.1: Flow chart of I-Frame Protection watermark embedding scheme

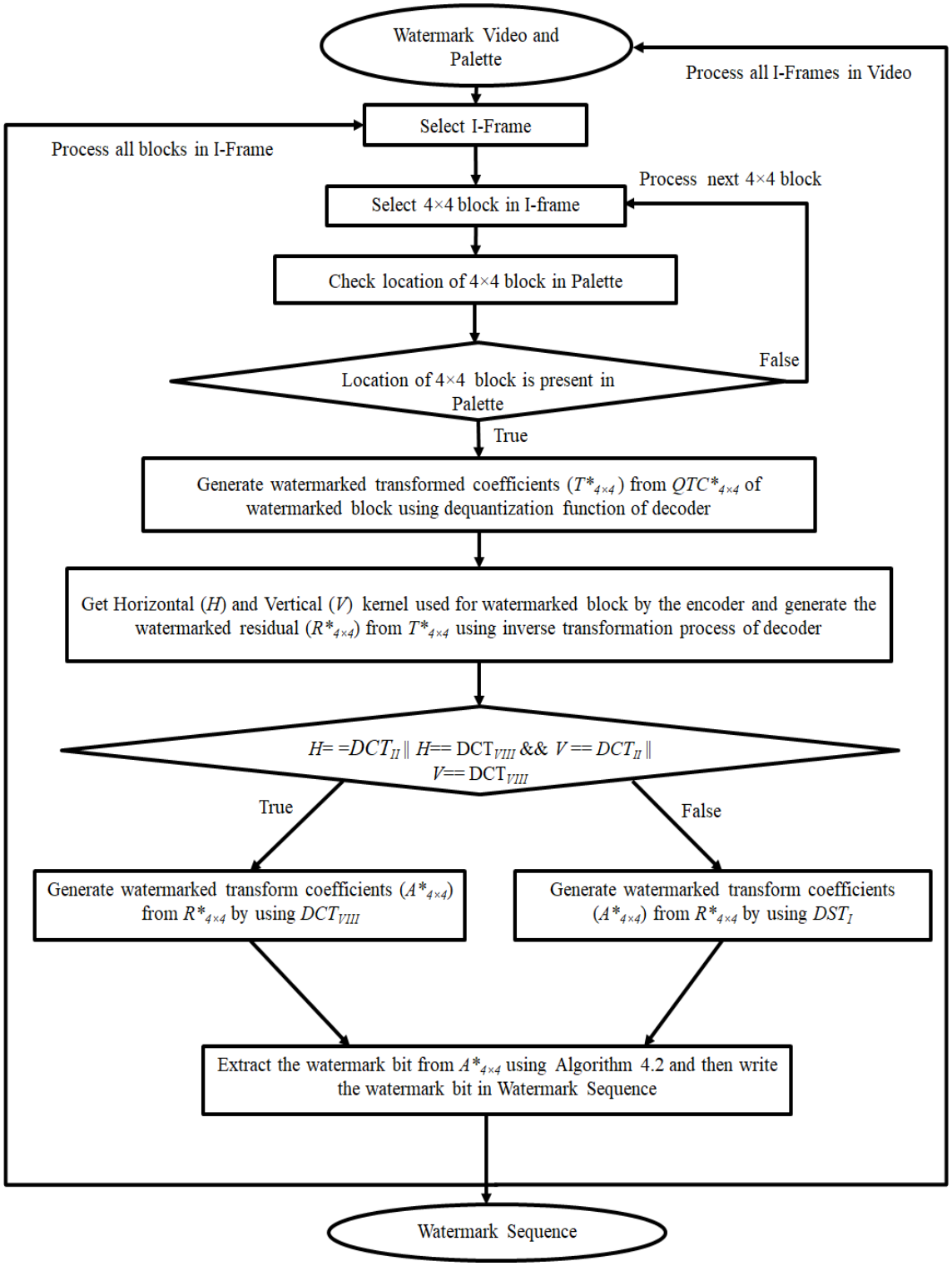


Figure 4.2: Flow chart of I-Frame Protection watermark extraction scheme

embedding together with extraction regarding the suggested I-frame protection scheme has been exemplified in Figures 4.1 and 4.2. The I-frames have been used for embedding watermarks in the present scheme considering I-frames provide a significant volume of details in contrast to inter-frames. An I-Frame could be segregated into 4×4, 8×8, 16×16, 32×32 or 64×64 sizes depending on the its texture. The Smooth areas can be divided across blocks of size 64×64, 32×32, and 16×16, while heavily textured parts are divided into blocks of 8×8 or 4×4. For watermark ingraining, the 4×4 block would be selected since HVS is less susceptible to more textured parts (Mansouri et al. 2010).

For fulfilling the criterion of imperceptibility as well as the robustness of the I-frame protection scheme, the threshold T_1 specifically the amount of non-zero QTC throughout the 4×4 block would be used to pick the block for ingraining of the watermark. The unnoticeable ocular quality is accomplished via choosing the 4×4 blocks of with high textured parts. The texture of the block is perceived to through the non-zero QTC meaning criterion. Blocks with a reasonably dense textured parts would have a far higher non-zero QTC's.

The robustness too is determined by the amount of non-zero QTC's within a 4×4 block. Moreover, the high non-zero QTC amount indicates that there is more unusual texture throughout the block. The more erratic texture implies that the block cannot be correctly predicted by utilizing the angular intra-prediction by dividing it into larger sizes. You can forecast this form of the textured block by subdividing it into of small size particularly 4×4 and hence its chances of flipping to a larger size are considerably less. Thus, for all the 4×4 blocks within a frame, we count the non-zero QTC through the block to determine, the threshold T_1. At runtime, the value of T_1 is chosen wielding (4.8) and (4.9) through the CCDF of the amount of blocks with dimensions 4×4 with different non-zero QTC count within a I-frame (Mansouri et al. 2010). At the running time, the optimum value for T_1 would be determined depending on the textures of the frame as well as the percentage of blocks needed to incorporate the watermark. Figure 4.3 indicates the deviations in block size from 4×4 to a larger block size relative to the non-zero QTC. Figure 4.3 reveals that the rate of change dwindled with a hike in the non-zero QTC ratio.

$$\overline{F}(T_1) > \nu \tag{4.8}$$

$$\overline{F}(T_1) = P(D > T_1) = 1 - F(T_1) = \sum_{D>T_1} p(D) \tag{4.9}$$

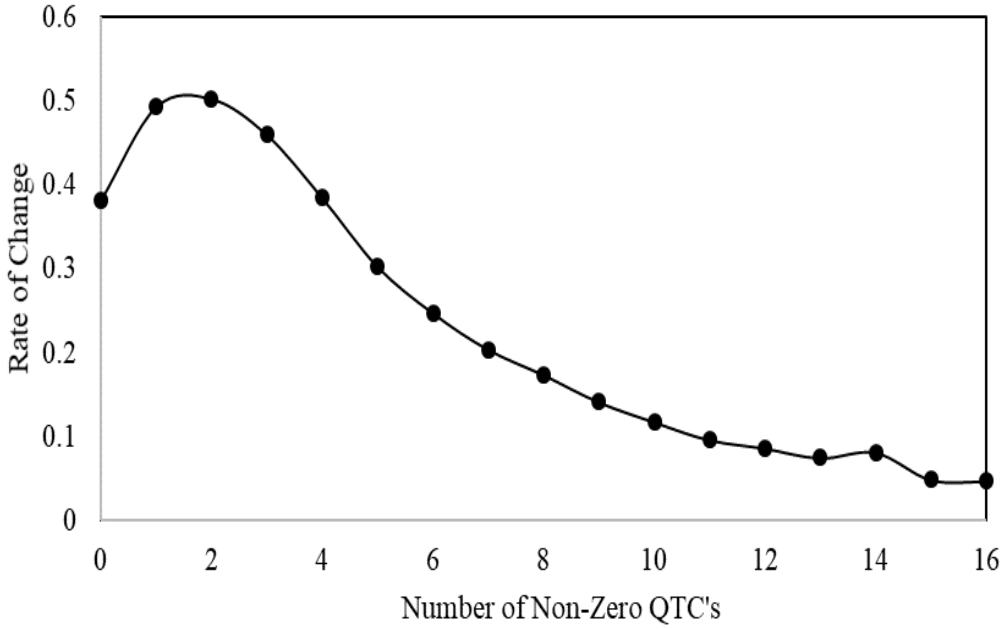


Figure 4.3: Rate of change in I-frame appertaining to deviations in the size of 4×4 to a larger size after re-encoding via Qp equivalent to 18 corresponding to non-zero QTC's

The formula of (Mansouri et al. 2010) has been employed to measure T_1. The T_1 in (4.8) as well as in (4.9) is the non-zero QTC count within 4×4 blocks moreover ν is the proportion of 4×4 blocks equivalent to T_1 amount of non-zero QTC's. Figure 4.4 and 4.5 emblazon the divergence as well as the CCDF of 4×4 blocks tallying to the non-zero QTC count in an I-frame of the videos with varied textures. Distinct T_1 values will be spawned depending on the texture of diverse video for the same value of ν. For reference, the same value of ν in Figure 4.5 specifically 0.1, the T_1 is 10 corresponding to BasketBallDrill on the other hand 9 corresponding to KristenAndSara. The optimum value of T_1 is then chosen depending on the appropriate robustness together with the number of blocks necessary for integrating the watermark.

The diverse kernels of DCT as well as DST have been employed in the EMT appertaining to process residual errors to translate into fewer coefficients. The DCT spawns a DC coefficient that is the mean of the coefficients available throughout the block, however, no such coefficient is provided by the DST. The values spawned alongside the DC coefficient are referred to as AC coefficients, whose steadiness reduces dramatically as we proceed from low frequency towards the high-frequency coefficients, though no such peculiarity is pursued by DST coefficients.

Dutta and Gupta (2016) incorporated a watermark through varying the association between the first two non-zero QTC's magnitudes within a block. The above scheme does

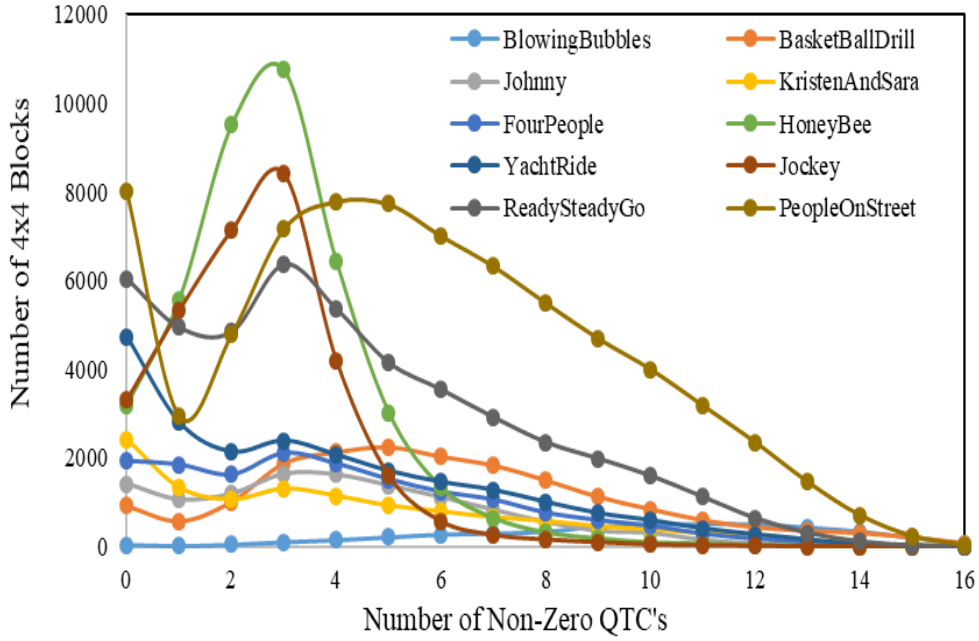


Figure 4.4: Diversities in 4×4 blocks appertaining to non-zero QTC's

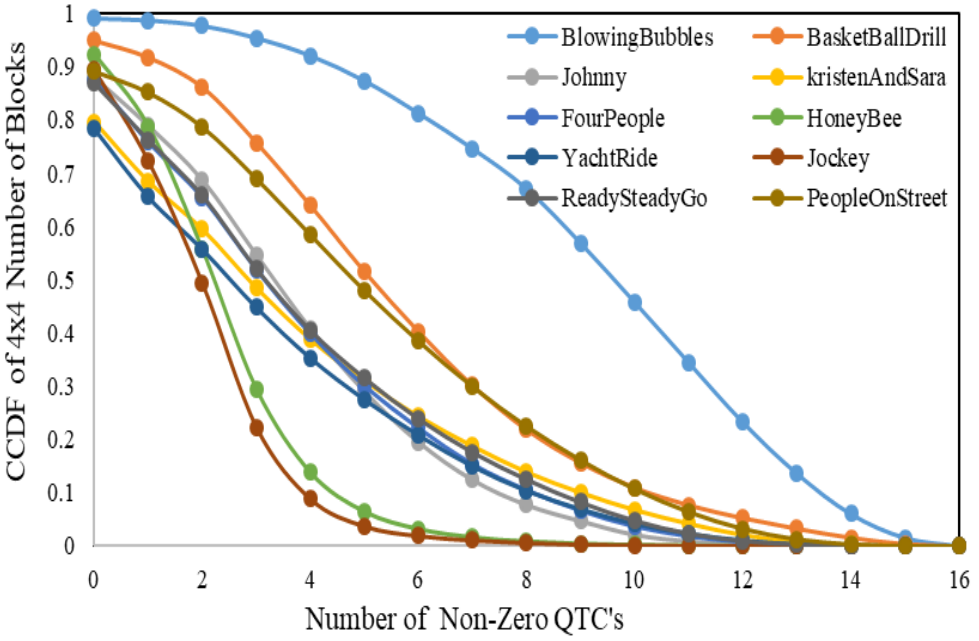


Figure 4.5: CCDF of 4×4 blocks appertaining to non-zero QTC's

not provide substantial robustness results when adapted for augmented HEVC owing to diversities between the DCT and DST peculiarities. The Gaj et al. (2017) strategy has been primarily grounded on DST transform's features, so this is not potent specifically for resilience towards attacks since the augmented HEVC employs DCT as well as the DST transformations. Furthermore, this does not have efficient ocular quality as the watermark ingraining is accomplished without retaining the signs of coefficients. A sequence of three Quantized Discrete Sine Transform Coefficients (QDST) that includes the first coefficient as well are revamped for implanting the watermark in (Liu et al. 2019). Further, the DCT together with the DST has been wielded in the EMT. In the case of the DCT, the very first coefficient is known as the DC coefficient, further minute amendment in this degrades the ocular output drastically culminating in noticeable artifacts. Moreover, the above method is not resilient towards the re-encoding attack this sequence of coefficient employed is not stable against re-encoding.

The method of watermarking developed in this study employs the peculiarities of the DCT together with the DST. The ingraining of the watermark is not explicitly done in the transformed coefficients since the transformations employed in the mechanism of EMT, specifically the DCT_{II}, DCT_{V}, and DST_{VII} are not symmetric. The adjustment in transformed coefficients which are produced through the above mentioned non-symmetric kernels induces ocular artifacts, particularly the distortion is accumulated and circulated further to the nearby areas owing to the prediction mechanism of the standard.

