INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

ENTROPY CODING FOR H.264 VIDEO COMPRESSION

Computer Science

Vijayalaxmi S Patil Research Scholar, PDA College of Engineering, Kalaburagi, India

Dr. Suvarna Nandyal Professor, PDA College of Engineering, Kalaburagi, India

ABSTRACT

H.264/AVC standard of the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group. The H.264/AVC standardization is now achieving significant improvement in rate-distortion efficiency for digital systems worldwide. This paper provides an overview of the technical features of H.264/AVC, describing its baseline profiles, motion estimation for inter, and intra prediction. To improve the human perceptual quality of videos deblocking filter and perceptual quality assessments are described in detail.

INTRODUCTION

ITU-T approves the H.264/AVC standard along with ISO/IEC. The MPEG-2 standard also referred to as H.262 video technology, is introduced before ten years to apply for television systems. H.264 helps in broadcasting and transmitting TV signals and storing HD videos on DVDs. However, the increase in services and the popularity of HD demands higher coding efficiency. Video coding helps maximize coding efficiency while dealing with network diversification. These standards give video shaping ability and broadens the digital environment range. H.264 standard uses either Inter or Intra prediction macroblock to exploits either temporal or spatial information to compress redundant data [1]. The intra scheme achieves higher efficiency compared to previous standards. In video recorder and camera, intra codec is as they cannot afford the inter scheme capability [2].

In video compression, Inter and Intra prediction exploits temporal and spatial redundancy. Inter prediction finds the best prediction using reference frames and gives better prediction for video in motion. The macroblocks of inter scheme are 16×8, 8×16, 16×16, and 8×8. For 8×8, the block is spitted by 8×4, 4×8, and 4×4 size. Intra prediction exploits spatial information via neighboring reconstructed pixels. H.264 consists of 4×4 and 16×16 intra blocks, respectively. 4×4 prediction mode is selected by using nine directional prediction filters. 16×16 prediction mode is selected using four directional modes. Nine prediction mode filters are Vertical, Horizontal, diagonal (Down-left), Mode 4: diagonal (Down-right), Vertical-right, Horizontal-down, Vertical Left, and Horizontal-up.
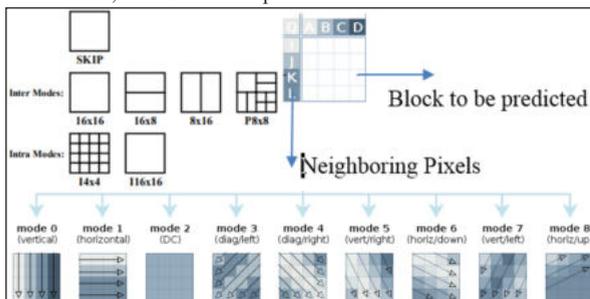

Figure 1: Various modes in H.264

I. LITERATURE SURVEY

Ajay Luthra et al. [3] proposed the H.264 coding standard of ITU-T and MPEG to improve the compression technique and provide conversational and non-conversational video representation applications. Few important H.264 points include motion-detection, the small size of the macroblock, match transform, deblocking filter, entropy coding methods.

Lalit k. Jiwani et al. [4] proposed an intra prediction for video compression using H.264/AVC. Here intra frame is split into 4x4 pixel blocks. Depending upon the phase of the nine possible modes, intra mode is chosen. To select the mode for predicting pixel values, two methods are preferred, such as the Forced method and neighboring 4x4 block orientation and energy. The force method is complicated; hence block energy and orientation methods are used for intra prediction mode.

Lalit K. Jiwani et al. [5] proposed least square prediction mode(LSPM) technique for intra prediction video compression, usually, for selecting intra mode, each frame is split into several 4x4 blocks, and the phase of the block is selected among nine possible cases which require pixel value comparison and Sum of absolute error (SAE). However, SAE is complex; therefore, LAPM is adopted to predict intra frame selection using neighboring pixel values. LAPM method achieves more than 30% reduction of cost while maintaining the quality of the visual screen.

Jiefeng Guo et al. [6] proposed an intra prediction technique for video compression. Long loop, data dependency of 4x4 block force method, and complex algorithms forced to select a novel method called pipelined method to address problems of 4x4 blocks. The pipelined method between MB enhances coding speed. Diagonal and Vertical mode for MB enhances bit rate. Predicting formula redundancy is reduced by incorporating a 4x4 block sharing formula and streamlined method. CAVLC encoder further increases entropy coding speed.

K Sridhar et al. [7] introduced objectives of inter and intraframe predictions. Interframe includes hiding data in motion vectors that are data embedding with matching macro-block phase among two CMVs. The other method is the 4x4 block force method. Thus by inter and intraframe prediction and by using embedding steganography scheme, data is embedded.

Gayathri et al. [8] proposed the H.264/HEVC standard for intra prediction video compression by exploiting spatial redundancy. H.264 allows 35 modes of intra prediction 33 phase modes with DC representing planar mode. This intra prediction method is efficient; however, when a complex structure arrives, then template matching is preferred in the case of H.264/AVC. Since decoder complexity was increased in template matching making it less prior in practical applications. To address decoder complexity, Fast Template matching is adopted with a 33% increase in decoder run time.

Abderrahmane Elyousfi et al. [9] proposed H.264/MPEG intra prediction coding using intra block decision for encoding optimization. This method exploits the spatial redundancy of macroblock based on the similarity of macroblock pixels amplitude. Experimental results reveal slight PSNR degradation and BR increase.