In the first phase, the residual coefficients are retrieved in the space domain to solve this problem. These coefficients in the next phase are revamped through the symmetric transform specifically either DCT_{VIII} or DST_{I}. The kernel spawned through either of the two above suggested symmetric transforms has coefficients in the upper triangle with identical amplitudes with the coefficients in the lower triangle with opposite signs. This specific peculiarity of the symmetric transformation mentioned above is being wielded for watermark ingraining for abolishing the accumulation of the error. Either of the two symmetric transforms available, namely DCT_{VIII} or DST_{I}, is chosen depending upon the horizontal together with the vertical transformations employed through the enhanced encoder.

$$A_{4\times4} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{bmatrix} \tag{4.10}$$

The matrix of dimension 4×4 in (4.10) contains the residual error gleaned through employing the chosen transform. From the above coefficients only two subgroups of the coefficients, namely A_{12}, A_{32} and A_{21}, A_{23}, are used for the embedding of the watermark. These subgroups are not employed for the intra-prediction mechanism for simulating the nearby blocks, therefore the error triggered through the implantation of the watermark will not be disseminated to the surrounding blocks. The specific subgroup from the two alternatives is chosen to utilize the gap among the magnitudes of the two coefficients in the subgroup. The subgroup with the minimum gap will be employed for the ingraining of the watermark bits. The minimum gap has been exploited to avoid the accumulation of error drift. Furthermore, for attaining the desired optical quality signs of coefficients are retained while altering only magnitudes because any manipulation with the signs skeptically impairs the ocular quality of the video. The modified matrix after ingraining the watermark is translated to watermarked QTC's through the augmented HEVC encoder. The information regarding the position as well as the subgroup within each block adopted for ingrain the watermark has been penned in the palette file at the end of the encoder. The file would be wielded for retrieving the watermark to prevent the de-synchronization at the end of the decoder. Hence, the recommended I-frame protection method is semi-blind, because rather than the complete un-watermarked video, particularly the palette is needed for the retrieval. The palette can be conveniently disseminated via a protected channel. The recommended I-frame protection watermark embedding algorithm is expounded in the Section 4.3.1 together with the algorithm for watermark extraction in the Section 4.3.2.

4.3.1 I-Frame Based Watermark Embedding Procedure

The series of steps involved in the watermark embedding mechanism of the recommended I-frame protection scheme are just as follows:

Step 1: Repeat Step 2 to Step 14 for each I-frame of the input video sequence.

Step 2: Repeat Step 3 to Step 14 for all 4×4 blocks in the I-frame.

Step 3: Reckon the number of non-zero QTC as well as compute pseudo-random bit τ wielding pseudo-randomly spawned secret key that is utilized particularly during embedding of the watermark.

Step 4: If $\tau == 1$ as well as number of non-zero QTC $\geq T_1$ which is computed through employing (4.8) and (4.9).

Step 5: Generate transformed coefficients $T_{4\times4}$ from the block's $QTC_{4\times4}$ utilizing the enhanced HEVC encoder de-quantization function.

Algorithm 4.1: Watermark Embedding in I-Frame

Data: Original Block $A_{4\times4}$ and Watermark Bit WB
Result: Watermarked Block $A^*_{4\times4}$, P: Pair used for embedding watermark

if *WB=1* **then**
- **if** $(|A_{12}| \geq |A_{32}|)\|(|A_{21}| \geq |A_{23}|)$ **then**
  - No change
  - **if** $(|A_{12}| \geq |A_{32}|)$ **then**
    - $P = 1$
  - **else**
    - $P = 2$
- **else**
  - $\Delta_1 = |(|A_{12}| - |A_{32}|)|$
  - $\Delta_2 = |(|A_{21}| - |A_{23}|)|$
  - **if** $(\Delta_1 \leq \Delta_2)$ **then**
    - $P = 1$
    - $Sign1 = Sign\ of\ A_{12}$
    - $A^*_{12} = Sign1 \times (|A_{12}| + \Delta_1)$
  - **else**
    - $P = 2$
    - $Sign1 = Sign\ of\ A_{21}$
    - $A^*_{21} = Sign1 \times (|A_{21}| + \Delta_2)$

else
- WB=0
- **if** $(|A_{12}| < |A_{32}|)\|(|A_{21}| < |A_{23}|)$ **then**
  - No change
  - **if** $(|A_{12}| < |A_{32}|)$ **then**
    - $P = 1$
  - **else**
    - $P = 2$
- **else**
  - $\Delta_1 = |(|A_{12}| - |A_{32}|)|$
  - $\Delta_2 = |(|A_{21}| - |A_{23}|)|$
  - **if** $(\Delta_1 \leq \Delta_2)$ **then**
    - $P = 1$
    - $Sign2 = Sign\ of\ A_{32}$
    - $A^*_{32} = Sign2 \times (|A_{32}| + (\Delta_1 + \Theta))$
  - **else**
    - $P = 2$
    - $Sign2 = Sign\ of\ A_{23}$
    - $A^*_{23} = Sign2 \times (|A_{23}| + (\Delta_2 + \Theta))$

All coefficients of $A_{4\times4}$ except pair changed for embedding are directly copied to $A^*_{4\times4}$ and Θ is constant equal to one, which is used when the value Δ_1 or Δ_2 is zero.

Algorithm 4.2: Watermark Extraction from I-Frame

Data: Watermarked Block $\boldsymbol{A}^*_{4\times4}$ and P: Pair used for embedding watermark
Result: WB: Extracted Watermarked Bit
if *P=1* **then**
  if $(|\boldsymbol{A}^*_{12}| \geq |\boldsymbol{A}^*_{32}|)$ **then**
    WB=1
  else
    WB=0
else
  if $(|\boldsymbol{A}^*_{21}| \geq |\boldsymbol{A}^*_{23}|)$ **then**
    WB=1
  else
    WB=0

Step 6: Retrieve the Horizontal (H) as well as the Vertical (V) transforms utilized in the block and then generate the residual $R_{4\times4}$ from the transformed coefficients $T_{4\times4}$ employing the enhanced HEVC encoder using (4.2).

Step 7: If $H ==$ DCT and also V $==$ DCT where DCT is $DCTII$ or $DCTV$ or $DCTVIII$, set the kernel equivalent to 1 otherwise kernel equivalent to 2.

Step 8: Generate Transform coefficients $A_{4\times4}$ from the residual $R_{4\times4}$ utilizing (4.11), if the kernel is equivalent to 1 then the kernel $K_{4\times4}$ in (4.11) is spawned by employing (4.5) by varying the value of i,j starting from 0 to N-1 where N = 4. The kernel $K_{4\times4}$ in (4.11) is spawned by employing (4.6) a varying the value of i,j starting from 0 to N-1 where N = 4. The kernel $K^T_{4\times4}$ is transpose of $K_{4\times4}$.

$$\boldsymbol{K}_{4\times4} \times \boldsymbol{R}_{4\times4} \times \boldsymbol{K}^T_{4\times4} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{bmatrix} \tag{4.11}$$

Step 9: Pick up the Watermark Bit (WB) from the Watermark Series W_m in which m represents the total of bits that are being implanted within the I-frame of the video. The W_m is provided as input to the algorithm, together with the video in which the watermark is to be ingrained.

Step 10: Insert WB in the transform coefficients $A_{4\times4}$ wielding the Algorithm 4.1 to realize the watermarked transform coefficients $A^*_{4\times4}$ furthermore record block the position

of the block along with the pair wielded in the Palette.

Step 11: Spawn the watermarked residual $R^*_{4\times4}$ from the watermarked transform coefficient ${A^*}_{4\times4}$ wielding (4.12), if the value of the kernel corresponds to 1 then the kernel $K_{4\times4}$ in (4.12) will be spawned utilizing (4.5) by varying i,j starting from 0 up to N-1 where N is equivalent to 4, The kernel $K_{4\times4}$ in (4.12) will be spawned utilizing (4.6) by varying i,j starting from 0 up to N-1 where N is equivalent to 4, $K^T_{4\times4}$ is the transpose of the kernel $K_{4\times4}$.

$$\boldsymbol{K}^T_{4\times4} \times \boldsymbol{A}^*_{4\times4} \times \boldsymbol{K}_{4\times4} = \begin{bmatrix} R^*_{11} & R^*_{12} & R^*_{13} & R^*_{14} \\ R^*_{21} & R^*_{22} & R^*_{23} & R^*_{24} \\ R^*_{31} & R^*_{32} & R^*_{33} & R^*_{34} \\ R^*_{41} & R^*_{42} & R^*_{43} & R^*_{44} \end{bmatrix} \tag{4.12}$$

Step 12: Generate watermarked transformed coefficients $T^*_{4\times4}$ from the watermarked residual $R^*_{4\times4}$ using (4.1) specifically the forward transformation function of the encoder.

Step 13: Generate watermarked $QTC^*_{4\times4}$ from $T^*_{4\times4}$ through wielding the encoder's quantization function.

Step 14: Generate watermarked blocks through reinstating $QTC_{4\times4}$ with $QTC^*_{4\times4}$.

4.3.2 I-Frame Based Watermark Extraction Procedure

The procedurefor retrieval of the watermark from each I-frame of the watermark videoare as follows:

Step 1: Repeat step 2 to Step 7 on all the4×4 blocks inthe I-frame.

Step 2: If the position of the watermarked block available within the palette then retrieve the P: Pair from the palette.

Step 3: Generate transformed watermarked coefficients ${T^*}_{4\times4}$ from $QTC^*_{4\times4}$ of the watermarked block wielding HEVC decoder'sde-quantificationfunction.

Step 4: Retrieve the horizontal (H), as well as the vertical (V), transforms utilized by the watermarked block then generate the watermarked residual ${R^*}_{4\times4}$ from the watermarked transformed coefficients $T^*_{4\times4}$ wielding (4.2) specifically the inverse transformation function of the decoder.

Step 5: If H = = DCT and alsoV = = DCT where DCT is DCT_{II} or DCT_V or DCT_{VIII}, then set the kernel =1 Otherwise kernel=2.

Step 6: Generate watermarked transform coefficients ${A^*}_{4\times4}$ from watermarked residual ${R^*}_{4\times4}$ wielding (4.11), If TransformUsed == 1 then the kernel $K_{4\times4}$ in (4.11) will be generated utilizing (4.5) by varying i,j starting from 0 to N-1 where N = 4 otherwise the kernel $K_{4\times4}$ in (4.11) will be generated utilizing (4.6) by varying i,j starting from 0 to N-1 where N = 4. Further the $K^T_{4\times4}$ in the following equation is transpose of the kernel $K_{4\times4}$.

$$\boldsymbol{K}_{4\times4} \times \boldsymbol{R^*}_{4\times4} \times \boldsymbol{K}^T_{4\times4} = \begin{bmatrix} A^*_{11} & A^*_{12} & A^*_{13} & A^*_{14} \\ A^*_{21} & A^*_{22} & A^*_{23} & A^*_{24} \\ A^*_{31} & A^*_{32} & A^*_{33} & A^*_{34} \\ A^*_{41} & A^*_{42} & A^*_{43} & A^*_{44} \end{bmatrix} \tag{4.13}$$

Step 7: Retrieve theWB employing Algorithm 4.2 and then append WB to the extracted watermark.

4.4 Experimental Results

The implementation of the introduced I-frame protection as well as the current schemes has been carried out using HM14KTA reference software of enhanced HEVC. Experiments are conducted on videos with varying textures and resolutions to test the new I-frame protection scheme. The test results are generated by employing the configuration file of intra-main profile by setting the $QP = 16$, $DecodingRefereshType = 2$, $GOPSize = 4$ and $IntraPeriod = 8$. In each I-frame 100 watermark bits are ingrained, further value of T_1 is computed at runtime by setting the value of $\nu = 0.1$, and the details of parameters for different videos are given in Table 4.3. The potency of the introduced I-frame protection scheme is assessed appertaining to robustness towards re-compression, noise and filtering attacks, frame dropping, degradation in visual quality as well as BIR. The introduced I-frame protection scheme is validated by comparing it with the cuurent schemes of Dutta and Gupta (2016), Gaj et al. (2017) as well as Liu et al. (2019). The average within Tables 4.4, 4.5 and 4.6 is the average of the outcomes gleaned for whole video set utilized for the simulations.

The robustness of the recommended I-frame protection method is contrasted with the state-of-the-art methods with respect to re-compression attack as well as the dropping of the frames for transcoding. The data for the above-methods attacks appertaining to the recommended I-frame protection scheme together with the current scheme is given in Table 4.4. The first attack particularly the re-compression has been carried out in two stages, a compressed watermarked video has been decoded first, thereafter the video is

Table 4.3: Data set of videos used for experimentation of the protection schemes

Serial Number of Video	Name of Video	Resolution	No of Frames	Frame Rate
1	BlowingBubbles	416×240	500	50
2	BasketBallDrill	832×480	500	50
3	Johnny	1280×720	600	60
4	KristenAndSara	1280× 720	600	60
5	FourPeople	1280×720	600	60
6	HoneyBee	1920×1080	600	120
7	YachtRide	1920×1080	600	120
8	Jockey	1920×1080	600	120
9	ReadySteadyGo	1920× 1080	600	120
10	PeopleOnStreet	2560× 1600	150	30

compressed again with separate QP values specifically 18, 14 and16. The strategy of the (Dutta and Gupta 2016) has been founded on the peculiarities of the DCT coefficients. The strategies of the Gaj et al. (2017) as well as of Liu et al. (2019) are dependent mostly on the peculiarities of DST coefficients. The augmented HEVC format, however, wields EMT for the residual error coding that specifically employs the diverse kernels of DCT as well as DST. So, established on the properties of both DCT as well as the DST, the introduced I-frame protection scheme functions accuratelier than the current schemes of (Dutta and Gupta 2016; Gaj et al. 2017; Liu et al. 2019).

The mean of the correlation coefficient with regards to Qp=14 for the recommended I-frame protection scheme as well as for the Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.8819, 0.5407, 0.8492 and 0.4111 respectively. The mean of the correlation coefficient with regards to Qp=16 for the recommended I-frame protection scheme as well as for the Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.9619, 0.8066, 0.9432 and 0.4694 respectively. The mean of the correlation coefficient with regards to Qp=18 for the recommended I-frame protection scheme as well as for the Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.8363, 0.5580, 0.7992 and 0.4005 respectively. The overall mean of the correlation coefficient with regards to all the QP specifically 14, 16 and 18 for the recommended I-frame protection scheme as well as for the Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are are 0.8934, 0.6351, 0.8639 and 0.4270 respectively.

The strategy of the Li et al. (2015) has been wielded for accomplishing an indiscriminate frame dropping attack. The binary frame number has been ingrained antecedent to implanting the actual watermark bits. The process of retrieval starts with extracting the ingrained frame number to classify the frame and thereafter the actual watermarked is retrieved. The overall mean of correlation coefficient particularly with respect to the 25 percent arbitrary frame dropping for the newly introduced I-frame protection scheme, Dutta and Gupta (2016), Gaj et al. (2017) as well as for the Liu et al. (2019) are 0.7920, 0.3440, 0.7524 and 0.3391 respectively. Hence, it explicates that the recommended I-frame protection scheme is far more potent against the re-encoding as well as frame dropping appertaining to the current schemes scrutinized for comparison.

The recommended I-frame protection scheme has been additionally tested for diverse video-processing attacks specifically Gaussian noise, salt and pepper noise, median and Gaussian filtering. Table 4.5 emblazons the overall correlation coefficient among the ingrained as well as the recovered watermark after the diversified video manipulation attacks, including Gaussian noise with average and variance equivalent to zero and 0.001 respectively, salt and pepper noise at density 0.01, 3×3 median filter and Gaussian filter with dimensions 3×3 with 0.3 value of sigma. The overall correlation coefficient appertaining to Gaussian noise for the I-frame protection scheme as well as for the Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.7649, 0.4948, 0.7467 and 0.4266 respectively. The overall correlation coefficient appertaining to salt and pepper noise for the I-frame protection scheme as well as for the Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.7538, 0.468, 0.7263 and 0.438 respectively. The overall correlation coefficient appertaining to median filter for the I-frame protection scheme as well as for the Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.8103, 0.5225, 0.7769 and 0.4923 respectively. The overall correlation coefficient appertaining to Gaussian filter for the I-frame protection scheme as well as for the Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.7826, 0.4996, 0.7527 and 0.4563 respectively. The afore-mentioned outcomes prove that the recommended I-frame protection scheme resilient towards the diverse video manipulation attacks and hence are superior to the prevailing schemes that have been scrutinized for comparison.

The Ocular output has been assessed through the subjective aspect as well as through the quantitative measurements covering the $VIFp$, $SSIM$, and $PSNR$. Figure 4.6 substantiates that there have been no optical glitches throughout thewatermarked video frames and perhaps an indiscernible difference in the quality of video frames following watermark embedding. Table 4.6 exhibits the simulation results of the loss of ocularquality after integrating thewatermark. The overall mean of δ_{PSNR} of the new I-frame protectionsystem

Table 4.4: Comparison of the correlation coefficient of the proposed I-frame protection scheme with existing schemes for re-compression and frame dropping attack

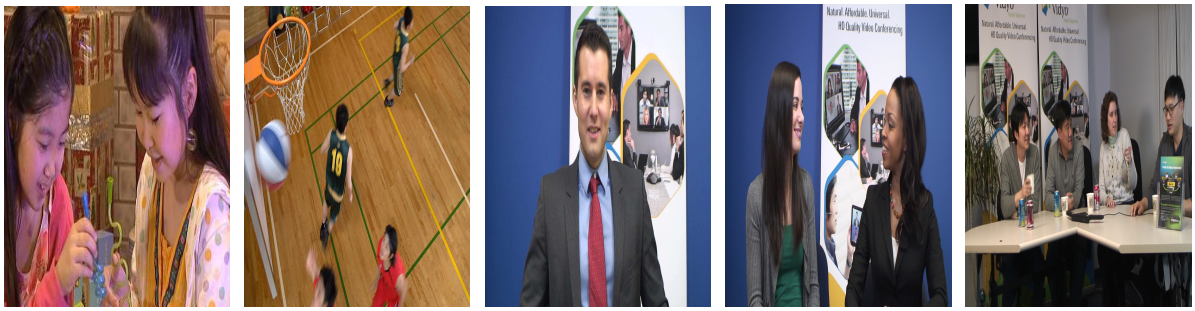
Video Sequence	Proposed I-Frame Protection Scheme				Dutta and Gupta (2016)				Gaj et al. (2017)				Liu et al. (2019)			
	Re-compression			Frame	Re-compression			Frame	Re-compression			Frame	Re-compression			Frame
	QP=14	QP=16	QP=18	Dropping	QP=14	QP=16	QP=18	Dropping	QP=14	QP=16	QP=18	Dropping	QP=14	QP=16	QP=18	Dropping
BlowingBubbles	0.8516	0.9609	0.7837	0.7908	0.8046	0.9531	0.7256	0.7543	0.5365	0.7948	0.4399	0.3802	0.3959	0.4431	0.3742	0.3384
BasketBallDrill	0.8997	0.9766	0.7608	0.7738	0.8725	0.9695	0.7165	0.7343	0.6506	0.8962	0.4732	0.3262	0.4199	0.4316	0.3627	0.3299
Johnny	0.8846	0.9715	0.7949	0.8278	0.8369	0.9283	0.7462	0.7911	0.5143	0.8966	0.5559	0.4142	0.4124	0.4487	0.3798	0.3569
KristenAndSara	0.8593	0.9781	0.8842	0.7398	0.8449	0.9605	0.8492	0.7078	0.4996	0.8868	0.5631	0.3822	0.3998	0.4933	0.4244	0.3129
FourPeople	0.8754	0.9205	0.8589	0.7868	0.8599	0.9061	0.8348	0.7495	0.5089	0.7718	0.6885	0.3662	0.4078	0.4807	0.4118	0.3364
HoneyBee	0.8675	0.9549	0.8171	0.8155	0.8147	0.9336	0.7895	0.7737	0.4452	0.5176	0.5047	0.3413	0.4039	0.4598	0.3909	0.3508
YachtRide	0.9229	0.9722	0.8999	0.8093	0.9038	0.9505	0.8602	0.7643	0.5951	0.8171	0.5895	0.2964	0.4316	0.5012	0.4323	0.3477
Jockey	0.8914	0.9419	0.8429	0.8048	0.8529	0.9264	0.8052	0.7626	0.5857	0.7684	0.5884	0.3257	0.4158	0.4727	0.4038	0.3454
ReadySteadyGo	0.8672	0.9776	0.8819	0.7393	0.8289	0.9524	0.8602	0.7019	0.5353	0.8508	0.5972	0.3175	0.4037	0.4922	0.4233	0.3127
PeopleOnStreet	0.8998	0.9652	0.8382	0.8323	0.8728	0.9513	0.8041	0.7843	0.5357	0.8659	0.5799	0.2903	0.4205	0.4703	0.4014	0.3592
Average	0.8819	0.9619	0.8363	0.792	0.8492	0.9432	0.7992	0.7524	0.5407	0.8066	0.558	0.344	0.4111	0.4694	0.4005	0.3391

Table 4.5: Comparison of the correlation coefficient of the proposed I-frame protection scheme with existing schemes for noise and filtering attacks

Video Sequence	Proposed I-Frame Protection Scheme				Dutta and Gupta (2016)				Gaj et al. (2017)				Liu et al. (2019)			
	Gaussian Noise	Salt and Pepper Noise	Median Filter	Gaussian Filter	Gaussian Noise	Salt and Pepper Noise	Median Filter	Gaussian Filter	Gaussian Noise	Salt and Pepper Noise	Median Filter	Gaussian Filter	Gaussian Noise	Salt and Pepper Noise	Median Filter	Gaussian Filter
BlowingBubbles	0.7609	0.7416	0.7326	0.7123	0.7449	0.6551	0.7191	0.6857	0.4959	0.4263	0.4716	0.4544	0.4206	0.4045	0.4432	0.4272
BasketBallDrill	0.7589	0.6643	0.7185	0.6733	0.7392	0.6476	0.6987	0.6604	0.4749	0.4168	0.4348	0.4356	0.4196	0.3819	0.4284	0.4077
Johnny	0.7697	0.7429	0.7964	0.7624	0.7508	0.6986	0.7824	0.7219	0.5867	0.4257	0.4829	0.4574	0.4651	0.4296	0.4953	0.4522
KristenAndSara	0.7673	0.7286	0.7847	0.7585	0.7421	0.6789	0.7993	0.6996	0.4903	0.4365	0.4963	0.4607	0.4438	0.4275	0.4795	0.4503
FourPeople	0.7739	0.7548	0.8456	0.8198	0.7624	0.7812	0.8351	0.8036	0.4589	0.4545	0.5189	0.4898	0.4271	0.4691	0.5156	0.4809
HoneyBee	0.7736	0.7934	0.8938	0.8825	0.7549	0.8243	0.8694	0.8456	0.5686	0.5213	0.6232	0.5995	0.4271	0.4852	0.5532	0.5123
YachtRide	0.7508	0.8078	0.8175	0.7721	0.7394	0.7014	0.7704	0.7567	0.4188	0.4636	0.5298	0.4881	0.4155	0.4246	0.4926	0.4571
Jockey	0.7684	0.7719	0.8712	0.8582	0.7493	0.7562	0.7649	0.7906	0.5124	0.5244	0.5574	0.5472	0.4743	0.4865	0.5346	0.5001
ReadySteadyGo	0.7754	0.7934	0.8352	0.8161	0.7606	0.7945	0.8276	0.8055	0.4464	0.5182	0.5615	0.5491	0.3978	0.4483	0.5087	0.4791
PeopleOnStreet	0.7506	0.7391	0.8075	0.7706	0.7236	0.7247	0.7025	0.7576	0.4946	0.4926	0.5489	0.5146	0.4054	0.4272	0.4715	0.4563
Average	0.7649	0.7538	0.8103	0.7826	0.7467	0.7263	0.7769	0.7527	0.4948	0.468	0.5225	0.4996	0.4266	0.4384	0.4923	0.4563

Table 4.6: Comparative investigation of the deterioration in the ocular quality of the proposed I-frame protection scheme with current schemes

Video Sequence	Proposed I-Frame Protection Scheme			Dutta and Gupta (2016)			Gaj et al. (2017)			Liu et al. (2019)		
	δ_{PSNR}(dB)	$SSIM$	$VIFp$	δ_{PSNR}(dB)	$SSIM$	$VIFp$	δ_{PSNR}(dB)	$SSIM$	$VIFp$	δ_{PSNR}(dB)	$SSIM$	$VIFp$
BlowingBubbles	0.1079	0.99902	0.94365	1.6355	0.99801	0.90553	0.1565	0.99853	0.91548	1.9225	0.99752	0.90185
BasketBallDrill	0.1204	0.99909	0.97128	1.8418	0.99846	0.95279	0.1774	0.99863	0.95692	2.0678	0.99789	0.93843
Johnny	0.1354	0.99969	0.9727	2.0536	0.99904	0.93683	0.1991	0.99953	0.95905	2.2826	0.99888	0.92318
KristenAndSara	0.1211	0.99994	0.98998	1.4579	0.99983	0.95957	0.1911	0.99991	0.98497	1.6369	0.99975	0.95456
FourPeople	0.1416	0.99934	0.97836	1.6832	0.99828	0.95191	0.1976	0.99902	0.96754	1.8232	0.99796	0.94109
HoneyBee	0.1499	0.99881	0.96286	2.2334	0.99536	0.92013	0.2161	0.99822	0.94429	2.5404	0.99477	0.90156
YachtRide	0.1309	0.99887	0.97175	2.0363	0.99586	0.92742	0.2052	0.99831	0.95762	2.1753	0.99531	0.91329
Jockey	0.1439	0.99844	0.96773	2.0462	0.99552	0.93712	0.2066	0.99766	0.95159	2.1852	0.99474	0.92098
ReadySteadyGo	0.1325	0.99945	0.97364	1.1242	0.99897	0.95037	0.1994	0.99918	0.96046	1.2632	0.99872	0.93719
PeopleOnStreet	0.0944	0.99976	0.97637	1.0939	0.99035	0.95443	0.1496	0.99964	0.96456	1.2329	0.99023	0.94262
Average	0.1278	0.99924	0.97083	1.7206	0.99696	0.93961	0.1899	0.99886	0.95625	1.9130	0.99658	0.92748



(a) Original (b) Original (c) Original (d) Original (e) Original

(f) Watermarked (g) Watermarked (h) Watermarked (i) Watermarked (j) Watermarked

(k) Original (l) Original (m) Original (n) Original (o) Original

(p) Watermarked (q) Watermarked (r) Watermarked (s) Watermarked (t) Watermarked

Figure 4.6: The Original as well as the watermarked I-frames for all the videos in data set with QP value 16

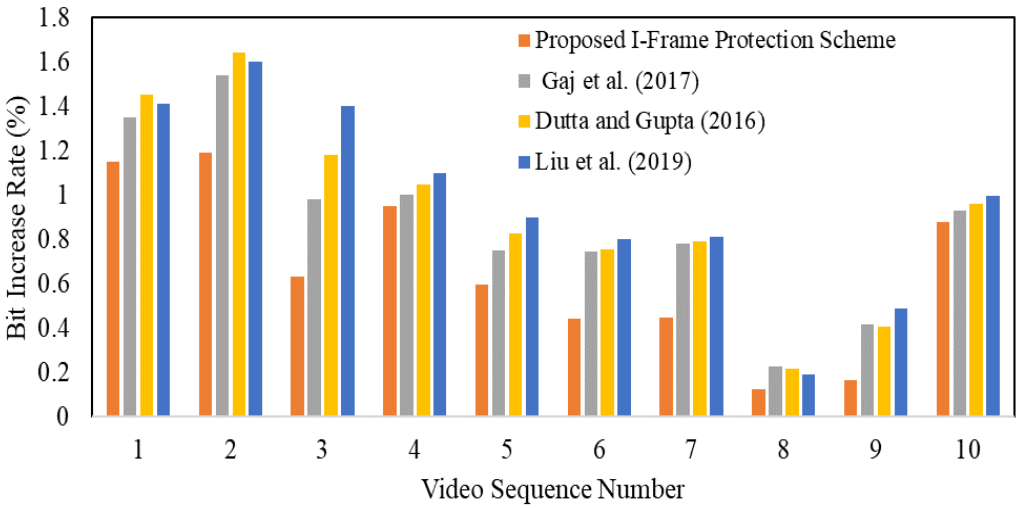


Figure 4.7: Comparative analysis of BIR for the proposed I-frame protection scheme versus current schemes

as well as for theDutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.1278 dB, 1.7206dB, 0.1899 dB and 1.9130 dB respectively. The overall mean of $SSIM$ for the new I-frame protection schemeas well as for theDutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.99924, 0.99696, 0.99886 and 0.99658 respectively. The overall mean of$VIFp$ for the new I-frame protection scheme as well as for theDutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.97083, 0.93961, 0.95625 and 0.92748, respectively. The findings in the table 4.6 indicate that the productivity of the new I-frame protectionscheme is superior to that of the Dutta and Gupta (2016); Gaj et al. (2017); Liu et al. (2019) in measures of $SSIM$, $VIFp$ as well as theδ_{PSNR}. The new I-frame protection scheme performs farbetter than the Dutta and Gupta (2016); Gaj et al. (2017); Liu et al. (2019) since just the intensities of both the transformed coefficients have beenmodified thereby retaining the signs.

In the recommended I-frame protection method, the watermark is implanted in 4×4 blocks having an abundant count of non-zero QTC's, retrenching rise in the bit-rate of the encoded video after imbuing of the watermark. The potency of the new I-frame protection scheme has been computed with regards to BIR. Figure 4.7 reveals the BIR of the new I-frame protection scheme as well as the current schemes of (Dutta and Gupta 2016; Gaj et al. 2017; Liu et al. 2019). The video number in Figure 4.7 corresponds to the number in the Table 4.3. The mean BIR for the new I-frame protection system as well as for the Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.7%, 1.10%, 0.9% and 1.15% respectively. Thereupon, it could be derived that the potency of the new I-frame protection proposed strategy is better with regards to the current techniques of

Dutta and Gupta (2016); Gaj et al. (2017); Liu et al. (2019), considering transformed coefficients magnitudes have been tweaked albeit their signs have been retained.

In the proposed I-frame protection scheme, a watermark has been inserted in the 4×4 blocks. Let us assume the total number of I and P frames in the video are i and p respectively. Further, the number of 4×4 blocks within each I-frame varies from one video to another. Therefore, we can assume the number of 4×4 blocks in a video to be n. The proposed watermark embedding method processes each 4×4 to count the non-zero QTC's to select the block for inserting the watermark bit. The watermark bit is embedded into the selected block in constant time which is not dependent on the number of blocks in an I-frame. Therefore, the worst-case complexity of the proposed I-frame protection scheme in terms of Big-O notation is O($i \times n$).

The Gaj et al. (2017) method firstly processes each 4×4 block to select and embed the watermark. Therefore, the complexity of this process is O($i \times n$). In the second step, in this method, a drift compensation is employed as signs of the coefficients are also altered during the embedding of the watermark. The drift compensation procedure is employed in two parts. In the first part, the homogeneous regions are identified from the prediction blocks present in the current I-frame. Let the number of prediction blocks within the I-frame are m. In the next part, the drift compensation is done in the P-frames corresponding to the m prediction blocks of the I-frames. So the complexity of the Gaj et al. (2017) scheme can be written as O($i \times n + i \times m + p \times m$). As we know, we can approximate the complexity to the value which is maximum. In the above-calculated complexity, the $p \times m$ is a maximum value because the number of P-frames in a video are always greater than I-frame and also the number of prediction blocks are greater than the 4×4 blocks within an I-frame. Therefore, the approximated complexity of the Gaj et al. (2017) scheme is O($p \times m$). Similarly, the Dutta and Gupta (2016) scheme first refines the motion vector of all the prediction block in the P-frames. Then, it estimates the pseudo motion vector of the blocks in I-frames and selects the 4×4 blocks for embedding the watermark. Therefore, the total complexity of the Dutta et scheme is O($p \times m + i \times n$). The approximated complexity of the Dutta and Gupta (2016) scheme is O($p \times m$). Lastly, the Liu et al. (2019) scheme first processes all the prediction block in I-frames to extract the intra prediction mode of the blocks. Then the watermark embedding method is employed to embed the watermark bits based on the intra prediction modes of the $4 \times$ block. Hence, the complexity of the Liu et al. (2019) scheme is O($i \times m + i \times n$s). The approximated complexity of the Liu et al. (2019) scheme is O($i \times m$). Thus, it could be ascertained from the above-mentioned analysis of the complexity of the schemes that the recommended I-frame protection scheme has minimal computational complexity.

4.5 Conclusion of the Chapter

In this chapter, the I-frame based protection scheme has been devised for augmented HEVC where the accumulated error is not spread to the nearby block. This standard reinforces diverse transforms for proper decorrelation of residual error within the I-frame's intra-luma blocks. Many of these transforms are not symmetric, and therefore should not be utilised straightforwardly for watermark embedding. Thus, symmetric transforms have been employed in this scheme to ingrain the watermark within the chosen luma blocks. To incorporate a watermark, the association among both the amplitudesof a pair of intra-luma block'stransform coefficients is modified. One such pair of coefficients is chosen in a manner that the distortion caused by the ingraining of the watermark does not circulate to the surrounding blocks through the standard'sintra-prediction process. The overall value of resilience toward re-compression for the proposed scheme, with various QP values, as well as for the schemes of Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.8934, 0.6351, 0.8639 and 0.4270 respectively. The drop in PSNR for the recommended as well as for the schemes of Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.1278 dB, 1.7206 dB, 0.1899 dB and 1.9130 dB respectively. The upsurge in bit-rates for the newly devised scheme as well as for the schemes of Dutta and Gupta (2016), Gaj et al. (2017) and Liu et al. (2019) are 0.7%, 1.10%, 0.9% and 1.15% respectively. Experimental data support the potency of the recommended scheme with regards to ocular quality, retrenched BIR as well as resistance towards attacks.

Chapter 5

A P-Frame Based Protection Scheme for Enhanced HEVC

The I-frame based scheme proposed in Chapter 4, is not sufficient for some application like a video surveillance system, video conferencing system etc. In these applications, videos are compressed using the LD configuration video coding structure, in which all the frames are P-frames except the first one that is I-frame. This scheme does not have adequate embedding capacity for these application as there is only one I-frame in the video. On the other hand, this scheme does not achieve satisfactory robustness when applied to P-frames. The blocks in the P-frames are compressed using both intra and inter prediction. This results in change of intra mode to inter mode and vice versa after the re-compression, which results in de-synchronization in the watermark retrieval process at the decode side. Thus, the primary aim of this chapter to improve the robustness of the I-frame protection scheme when applied to P-frames. Considering this, a robust P-frame bases protection scheme is proposed for the LD configuration of the enhanced HEVC standard. The robustness for the recommended P-frame protection scheme has been accomplished by employing a threshold that is determined at running time depending on the texture in the P-frames. This threshold is based on the observation that, the blocks selected with more number of non-zero QTCs with magnitude greater than one are robust to synchronization error.

The limitation of existing P-frame based schemes are analysed in Section 5.1. The various transforms used for P-frames in enhanced HEVC are explained in Section 5.2. The explanation of the proposed P-frame protection the scheme has been enunciated in Section 5.3. The attained accuracy for the improved scheme has been measured by executing series of experiments and are analysed in Section 5.4. Lastly, the potency of the recommended P-frame protection scheme is validated appertaining to the increase in bit-rate, robustness, and imperceptibility, and the results are concluded in Section 5.5.

5.1 Introduction

The performance of the existing P-frame schemes is analyzed for enhanced HEVC standard in the following text. The LSB's of QTC's are altered to insert the watermark in the Swati et al. (2014)'s scheme. The scheme of Van et al. (2015) embeds a watermark by changing motion vector difference and the last non-zero transformed residual coefficient of the 4×4, 8×8 and 16×16 inter blocks are manipulated to embed the watermark. Dutta and Gupta (2017) embedded robust watermark in luma blocks by manipulating the magnitudes of DCT_{II} coefficients. Shanableh (2018) scheme embedded watermark by modifying the split decision of blocks of size 32×32 and 16×16. Yang and Li (2018) scheme embedded watermarking by adjusting either horizontal or vertical motion vector component of the smallest block in the CTU. Li et al. (2019) scheme embedded watermark by altering the size of the 8×8 and 16×16 blocks.

Shanableh (2019) improved the scheme of Shanableh (2018), by altering the size of only 16×16 blocks for embedding watermark. The improved scheme has better visual quality with minimum changes in the splits decisions of the blocks. In Yang et al. (2019), watermark embedding is done by altering the size of 64×64 to 8×8 blocks. The final block size used for inserting the watermark is selected based on the QP used by the encoder. Xu (2019) proposed a protection scheme in which watermark embedding is combined with the encryption process. The absolute value of remaining coefficient levels encoded using golomb rice are altered to embed the watermark in QTCs of both intra and inter blocks.

The main disadvantage of the schemes in (Li et al. 2019; Shanableh 2018, 2019; Swati et al. 2014; Van et al. 2015; Yang and Li 2018; Yang et al. 2019) is that, the motion vector components, LSB's of QTC's, the last non-zero DCT coefficients, and the split decision of blocks are altered by the re-compression process of the encoder. Dutta and Gupta (2017) ingrained the watermark via adjusting the DCT_{II} transformed coefficients present in the intra as well as inter blocks of P-frames. However, in the enhanced HEVC standard, EMT uses different variants of DCT and DST, namely DCT_{II}, DCT_{VIII} and DST_{VII}. On the other hand, Xu (2019) scheme does not apply any criteria for the selection of blocks for embedding the watermark. This results in the de-synchronization of watermark after re-compression.

It could be gleaned from the aforementioned scrutiny that current methods are not effective for watermarking the P-frames appertaining to augmentedHEVC. So, the I-frame protection scheme proposed in Chapter 4 is modified and extended for P-frames. The direct extension of this scheme to P-frames does not give optimal results because of the

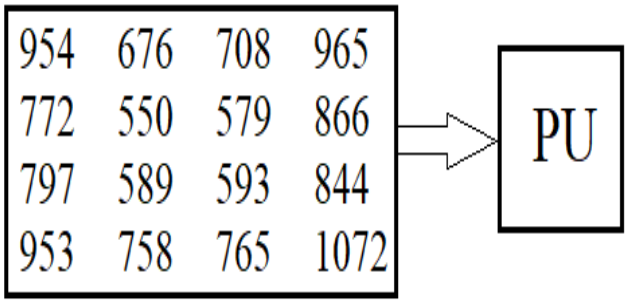


Figure 5.1: Residual error distribution in 4×4 PB in which size of TB is same as PB

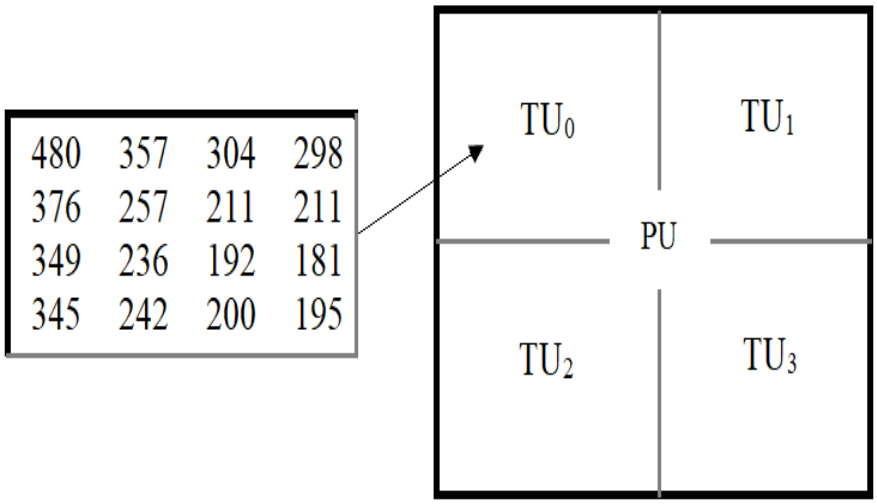


Figure 5.2: Residual error distribution of top left 8×8 where PB is split into four 4×4 TB's

co-existence of intra together with the inter modes. It is difficult to monitor changes in prediction modes and sizes of block in P-frames. The proper selection of 4×4 blocks for the embedding of the watermark in the P-frame is therefore very necessary for the reliability of the scheme in terms of the resistance to re-compression.The 4×4 blocks are chosen on the basis of a threshold that is determined using a non-zero QTC number with a magnitude greater than one present in a block. In addition, the video bit rate can be increased substantially when the watermark is inserted into the P-frames Noorkami and Mersereau (2008), to combat with the increase in bit-rate, only non-zero pair of coefficients are modified to embed watermark.

5.2 Overview of EMT for P-Frames

In HEVC, only the kernel of DCT_{II} is used for residual error coding in P-frames. However, DCT_{II} does not efficiently compact the residual error into a fewer transform coefficient for all the different types of data with edges and curves present in natural videos. So, in enhanced HEVC, the kernels DCT_{II}, DCT_{VIII} and DST_{VII} are used for transform coding of residual error in P-frame (Zhao et al. 2016). The selection of the optimal transform is made based on data present in the PB. There are two types of blocks, which are used to convert residual error into transformed coefficients. In the first category of blocks, the size of PB is the same as TB used for transform coding. Figure 5.1 shows this type of 4×4 PB which is predicted using inter prediction, the residual error is larger near the boundaries of the PB than in the middle of the PB (An et al. 2011). On the other hand, PB of a size larger than 4×4 can further split into four TB's of equal size. Figure 5.2 shows the data in the top-left corner 4×4 block of a 8×8 PB, which is split into four TB's of size 4×4 (An et al. 2011). The residual error in this type of block is lager near the PB boundaries than the TB boundaries. Further, the residual error decreases gradually from top to bottom as well as from left to right. So, EMT is used to handle this type of uneven residual error distribution in a block.

In this scheme, 4×4 intra and inter luma TB's are used for embedding the watermark bits and are addressed as 4×4 blocks in the rest of the text. In EMT, the residual error of 4×4 block generated by the dual prediction process is converted into transform coefficients using two subsequent steps. In the first step, one dimensional horizontal transformation is applied to the residual error. Then the first steps result is transformed by a vertical one-dimensional transform. The residual error of a 4×4 block is denoted by $R_{4\times4}$ and $T_{4\times4}$ are the corresponding transformed coefficients. The horizontal one dimensional kernel is $H_{4\times4}$. Further, $V_{4\times4}$ and $V_{4\times4}^{T}$ are one dimensional vertical transform and transpose of the vertical transform respectively. The kernels of DCT as well as the DST transforms are orthogonal specifically the inverse of $H_{4\times4}$ and $V_{4\times4}$ is equal to the transpose of $H_{4\times4}$ and $V_{4\times4}$ respectively. The $T_{4\times4}$ matrix and residual error matrix can be calculated using (5.1) and (5.2). The kernels in (5.1) and (5.2) are calculated using (4.3), (4.5) and (4.7) by substituting the value of *N=4* and *i, j= 0, 1, 2,3* (Zhao et al. 2016). The kernel of DCT_{II} is used in blocks with flat residual error or residual error with little variations where size of TB is same as PB. Alternatively, DCT_{VIII} and DST_{VII} are used for blocks with more texture regions where size of TB is different than PB as shown in Figure 5.2. The selection from these two transforms has been accomplished via rate-distortion cost

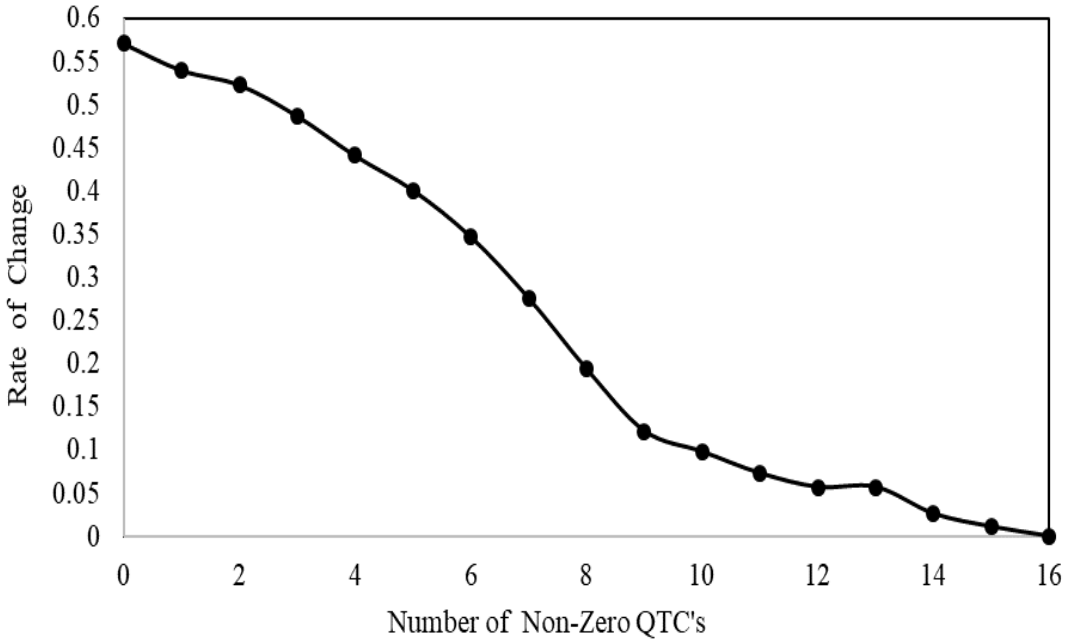


Figure 5.3: Rate of change in P-frame appertaining to deviations in the size of 4×4 to a larger size after re-encoding via Qp equivalent to 18 corresponding to non-zero QTC's

appertaining to every individual TB.

$$\boldsymbol{H}_{4\times4} \times \boldsymbol{R}_{4\times4} \times \boldsymbol{V}^{T}_{4\times4} = \begin{bmatrix} T_{11} & T_{12} & T_{14} & T_{14} \\ T_{21} & T_{22} & T_{23} & T_{24} \\ T_{31} & T_{32} & T_{33} & T_{34} \\ T_{41} & T_{42} & T_{43} & T_{44} \end{bmatrix} \tag{5.1}$$

$$\boldsymbol{H}^{T}_{4\times4} \times \boldsymbol{T}_{4\times4} \times \boldsymbol{V}_{4\times4} = \begin{bmatrix} R_{11} & R_{12} & R_{14} & R_{14} \\ R_{21} & R_{22} & R_{23} & R_{24} \\ R_{31} & R_{32} & R_{33} & R_{34} \\ R_{41} & R_{42} & R_{43} & R_{44} \end{bmatrix} \tag{5.2}$$

5.3 Proposed P-Frame Protection Scheme

The proposed P-frame protection scheme is articulated in this section. The P-frame is divided into variable size TB's from 64×64, 32×32, 16×16, 8×8 or 4×4 on the basis of the texture. The blocks with dimension 4×4 have been employed for ingraining the watermark because the areas with dense textured are partitioned into the smallest blocks, and changes in the dense luminance component are less visible to HVS. The selection of

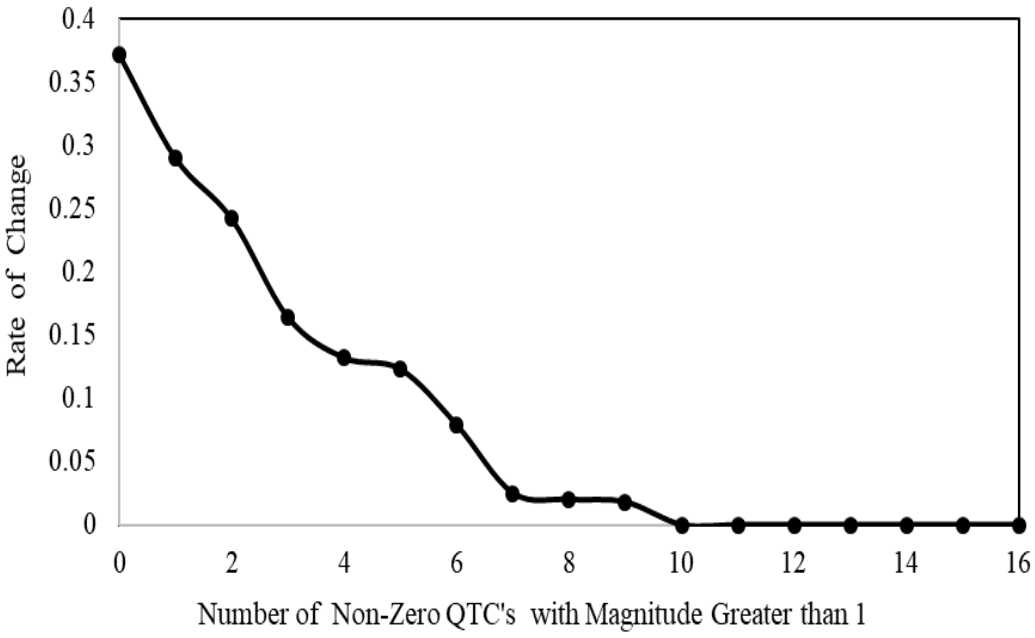


Figure 5.4: Rate of change in P-frame appertaining to deviations in the size of 4×4 to a larger size after re-encoding via Qp equivalent to 18 corresponding to non-zero QTC's with magnitude greater than one

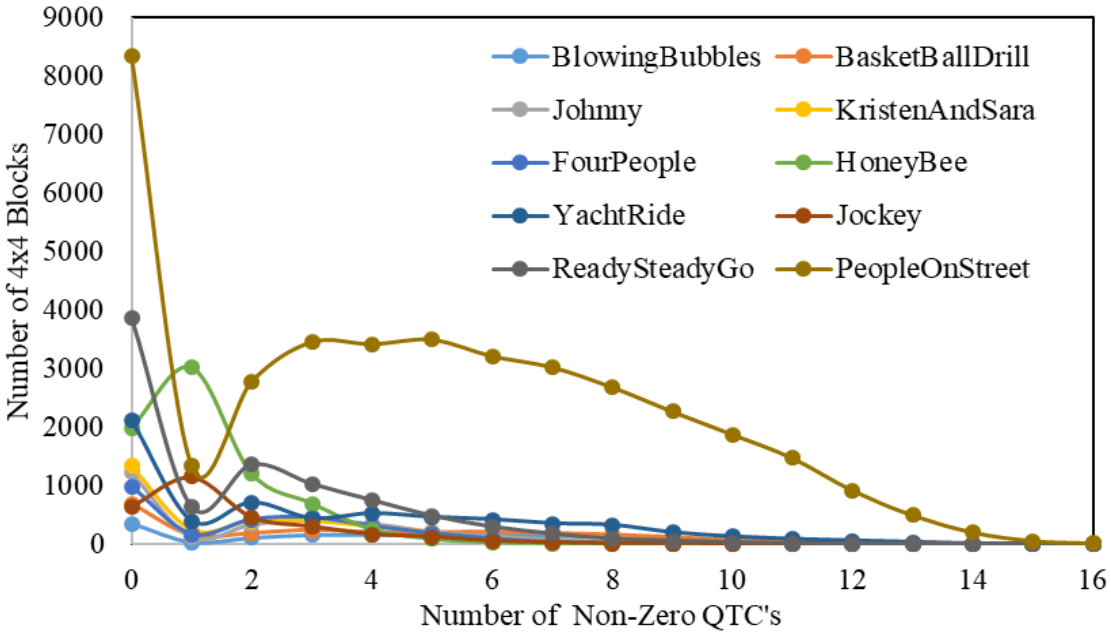


Figure 5.5: Diversities in 4×4 blocks appertaining to non-zero QTC's

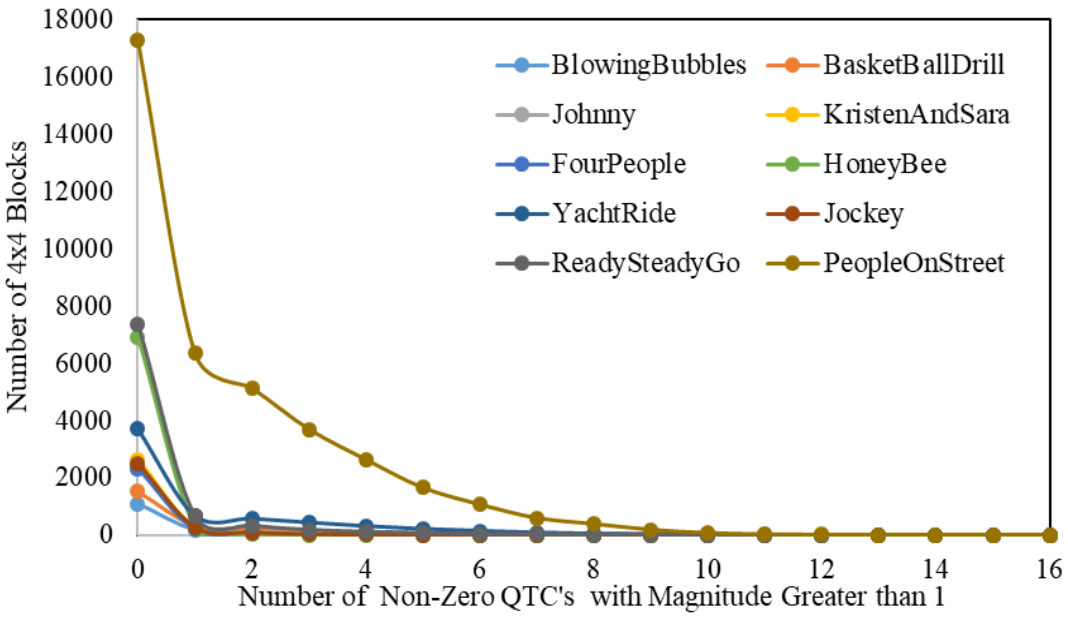


Figure 5.6: Diversities in 4×4 blocks appertaining to non-zero QTC's magnitude greater than one

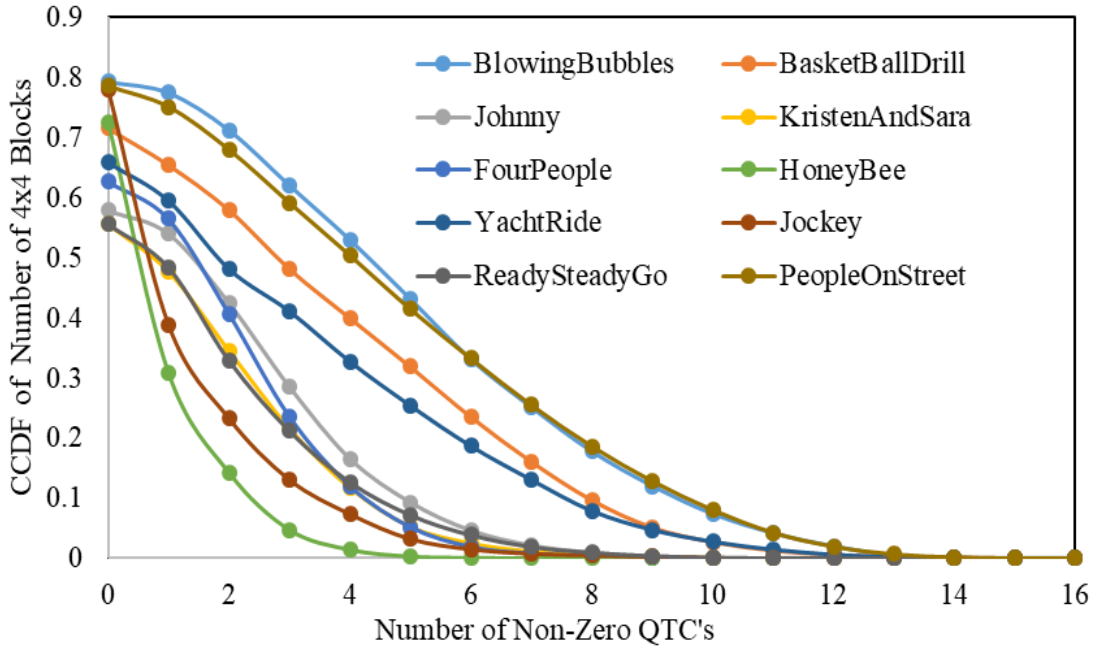


Figure 5.7: CCDF of 4×4 blocks appertaining to non-zero QTC's

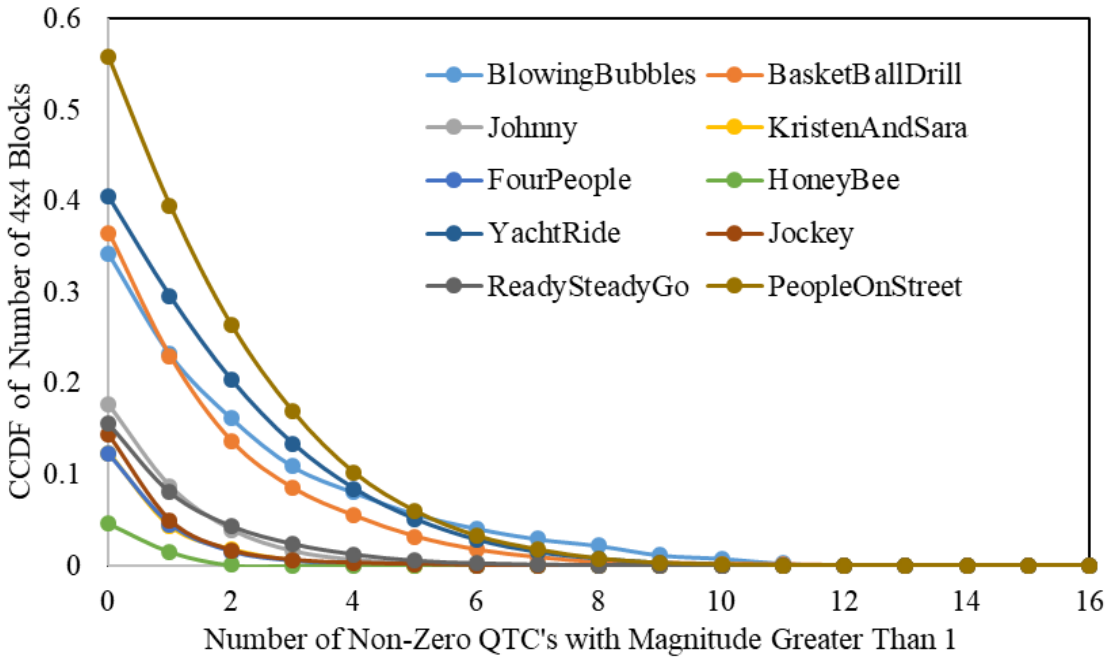


Figure 5.8: CCDF of 4×4 blocks appertaining to non-zero QTC's with magnitude greater than one

appropriate 4×4 block for embedding the watermark is vital for the efficiency of P-frame watermarking schemes in terms of visual imperceptibility and robustness against re-compression attacks. The indistinguishable ocular quality for the recommended P-frame protection scheme has been attained via choosing the 4×4 blocks using threshold T_1, which is reckoned by utilizing the count of non-zero QTC's within a 4×4 blocks. Since, more non-zero QTC's corresponds to the 4×4 blocks with non-flat areas specifically the 4×4 blocks with dense texture.

In P-frames, the prediction is rendered through either utilising 1 of the 34 modes for intra-prediction or motion estimation and compensation process (inter prediction). The dual prediction process in P-frames results in more changes in modes after re-compression. Consequently, it affects the robustness efficiency of the scheme with respect to re-encoding with different QP values. Further, the I-frame protection scheme in Chapter 4 utilizes the threshold calculated using non-zero QTC's for both imperceptibility and robustness. However, it does not give optimal results for P-frames in terms of robustness against re-encoding. There are more flexible partitioning options for inter blocks. Therefore, it results in more changes in the size of inter prediction blocks after re-compression. Figures 4.3 and 5.3 illustrates the rate of changes in I and P frames with respect to the non-zero QTC's present in a 4×4 block. It can be observed from the results that the rate of changes is more in P-frames corresponding to same count of non-zero QTCs value with

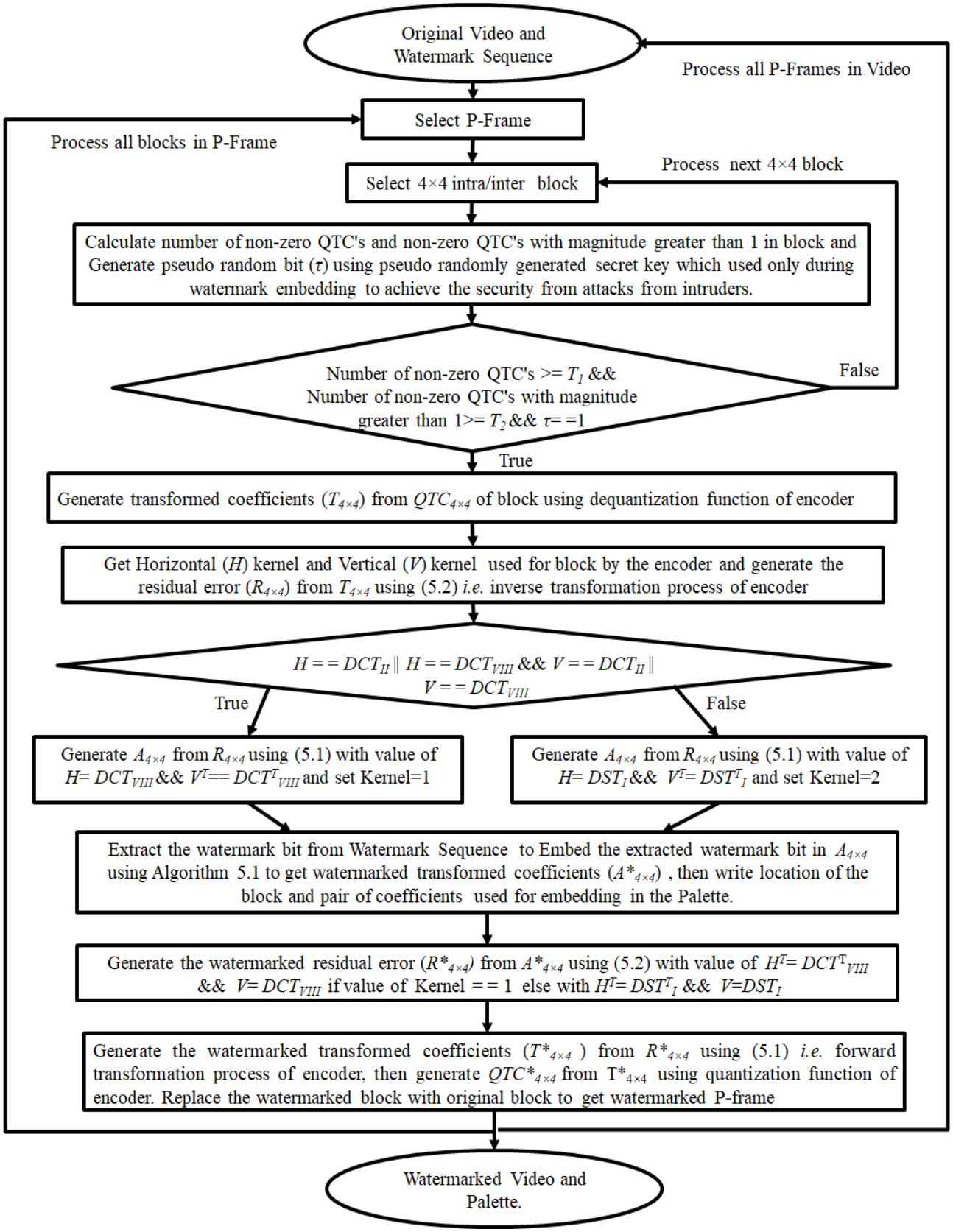


Figure 5.9: Scheme for embedding watermark in P-frames

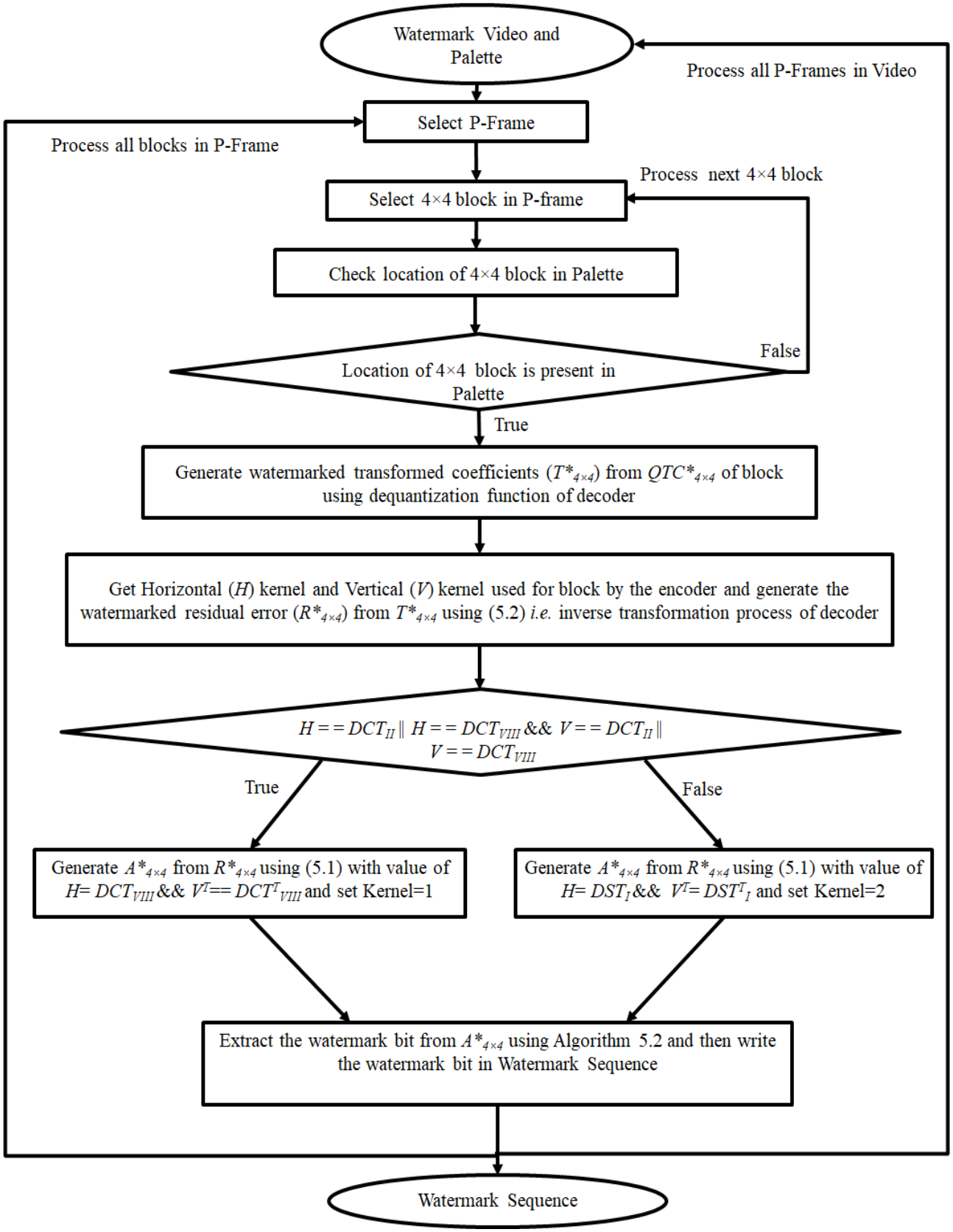


Figure 5.10: Scheme for extraction of watermark from P-frames

Algorithm 5.1: Watermark Embedding in P-Frame

Data: Original Block $A_{4\times4}$ and Watermark Bit WB
Result: Watermarked Block $A^*_{4\times4}$, P: Pair used for embedding watermark

if *WB==1* **then**
- **if** $(|A_{12}| \geq |A_{32}|)\|(|A_{21}| \geq |A_{23}|)$ **then**
  - No change
  - **if** $(|A_{12}| \geq |A_{32}|)$ **then**
    - $P = 1$
  - **else**
    - $P = 2$
- **else**
  - $\Delta_1 = |(|A_{12}| - |A_{32}|)|$
  - $\Delta_2 = |(|A_{21}| - |A_{23}|)|$
  - $P1 = |A_{12}|>0\&|A_{32}|>0?1:0$
  - $P2 = |A_{21}|>0\&|A_{23}|>0?1:0$
  - **if** $(P1 == 1\&P2 == 0)\|P1 == 1\&(\Delta_1 \leq \Delta_2)$ **then**
    - $P = 1$
    - $Sign1 = Sign\ of\ A_{12}$
    - $A^*_{12} = Sign1 \times (|A_{12}| + \Delta_1)$
  - **else**
    - **if** $P2 == 1$ **then**
      - $P = 2$
      - $Sign1 = Sign\ of\ A_{21}$
      - $A^*_{21} = Sign1 \times (|A_{21}| + \Delta_2)$
    - **else**
      - Search for next block

else
- WB==0
- **if** $(|A_{12}| < |A_{32}|)\|(|A_{21}| < |A_{23}|)$ **then**
  - No change
  - **if** $(|A_{12}| < |A_{32}|)$ **then**
    - $P = 1$
  - **else**
    - $P = 2$
- **else**
  - $\Delta_1 = |(|A_{12}| - |A_{32}|)|$
  - $\Delta_2 = |(|A_{21}| - |A_{23}|)|$
  - $P1 = |A_{12}|>0\&|A_{32}|>0?1:0$
  - $P2 = |A_{21}|>0\&|A_{23}|>0?1:0$
  - **if** $(P1 == 1\&P2 == 0)\|P1 == 1\&(\Delta_1 \leq \Delta_2)$ **then**
    - $P = 1$
    - $Sign2 = Sign\ of\ A_{32}$
    - $A^*_{32} = Sign2 \times (|A_{32}| + (\Delta_1 + \Theta))$
  - **else**
    - **if** $P2 == 1$ **then**
      - $P = 2$
      - $Sign2 = Sign\ of\ A_{23}$
      - $A^*_{23} = Sign2 \times (|A_{23}| + (\Delta_2 + \Theta))$
    - **else**
      - Search for next block

comparison to the rate of changes in I-frames. To solve this problem, an additional threshold T_2 based on the number of non-zero QTC's with magnitude greater than one is used to select 4×4 blocks to ingrain the watermark bits. Figure 5.4 demonstrates the rate of change appertaining to to the count of QTC's with magnitude greater than one. The higher the amount of coefficients with a amplitudes greater than one correlates to the less deviations in the size of 4×4 block to the larger size.

$$\overline{F}(T_2) > \rho \tag{5.3}$$

$$\overline{F}(T_2) = P(D > T_2) = 1 - F(T_2) = \sum_{D>T_2} p(D) \tag{5.4}$$

The distribution of the non-zero QTC's varies depending on the texture of a video; therefore, it wil not be pertinent to fix certain invariable values for the thresholds T_1 as well as T_2. Figures 5.5 and 5.6 show the diversities of 4×4 blocks appertaining to different amounts of the count of non-zero QTC's as well as the count of non-zero QTC's with magnitude greater than one. It can be analyzed from these graphs that the distribution of 4×4 varies for different videos. Hence, the thresholds should be selected based on the texture of the P-frames of different videos. The statistical information of each P-frame *i.e.* CCDF is used to calculate and select the thresholds at runtime (Mansouri et al. 2010).

The T_1 will be reckoned on the fly employing (4.8) together with (4.9). To begin with T_1, is the count of non-zero QTC's within 4×4 block on the other hand ν is the percentage of the 4×4 blocks with number of non-zero QTC's equal to T_1 in a P-frame. The $\overline{F}$ *i.e.* CCDF of 4×4 blocks corresponding number of non-zero QTC in P-frame of videos having diverse textures has been emblazoned through the Figure 5.7. The robustness threshold T_2 is calculated at runtime using (5.3) and (5.4). The threshold T_2 is the number of non-zero QTC's with magnitude greater than one in a 4×4 block, and ρ is the percentage of the 4×4 blocks with the number of non-zero QTC with magnitude greater than one equal to T_2 in a P-frame.

The CCDF of 4×4 blocks corresponding to varied count of non-zero QTC's with magnitude greater than one in P-frame of videos with different textures is shown in Figure 5.8. The varied value will be spawned for $T2$ depending on the spatial as well as temporal texture of diverse videos appertaining to the same value of ρ. For illustration, it can be detected from Figure 5.7, ρ equivalent to 0.01, the $T2$ comes out to be 8 with regards to BasketBallDrill contrary to this 10 with regards to PeopleOnStreet. The $T2$ will be elected depending on the necessary strength as well as count of blocks necessary for in-

graining the watermark. The optimal value of $T1$ and $T2$ will be determined on the fly depending upon the count of block required together with the texture of the video under consideration.

The residual error coding of P-frames use EMT explained is Section 5.2 to convert residual error into transform coefficients. The EMT employs different kernels of DCT and DST as these exhibit different properties in terms of compacting the residual error energy into fewer transform coefficients Salomon (2007). In DCT, one DC coefficient is generated, this is the mean of all the coefficients within block. The DC coefficient contains the maximum energy, and the remaining coefficients are known as AC coefficients. The consistency appertaining to the AC coefficients decreases from low to high frequency coefficients *i.e.* coefficients close to the DC coefficients are more consistent. Contrary to DCT, the DST transform generates coefficients near to the boundary where the maximum residual error is concentrated. Dutta and Gupta (2017) manipulated the first two non-zero AC coefficients to insert watermark in a block. So, this scheme does not provide proficient output for enhanced HEVC due to the implementation of EMT for residual error coding. To solve this problem, the watermark embedding is grounded on the characteristics of transforms used by EMT.

Further, some of the transforms used in EMT are non-symmetric, namely DCT_{II}, DCT_V, and DST_{VII} are not symmetric. The alteration in the coefficients of these transforms generate drift error that is spread to the neighbouring blocks can create visible artifacts. The degradation in the visual quality can be fixed by using kernels of the symmetric transforms DCT_{VIII} and DST_I for embedding the watermark. The coefficients in symmetric transforms have the same magnitude in the upper and lower triangle with opposites signs. The problem of the visual artifacts is solved by appropriately selecting the pair of coefficients generated by symmetric transforms for embedding watermark. Additionally, to control the increase in bit-rate only non-zero pairs of coefficients are used for embedding watermark as any changes in zero coefficients increase the bit-rate drastically. The watermark embedding method of the P-frame protection scheme is specified in section 5.3.1 and the extraction method of the P-frame protection scheme is given in Section 5.3.2.

5.3.1 P-Frame Based Watermark Embedding Procedure

The complete mechanism for implantation of the watermark within the P-frames is given in Figure 5.9. The procedure for inserting the watermark is implemented in the video encoding process of Enhanced HEVC compression standard. The inputs to the procedure are original video and watermark sequence (W_m) with the length of m bits that would be

ingrained in the P-frames belonging to the video. To start with, the procedure computes the count of non-zero QTCs as well as the QTCs with magnitude greater than one in a 4×4 block. Additionally, to provide security of the procedure a bit (0 or 1) is generated by a pseudo randomly derived secret key. These three constraints have been employed to choose the block for ingraining the watermark bit, which has been extracted from W_m. Initially, the QTCs of the selected block are dequantized and inverse transformed using the encoder. Thenceforth the watermark bit would be implanted in the block depending upon the EMT mechanism using Algorithm 5.1. The value of Θ is set to zero by default, but it is set to 1 when either Δ_1 or Δ_2 is zero. The 4×4 watermarked coefficients are generated by replacing the original coefficients with a modified pair of coefficients in the original block. In the palette, the position of the corresponding block as well as the pair wielded to ingrain the watermark is penned. Further, 4×4 watermarked residual error is generated from watermarked coefficients. Next, the watermarked 4×4 blocks is generated using the forward transformation and quantization process of the encoder. Finally, we get the watermarked video and palette containing the location of blocks used for embedding the W_m. This palette is given as input to the extraction process to overcome the problem of the de-synchronization during the decoding of the W_m. Hence, the proposed scheme is semi-blind *i.e.* palette is used by the extraction process in place of the original video. The palette can be communicated from one end to another by a secure channel.

Algorithm 5.2: Watermark Extraction from P-Frame

Data: Watermarked Block $\boldsymbol{A}^*_{4\times4}$ and P: Pair used for embedding watermark
Result: WB: Extracted Watermarked Bit
if *P==1* **then**
  if $(|\boldsymbol{A}^*_{12}| \geq |\boldsymbol{A}^*_{32}|)$ **then**
    WB=1
  else
    WB=0
else
  if $(|\boldsymbol{A}^*_{21}| \geq |\boldsymbol{A}^*_{23}|)$ **then**
    WB=1
  else
    WB=0

5.3.2 P-Frame Based Watermark Extraction Procedure

The process of extraction of the W'_m from P-frames of the watermarked video is given in Figure 5.10. The W'_m is extracted using the decoder of the enhanced HEVC compression

standard. The inputs to the procedure are watermarked video and the palette. The location of the current 4×4 block is compared with the location in the palette. If the location of the block is present in the palette, it will be wielded to retrieve the ingrained watermark bit. Then, the watermarked QTC's are dequantized and inverse transformed using the decoder to get the watermarked residual error. Next, the watermark bit is extracted from the watermarked residual error based on the EMT using Algorithm 5.2. Further, the extracted watermark bit is written in the $W_m^{'}$.

5.4 Experimental Results

The enhanced HEVC compression software HM14KTA (2016) is used for carrying out the experiments for the proposed P-frame protection scheme. The dataset wielded in the investigation of the P-frame protection scheme is provided through Table 4.3. The data set contains videos with varying resolutions and textures. The experiments are carried out using LD configuration with the value of QP is set to 16 as well as the size of GOP=4. The average count of bits embedded within a one GOP is equivalent to 25, the value corresponding to the ν for calculating T_1 is set to 0.1, and the value of ρ for calculating T_2 is set to 0.01. The comparative analysis of the output of the experimental outcomes of the recommended P-frame protection scheme is done with the existing Dutta and Gupta (2017) and Xu (2019) schemes for P-frames. The proficiency of the schemes is compared in terms of different parameters like deterioration in ocular quality, changes in the bit-rate of the video, and resilience towards attacks.

The analysis of the robustness for all the three schemes has been done by re-compressing the watermarked videos with different QP values. Further, to examine the robustness against signal processing attacks, the watermarked videos are manipulated with Gaussian noise with mean equivalent to zero as well as the variance equivalent to 0.001 and Gaussian filter of radius 3×3 with the value of sigma equal to 0.3. Table 5.1 gives the results for the proposed P-frame protection scheme ,Dutta and Gupta (2017) scheme and Xu (2019) scheme.The above-mentioned outcomes have been reckoned by applying the correlation coefficient among the originally ingrained watermark as well as the watermark culled subsequently after the attacks. The average revealed in Table 5.1 is the mean value of the correlation coefficients appertaining to a specific attack for all the videos in the data set.

For the re-compression, with a smaller QP value of 14, the average for the proposed scheme is 0.8678, Dutta and Gupta (2017) scheme is 0.7518 and Xu (2019) scheme is 0.4108 . Next, for the re-compression with the same QP value of 16, the average for

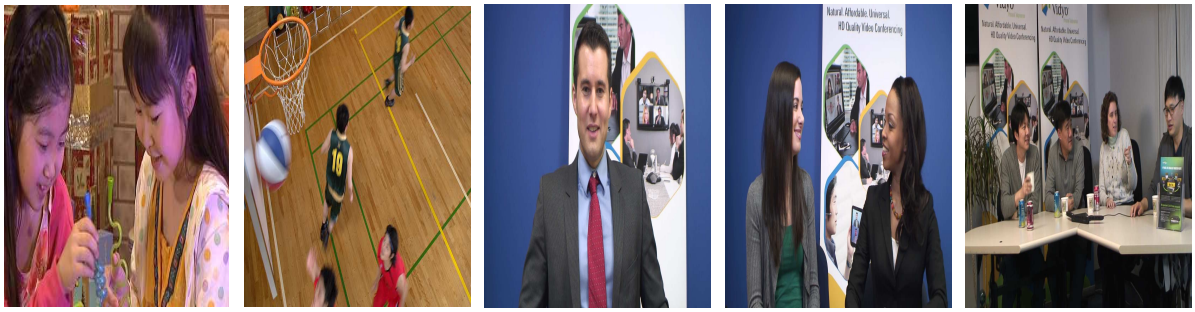
Table 5.1: The correlation coefficient among original and retrieved watermark after attacks for the proposed P-frame protection scheme and existing schemes

Video Sequence	Proposed P-Frame Protection Scheme					Dutta and Gupta (2017)					Xu (2019)				
	Re-compression			Gaussian Filter	Gaussian Noise	Re-compression			Gaussian Filter	Gaussian Noise	Re-compression			Gaussian Filter	Gaussian Noise
	QP=14	QP=16	QP=18			QP=14	QP=16	QP=18			QP=14	QP=16	QP=18		
BlowingBubbles	0.841	0.9412	0.7379	0.7327	0.7517	0.739	0.8571	0.6679	0.6159	0.6733	0.3638	0.5038	0.2913	0.5423	0.5264
BasketBallDrill	0.881	0.958	0.7715	0.7719	0.7836	0.7296	0.8717	0.6738	0.6384	0.6548	0.4865	0.5699	0.3457	0.5283	0.5165
Johnny	0.8547	0.9126	0.7981	0.756	0.7424	0.7091	0.8192	0.6316	0.6499	0.6407	0.4448	0.6301	0.3177	0.5536	0.4756
KristenAndSara	0.8794	0.9263	0.8056	0.7453	0.7566	0.7865	0.8671	0.6626	0.6289	0.6494	0.5414	0.6671	0.4668	0.5974	0.5128
FourPeople	0.8619	0.9583	0.8365	0.7626	0.7407	0.7643	0.8438	0.6746	0.6568	0.6368	0.2981	0.4619	0.2105	0.5047	0.4789
HoneyBee	0.8769	0.9148	0.7837	0.7647	0.7414	0.7189	0.8025	0.6629	0.6485	0.6018	0.3711	0.5543	0.2375	0.5124	0.5014
YachtRide	0.8615	0.9358	0.8396	0.7816	0.7519	0.7957	0.8647	0.6705	0.6761	0.6317	0.4678	0.5298	0.3479	0.5939	0.5473
Jockey	0.8619	0.9325	0.8155	0.7767	0.7472	0.7582	0.8219	0.6635	0.6543	0.6251	0.3685	0.5774	0.2713	0.5919	0.5208
ReadySteadyGo	0.8935	0.952	0.8455	0.739	0.7533	0.7368	0.8347	0.6597	0.6327	0.6407	0.3534	0.5252	0.246	0.5832	0.5265
PeopleOnStreet	0.8657	0.9062	0.8119	0.7418	0.7603	0.7798	0.8439	0.6725	0.6482	0.6164	0.413	0.5594	0.353	0.5935	0.5502
Average	0.8678	0.9338	0.8046	0.7572	0.7529	0.7518	0.8427	0.6640	0.6450	0.6371	0.4108	0.5579	0.3088	0.5601	0.5156

Table 5.2: Comparison of objectives metrics for quality assessment and increase in bit-rate of the proposed P-frame protection scheme and existing scheme

Video Sequence	Proposed P-Frame Protection Scheme				Dutta and Gupta (2017)				Xu (2019)			
	δ_{PSNR}(dB)	$SSIM$	$VIFp$	BIR(%)	δ_{PSNR}(dB)	$SSIM$	$VIFp$	BIR(%)	δ_{PSNR}(dB)	$SSIM$	$VIFp$	BIR(%)
BlowingBubbles	0.067	0.99839	0.91759	0.87	0.841	0.99823	0.90872	1.23	0.155	0.99824	0.91455	0.92
BasketBallDrill	0.041	0.99887	0.93524	1.42	0.902	0.99858	0.92345	1.58	0.096	0.99885	0.93265	1.35
Johnny	0.053	0.99963	0.96357	0.89	0.818	0.99957	0.95152	1.78	0.115	0.99959	0.96058	0.95
KristenAndSara	0.082	0.99992	0.99858	0.96	0.692	0.99977	0.98112	1.14	0.166	0.99979	0.99554	1.07
FourPeople	0.057	0.99973	0.97062	1.02	0.446	0.99861	0.95947	1.79	0.13	0.99969	0.96747	1.14
HoneyBee	0.034	0.99961	0.95163	0.55	0.478	0.99955	0.94889	0.82	0.086	0.99958	0.94904	0.63
YachtRide	0.054	0.99909	0.93559	0.62	0.709	0.99899	0.92317	1.03	0.127	0.99901	0.93246	0.76
Jockey	0.065	0.99974	0.96519	0.37	0.712	0.99889	0.95143	0.96	0.15	0.99962	0.96221	0.42
ReadySteadyGo	0.047	0.99948	0.96498	1.12	0.618	0.99826	0.94836	1.34	0.099	0.99946	0.96267	1.16
PeopleOnStreet	0.097	0.99961	0.93649	0.56	0.552	0.99856	0.91009	0.63	0.177	0.99951	0.93399	0.61
Average	0.06	0.99940	0.95395	0.84	0.677	0.9989	0.94062	1.23	0.13	0.99933	0.95112	0.90

(a) Original (b) Original (c) Original (d) Original (e) Original



(f) Watermarked (g) Watermarked (h) Watermarked (i) Watermarked (j) Watermarked

(k) Original (l) Original (m) Original (n) Original (o) Original

(p) Watermarked (q) Watermarked (r) Watermarked (s) Watermarked (t) Watermarked

Figure 5.11: Original as well as the watermarked P-frames for all the videos in the dataset

the proposed scheme is 0.9338, for the Dutta and Gupta (2017) scheme is 0.8427 and for the Xu (2019) scheme is 0.5579. Finally, a larger QP value of 18, the average for the proposed scheme is 0.8046, for the Dutta and Gupta (2017) scheme is 0.6640 and for the Xu (2019) scheme is 0.3088. The efficacy of the proposed P-frame protection scheme appertaining to the re-compression with distinctive QP values is 0.8687, which is comparatively higher than the existing schemes of Dutta and Gupta (2017) and Xu (2019) *i.e* 0.7528 and 0.4258 respectively. For the signal processing attacks, Gaussian filter and noise, the average for the proposed P-frame protection scheme are 0.7572 and 0.7529, respectively. Comparatively, for the Dutta and Gupta (2017) scheme are 0.6450 and 0.6371, and for the Xu (2019) scheme are 0.5601 and 0.5156 respectively.

Hence, the afore-mentioned results validate that the robustness of the recommended P-frame protection scheme is higher in comparison to the existing schemes. The Dutta and Gupta (2017) embedded watermark in the P-frame by altering the AC coefficient occurring in the initial two positions in the zig-zag order with non-zero intensities generated by the kernel of the DCT_{II} transform. But the enhanced HEVC used EMT, which used different kernels of both DCT and DST for efficiently compacting the residual error spawned through the dual prediction mechanism of the encoder. The recommended P-frame protection scheme has been devised considering the characteristics of the DCT as well as the DST together with a robustness threshold to select the blocks for embedding the watermark. On the other hand, the Xu (2019)does not pursue such a criterion for choosing the block for ingraining the watermark consequently results in de-synchronization of watermark after re-encoding. Therefore, the proposed p-frame protection scheme performs better than the Dutta and Gupta (2017) and Xu (2019) schemes.

The imperceptibility of the scheme is assessed through degradation in the subjective quality and calculating the objective metrics for original and watermarked videos. For evaluating the subjective ocular quality of the watermarked videos, Figure 5.11 reveals the original P-frames, as well as the watermarked P-frames of all the videos in the dataset.From eyeballing on the original as well as the watermarked video frames, it could be scrutinized that there are no spectacular artifacts within the watermarked P-frames. Hence, the proposed scheme has visually imperceptible quality.

The comparison in terms of objective metrics is done in Table 5.2. The average variation in PSNR (δ_{PSNR}) of original as well as the watermarked video for the introduced P-frame protection scheme is 0.060, for the Dutta and Gupta (2017) and Xu (2019) schemes are 0.677 and 0.130 respectively.The average of the $SSIM$ index value for the recommended P-frame protection scheme is 0.99940, for the Dutta and Gupta (2017) and Xu (2019) schemes are 0.99890 and 0.99933 respectively. The average of the $VIFp$ index is 0.95395

for the proposed P-frame protection scheme and for the Dutta and Gupta (2017) and Xu (2019) schemes are 0.94062 and 0.95112 respectively. The average of the percentage increase in bit rates of all videos in data set after embedding the watermark for the proposed scheme, Dutta and Gupta (2017) and Xu (2019) schemes are 0.84%, 1.23% and 0.90% respectively. The BIR of the scheme is comparatively less than the Dutta and Gupta (2017) and Xu (2019) schemes because changes in the transform domain induce less increase in the bit rate than directly altering the QTCs. From the comparative assessment in the above text in terms of subjective and objective metrics, it can be determined that the proposed scheme for watermarking of P-frames is proficient than the Dutta and Gupta (2017) and Xu (2019) schemes.

In the proposed P-frame protection scheme, a watermark has been inserted in the 4×4 blocks. Let us assume the total number of P-frames in the video are p. Further, the number of 4×4 blocks within each P-frame varies from one video to another. Therefore, we can assume the number of 4×4 blocks in a video to be n. The proposed watermark embedding method processes each 4×4 to count the non-zero QTC's along with the QTC's with magnitude greater than one in order to select the block for inserting the watermark bit. The watermark bit is embedded into the selected block in constant time which is not dependent on the number of blocks in a P-frame. Therefore, the worst-case complexity of the proposed P-frame protection scheme in terms of Big-O notation is $O(p \times n)$. The Dutta and Gupta (2017) scheme first computes the average motion vector for all the prediction block in the P-frames and then employs criteria based on the average motion vector and number of non-zero QTC's in a block to select the 4×4 blocks for embedding the watermark. Therefore, the total complexity of the Dutta and Gupta (2017) scheme is $O(p \times m + p \times n)$. As we know, we can approximate the complexity to the maximum value. In the above-calculated complexity, the number of prediction blocks are greater than the 4×4 blocks within a P-frame. Accordingly, the approximated complexity of the Dutta and Gupta (2017) scheme is $O(p \times m)$. Further, the Xu (2019) scheme selects the block for embedding the watermark based on absCoeffLevel of coefficients in a 4×4 blocks of P-frames. This scheme traverses each 4×4 only once. Therefore, the complexity of the Xu (2019) scheme is $O(p \times n)$. The complexity of the proposed P-frame protection scheme is not higher than the existing schemes but it performs better against different attacks.

5.5 Conclusion of the Chapter

In this chapter, a protection scheme is proposed for P-frames. The watermark is embedded in the intra and inter blocks. These blocks are selected using a robustness thresh-

old. This threshold is calculated at runtime depending on the texture available within a frames. The texture exploration has been carried out by employing the count of non-zero QTC's with a value higher than one. The insertion of the watermark is done via selected blocks using symmetric transforms. Firstly, the transform coefficients are generated using a symmetric transform. Then, the relationship between the amplitudes of the non-zero coefficients has been altered for inserting the watermark bits. The robustness of the recommended P-frame protection scheme for re-compression with different QP values is 0.8687, for Dutta and Gupta (2017) and Xu (2019) schemes is 0.7528 and 0.4258 respectively. The average loss in $PSNR$ for the proposed scheme is 0.060 dB, whereas for Dutta and Gupta (2017) as well as for the Xu (2019) schemes are 0.677 dB and 0.130 dB. The overall mean of increase in bit-rate is 0.84%, 1.23% and 0.90% for the proposed, Dutta and Gupta (2017) and Xu (2019) schemes, respectively. Thus, it could be derived through the aforementioned outcomes that the introduced P-frame protection scheme is more effective than the Dutta and Gupta (2017) and Xu (2019) schemes with regards to the robustness, loss in visual quality as well as for an increase in bit-rate.

Chapter 6

Conclusions and Future Work

This chapter is perhaps the final section of the study and also provides several proposals for further enhancement of the existing work. Section 6.1 outlines the key results of the experimental work conducted forth in this research. Suggestions for prospective avenues for study and potential extensions of work discussed in the thesis are given in Section 6.2.

6.1 Conclusions

In this thesis, efficient watermarking schemes are developed for the authentication and protection of HEVC encoded videos. The proposed schemes are summarized as follows:-

- In chapter 3, a semi-fragile scheme have been developed for the authentication of HEVC videos. This scheme divides 4×4I-frame intra-luma blocks into non overlapping subgroups. One subgroup is being used to create authentication code and a second subgroup is used to embed the generated code. The method is robust and visually imperceptible, this is achieved by using threshold value for carefully selecting blocks for the generation and insertion of authentication code. The calculation of these thresholds is carried out at run time using only non-zero QDST in 4×4 blocks. The developed authentication method is investigated for video with divergent textures and resolutions. The average result for the re-compression with varying values of QP is 0.8079. The average decrease in $PSNR$ and BIR are 0.1253 dB and 1.54%, respectively. The findings above suggest that the recommended authentication strategy is effective in contexts to re-compression robustness, minor loss of video quality and slight increase of the bit rate.
- In this Chapter 4, I-frame based protection scheme is designed for enhanced HEVC. The proposed I-frame protection scheme uses the residual error in the intra luma 4×4 blocks to embed the watermark.This standard applies EMT on the residual error obtained throughout the encoder by that of the prediction process. Some of the transforms used in EMT are not symmetric, so they should not be utilized

directly for inserting watermark. The change in transformed coefficients produced by transforms that are not symmetrical induces perceptible disturbances, specifically a drift error which is somewhat transmitted often to nearby blocks. In this scheme, symmetric transforms are being utilized to insert the watermark in blocks. The correlation among the magnitudes of the chosen subgroup of transform coefficients within a block has been used for the incorporation of a watermark. The imperceptibility of the proposed scheme has been obtained via maintaining the sign for both the transform coefficients while inserting the watermark. The outcomes for the proposed scheme for the drop in $PSNR$, bit-rate escalation and robustness towards re-encoding with different QP values are 0.1278 dB, 0.7%, and 0.8934. Simulation results demonstrate the effectivenessof the suggested scheme for the indistinguishable visual deterioration, nominalinflation ofbit-rate and robustness towards re-compression.

- In Chapter 5, a robust protection scheme is proposed for P-frames. First, the 4×4 luma transform blocks will be choosen employing robustness threshold that is calculated utilising the number of non-zero QTC with a magnitude greater than one. Then, the watermark would be embedded through modifying the transform coefficients generated by the selected symmetric transform based on the multiple transforms used by the current block. Furthermore, only non-zero transform coefficients have been employed to incorporate the watermark in attempt to manage its bit-rate increment. The efficiency of this proposed scheme is verified by carrying out a series of experiments on videos with diverse textures and resolutions. The efficiency of the proposed scheme towards the re-compression attack is 0.8687. The average increase in bit-rate and the reduction in $PSNR$ for the suggested method is 0.84% and 0.060 dB. The experimental findings indicate whereby the proposed solution is effective in aspects regarding to therobustness, bit-rate increase as well as visual quality deterioration.

6.2 Scope for Future Work

The research findings of this study can be broadened in the following several directions:

1. In the proposed authentication scheme, AIM is generated to detect any malicious tampering in the frame of the video. The generated AIM is processed to remove non-malicious tampered elements and localize the maliciously tampered areas. The AIM processing does not remove all non-malicious tampered elements. So, the existing procedure can be enhanced for more accurately differentiating between the

non-malicious and malicious tampered elements.

2. Videos require high-end infrastructure for processing and storage. The installation and maintenance of this infrastructure are very costly. Nowadays, a cost-effective alternative approach is using cloud-based computation and storage systems for inserting a watermark in videos. However, in these systems, sensitive data is shared with the servers of the third party. This problem is solved by using secure access control systems available in the cloud infrastructure. In the future work, authentication and protection of HEVC videos can be done by extending the cloud-based information encryption and hiding methods for videos encoded through HEVC.

3. Recently, the extension has been made in HEVC standard to support multi-view 3D videos. For this, new features are added in the standard as the 2D video contains only texture data, but the 3D video contains depth maps as well in addition to texture data. So, the proposed authentication and protection schemes for HEVC encoded videos can be extended for the 3D HEVC standard with modifications.

References

Alattar, A. M., Lin, E. T., and Celik, M. U. (2003). Digital watermarking of low bit-rate advanced simple profile mpeg-4 compressed video. *IEEE transactions on Circuits and Systems for Video Technology*, Vol 13(8):pp. 787–800.

An, J., Zhao, X., Guo, X., and Lei, S. (2011). Non-ce7: Boundary-dependent transform for inter-predicted residue.

Bayram, B., Çavdaroğlu, G. Ç., Şeker, D. Z., and Külür, S. (2017). A novel approach to automatic detection of interest points in multiple facial images. *International Journal of Environment and Geoinformatics*, Vol 4(2):pp. 116–127.

Chang, P.-C., Chung, K.-L., Chen, J.-J., Lin, C.-H., and Lin, T.-J. (2014). A dct/dst-based error propagation-free data hiding algorithm for hevc intra-coded frames. *Journal of Visual Communication and Image Represention*, Vol 25(2):pp. 239–253.

Chen, J., Chen, Y., Karczewicz, M., Li, X., Liu, H., Zhang, L., and Zhao, X. (2015). Coding tools investigation for next generation video coding based on hevc. In *Proceedings of SPIE Optical Engineering+ Applications*, pages 95991B1–95991B9.

Chen, T.-Y., Chen, T.-H., Lin, Y.-T., Chang, Y.-C., and Wang, D.-J. (2008). H. 264 video authentication based on semi-fragile watermarking. In *Proceedings of IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 659–662.

Cox, I., Miller, M., Bloom, J., Fridrich, J., and Kalker, T. (2007). *Digital Watermarking and Steganography*. The Morgan Kaufmann Series in Multimedia Information and Systems. Elsevier, San Francisco, second edition.

Cox, I. J., Kilian, J., Leighton, F. T., and Shamoon, T. (1997). Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, Vol 6(12):pp. 1673–1687.

Cross, D. and Mobasseri, B. G. (2002). Watermarking for self-authentication of compressed video. In *Proceedings of IEEE International Conference on Image Processing*, volume 2, pages II913–II916.

Derf (Accessed:2015). https://media.xiph.org/video/derf/.

Digiturk (Accessed:2016). http://ultravideo.cs.tut.fi/#testsequences.

Dutta, T. and Gupta, H. P. (2016). A robust watermarking framework for high efficiency video coding (hevc)–encoded video with blind extraction process. *Journal of Visual Communication and Image Represention*, Vol 38:pp. 29–44.

Dutta, T. and Gupta, H. P. (2017). An efficient framework for compressed domain water-

marking in p frames of high-efficiency video coding (hevc)–encoded video. *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol 13(1):pp. 1–12.

Dutta, T., Sur, A., and Nandi, S. (2013). A robust compressed domain video watermarking in p-frames with controlled bit rate increase. In *Proceedings of IEEE National Conference on Communications*, pages 1–5.

Elrowayati, A. A., Abdullah, M., Manaf, A. A., and Alfagi, A. S. (2016). Robust hevc video watermarking scheme based on repetition-bch syndrome code. *International Journal of Software Engineering and Its Applications*, Vol 10(1):pp. 263–270.

Farfoura, M. E., Horng, S.-J., Guo, J.-M., and Al-Haj, A. (2016). Low complexity semi-fragile watermarking scheme for h. 264/avc authentication. *Multimedia Tools and Applications*, Vol 75(13):pp. 7465–7493.

Furht, B. and Marques, O. (2003). *Handbook of video databases: design and applications*. CRC Press, Boca Raton, first edition.

Gaj, S., Kanetkar, A., Sur, A., and Bora, P. K. (2017). Drift-compensated robust watermarking algorithm for h. 265/hevc video stream. *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol 13(1):pp. 13–24.

Gaj, S., Sur, A., and Bora, P. K. (2015). A robust watermarking scheme against recompression attack for h. 265/hevc. In *Proceedings of National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*, pages 1–4.

Hartung, F. H. and Girod, B. (1996). Digital watermarking of raw and compressed video. In *Proceeding of Digital Compression Technologies and Systems for Video Communications*, pages 205–213. International Society for Optics and Photonics.

He, D., Sun, Q., and Tian, Q. (2004). A secure and robust object-based video authentication system. *EURASIP Journal on Advances in Signal Processing*, Vol 2004(14):pp. 21852200.

HM14KTA (Accessed:2016). Hm reference software available at https://jvet.hhi.fraunhofer.de/svn/ svn_hmjemsoftware/tags/hm-14.0-kta/.

HM16 (Accessed:2015). Hm reference software available at https://hevc.hhi.fraunhofer.de/trac/hevc/browser/tags/.

Hsu, C.-T. and Wu, J.-L. (1998). Dct-based watermarking for video. *IEEE Transactions on Consumer Electronics*, Vol 44(1):pp. 206–216.

Kim, D. W., Choi, Y. G., Kim, H. S., Yoo, J. S., Choi, H. J., and Seo, Y. H. (2010). The problems in digital watermarking into intra-frames of H.264/AVC. *Image and Vision Computing*, Vol 28(8):pp. 1220–1228.

Konyar, M. Z., Akbulut, O., and Öztürk, S. (2020). Matrix encoding-based high-capacity and high-fidelity reversible data hiding in hevc. *Signal, Image and Video Processing*,

pages pp. 1–9.

Krishnamoorthy, R. and Devi, S. S. (2013). Image retrieval using edge based shape similarity with multiresolution enhanced orthogonal polynomials model. *Digital Signal Processing*, Vol 23(2):pp. 555–568.

Langelaar, G. C. and Lagendijk, R. L. (2001). Optimal differential energy watermarking of dct encoded images and video. *IEEE Transactions on Image Processing*, Vol 10(1):pp. 148–158.

Li, C.-T. (2005). Digital watermarking schemes for multimedia authentication. In *Digital watermarking for digital media*, pages 30–51. IGI Global.

Li, L., Dong, Z., Lu, J., Dai, J., Huang, Q., Chang, C.-C., and Wu, T. (2015). An h. 264/avc hdtv watermarking algorithm robust to camcorder recording. *Journal of Visual Communication and Image Representation*, Vol 26:pp. 1–8.

Li, Z., Meng, L., Jiang, X., and Li, Z. (2019). High capacity hevc video hiding algorithm based on emd coded pu partition modes. *Symmetry*, Vol 11(8):pp. 1015–1–1015–18.

Lin, C.-Y., Wu, M., Bloom, J. A., Cox, I. J., Miller, M. L., and Lui, Y. M. (2001). Rotation, scale, and translation resilient watermarking for images. *IEEE Transactions on image processing*, Vol 10(5):pp. 767–782.

Liu, Y., Liu, S., Zhao, H., and Liu, S. (2019). A new data hiding method for h.265/hevc video streams without intra-frame distortion drift. *Multimedia Tools and Applications*, Vol 78(6):pp. 6459–6486.

Liu, Y., Zhao, H., Liu, S., Feng, C., and Liu, S. (2018). A robust and improved visual quality data hiding method for hevc. *IEEE Access*, Vol 6:pp. 53984–53997.

Luo, T., Zuo, L., Jiang, G., Gao, W., Xu, H., and Jiang, Q. (2018). Security of mvd-based 3d video in 3d-hevc using data hiding and encryption. *Journal of Real-Time Image Processing*, pages 1–13.

Mansouri, A., Aznaveh, A. M., Torkamani-Azar, F., and Kurugollu, F. (2010). A low complexity video watermarking in h. 264 compressed domain. *IEEE Transactions on Information Forensics and Security*, Vol 5(4):pp. 649–657.

Mobasseri, B. G. and Raikar, Y. N. (2007). Authentication of h. 264 streams by direct watermarking of cavlc blocks. In *Proceedings of Security, steganography, and watermarking of multimedia contents IX*, pages 65051W1–65051W5. International Society for Optics and Photonics.

Mobasseri, B. G., Sieffert, M. J., and Simard, R. J. (2000). Content authentication and tamper detection in digital video. In *Proceedings of IEEE International Conference on Image Processing*, volume 1, pages 458–461.

Niu, S., Tu, S., and Huang, Y. (2015). An effective and secure access control system scheme in the cloud. *Chinese Journal of Electronics*, Vol 24(3):pp. 524–528.

Noorkami, M. and Mersereau, R. M. (2007). A framework for robust watermarking of h. 264-encoded video with controllable detection performance. *IEEE Transactions on Information Forensics and Security*, Vol 2(1):pp. 14–23.

Noorkami, M. and Mersereau, R. M. (2008). Digital video watermarking in p-frames with controlled video bit-rate increase. *IEEE Transactions on Information Forensics and Security*, Vol 3(3):pp. 441–455.

Ogawa, K. and Ohtake, G. (2015). Watermarking for hevc/h. 265 stream. In *Proceedings of IEEE International Conference on Consumer Electronics*, pages 102–103.

Peng, Y.-J., Hsieh, Y.-C., Hsueh, C.-W., and Wu, J.-L. (2017). Cloud-based buyer-seller watermarking protocols. In *Proceedings of IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*, pages 1–9.

Podder, P. K., Paul, M., Murshed, M., and Chakraborty, S. (2014). Fast intermode selection for hevc video coding using phase correlation. In *Proceedings of IEEE International Conference on Digital Image Computing: Techniques and Applications*, pages 1–8.

Pröfrock, D., Richter, H., Schlauweg, M., and Müller, E. (2005). H. 264/avc video authentication using skipped macroblocks for an erasable watermark. In *Proceedings of Visual Communications and Image Processing*, pages 59604C1–59604C10. International Society for Optics and Photonics.

Pröfrock, D., Schlauweg, M., and Müller, E. (2006). A new uncompressed-domain video watermarking approach robust to h. 264/avc compression. In *Proceeding of SPPRA*, pages 99–104. Citeseer.

Ren, H., Niu, S., and Wang, X. (2019). Reversible data hiding in encrypted images using pob number system. *IEEE Access*, Vol 7:pp. 149527–149541.

Richardson, I. E. (2014). https://www.vcodex.com/hevc-an-introduction-to-high-efficiency-coding/. *VCODEX*.

Salomon, D. (2007). *Data Compression*. Springer, London, fourth edition.

Saxena, A. and Fernandes, F. C. (2013). Dct/dst-based transform coding for intra prediction in image/video coding. *IEEE Transactions on Image Processing*, Vol 22(10):pp. 3974–3981.

Shanableh, T. (2018). Altering split decisions of coding units for message embedding in hevc. *Multimedia Tools and Applications*, Vol 77(7):pp. 8939–8953.

Shanableh, T. (2019). Data embedding in high efficiency video coding (hevc) videos by modifying the partitioning of coding units. *IET Image Processing*, Vol 13(11):pp. 1909–1913.

Sullivan, G. J., Ohm, J., Han, W.-J., and Wiegand, T. (2012). Overview of the high efficiency video coding (hevc) standard. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol 22(12):pp. 1649–1668.

Swati, S., Hayat, K., and Shahid, Z. (2014). A watermarking scheme for high efficiency video coding (hevc). *PloS one*, Vol 9(8):pp. e105613–e105621.

Sze, V., Budagavi, M., and Sullivan, G. J. (2014). *High efficiency video coding (HEVC)*. Springer.

Tew, Y., Wong, K., Phan, R. C.-W., and Ngan, K. N. (2016). Multi-layer authentication scheme for hevc video based on embedded statistics. *Journal of Visual Communication and Image Representation*, Vol 40:pp. 502–515.

Tew, Y., Wong, K., Phan, R. C.-W., and Ngan, K. N. (2018). Separable authentication in encrypted hevc video. *Multimedia Tools and Applications*, Vol 77(18):pp. 24165–24184.

Upadhyay, A. and Dave, M. (2016). Fuzzy-bpn based improved technique for color image watermarking. In *Proceedings of the International Conference on Informatics and Analytics*, pages 1–10.

Van, L. P., De Praeter, J., Van Wallendael, G., De Cock, J., and Van de Walle, R. (2015). Out-of-the-loop information hiding for hevc video. In *Proceedings of IEEE International Conference on Image Processing*, pages 3610–3614.

VQMT (Accessed:2015). https://mmspg.epfl.ch. *Multimedia Signal Processing Group*.

Wang, J. J., Wang, R.-D., Xu, D.-W., and Li, W. (2015). An information hiding algorithm for HEVC based on angle differences of intra prediction mode. *Journal of Software*, Vol 10(2):pp. 213–221.

Wiegand, T., Sullivan, G., Bjontegaard, G., and Luthra, A. (2003). Overview of the H.264/AVC video coding standard. *IEEE Transactions on Circuits System and Video Technology*, Vol 13(7):pp. 560–576.

Wien, M. (2015). *High efficiency video coding*. Springer.

Xu, D. (2019). Commutative encryption and data hiding in hevc video compression. *IEEE Access*, 7:pp. 66028–66041.

Xu, D., Wang, R., and Wang, J. (2011). A novel watermarking scheme for H.264/AVC video authentication. *Signal Processing: Image Communication*, Vol 26(6):pp. 267–279.

Yang, J. and Li, S. (2018). An efficient information hiding method based on motion vector space encoding for hevc. *Multimedia Tools and Applications*, Vol 77(10):pp. 11979–12001.

Yang, Y., Li, Z., Xie, W., and Zhang, Z. (2019). High capacity and multilevel information hiding algorithm based on pu partition modes for hevc videos. *Multimedia Tools and Applications*, Vol 78(7):pp. 8423–8446.

Zhang, J. and Ho, A. T. (2006). Efficient video authentication for h. 264/avc. In *Proceed-*

ings of IEEE First International Conference on Innovative Computing, Information and Control, pages 46–49.

Zhao, X., Chen, J., Karczewicz, M., Zhang, L., Li, X., and Chien, W.-J. (2016). Enhanced multiple transform for video coding. In *Proceedings of IEEE Data Compression Conference*, pages 73–83.

Zou, D. and Bloom, J. A. (2009). H. 264/avc substitution watermarking: a cavlc example. In *Proceedings of Media Forensics and Security*, pages 72540Z1–72540Z12. International Society for Optics and Photonics.

Zou, D. and Bloom, J. A. (2010). H. 264 stream replacement watermarking with cabac encoding. In *Proceedings of IEEE International Conference on Multimedia and Expo*, pages 117–121.