METHODOLOGY

In H.264, Inter or Intra prediction mode is chosen as a compensation scheme to use temporal or spatial redundancy. Intra frame exploits spatial redundancy within a frame as the name intra reveals within or inside. While inter prediction exploits temporal redundancy between successive frames as the name inter reveal among or between. A video frame in H.264 is portioned into several macroblocks, a sequence of the macroblock is called slices.
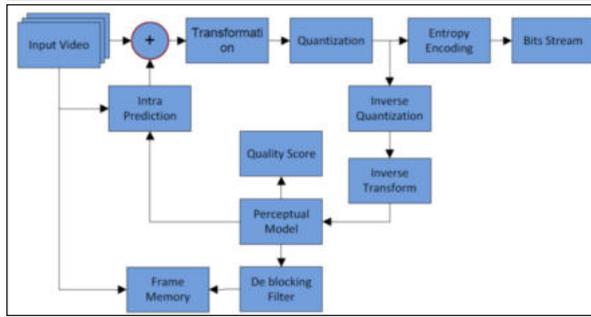
Figure 2: Architecture of the Proposed systemH.264 consists of different slices, namely I, P, B, SP, and SI slices. I slices coded by intra prediction, P slices coded by inter scheme using one motion predicted, B slices coded using inter scheme using two motion predicted. Hence H.264 encode frames using I frame, P frame, and B frames. Video compression using H.264 intra-frame encoding is shown in Figure1, respectively. To encode I-frame, the first step involves mode prediction. Residuals obtained after intra prediction are encoded by entropy encoder after the transform and quantization process. Here the quantized coefficients and redundancy of mode are used by the entropy method to encode quantized coefficients. At the decoder side, the decoding process retrieves residuals using decoder, inverse quantization, and inverse transform, respectively.

Transform
Different transforms used for video coding are DCT, DWT, Haar wavelet, Hierarchical, match inverse, and Short Word Length (SWL) transform, respectively. DCT applies in all video coding standards having an 8x8 block. In H.264/AVC 4x4 block helps the encoder to predict frame adaptively, thereby reducing artifacts. DWT, due to data reduction capability, is used for frequency spatial domain suitable for HVS properties.

The H.264/AVC block size is extended by Hierarchical transform to 8x8 for low-frequency chroma. All standards encoders and decoders use complex processing for transform computation. In H.264, 4x4 DCT is further categorized into two parts scalar multiplication and 4x4 integer transform, combined in quantization. DC coefficients are processed using 4x4 DHT or 2x2 DHT (Hadamard Transform) for 16x16 and 8x8 intra-predictions. In the H.264 quantization step, 52 quantization parameter values correspond to quantization steps are used.

Quantization
The primary reason for data loss is quantization; transformed coefficients are quantized in the matrix to describe the visual perception of humans. Domain reduces effectively when higher frequency results zero after quantization. The quantization matrix parameter decides the quality and compression degree. High compression shows artifacts and more zeros due to blocks compressed without correlation. In videos, the effect of the artifact is more as frame-blocks change individually over time, causing the worst case.

Entropy Coding
Entropy method such as Run Length, Huffman, and arithmetic coding respectively is an effective coding technique to make robust coding method. CAVLC (Context-Adaptive Variable-Length Coding) and CABAC (Context-Adaptive Binary Arithmetic Coding) are two entropy coding in H.264.

CABAC consists of binarized of m-ary data, context modeler, and arithmetic codec. Non-binary data is translated into a sequence of binary symbols by the binarized. The idea of binarization is very similar to variable-length coding (e. g. Huffman coding); however, CABAC uses adaptive arithmetic coding to reduce the existing inter-binary symbol redundancy. Binary symbols are fed to arithmetic codec core that processes the input symbols with the conditional probability of their occurrence in the data stream. The context modeler estimates the conditional probabilities of binary symbols. The way of calculating these probabilities strongly influences the compression ratio of the entropy encoder.

CVLC process AC Coefficients, by considering the transform coefficients, the CVLC table is selected to perform AC co-efficient

encoding. For each trial, the coefficient token bit is modified as 0 and 1 for + and – sign. Encoding of Non-zero Ac co-efficient in CVLC is done in a reverse way. The CVLC table encodes adaptively based on the magnitude co-efficient. The zero co-efficient before the non-zero coefficient is called run before encoded in reverse. A run_before is encoded with two criteria:

(a)　　　　If run_before = no more zeros, then encoding is not necessary.
(b)　　　　Encoding of run_before is not required for the non-zero coefficient. The CVLC is chosen based on zeros, which are not encoded [12].

Deblocking Filter
The input video image or frame is divided into blocks using block based transformation (DCT). Quantization process DCT coefficients using table resulting low bit rate DC and Ac coefficients. The correlation loss, discontinuities, retrieve images to suffer from blocking effects. Blocking artifacts also cause motion prediction of an interpolated pixel from different locations of different frames resulting in edge discontinuities.

Table 1: Deblocking Filter for various standard

Standards	Deblocking Filter
H.261	O (optional)
MPEG-1	NO
MPEG-2	NO
H.263	NO
MPEG-4	Suggested
H.264	Recommended

Reference frame edge discontinuities exist in the interior block is compensated, which causes disturbing artifacts. To counter blocking artifacts and to remove blocking artifacts deblocking filter is adopted. The Deblocking filter for various standards is shown in the below table [13]. The Deblocking works in different ways to remove blocking artifacts.

Deblocking Filter Algorithm Based on the AC coefficients regularization,
step 1. AC and DC coefficients are analyzed.
step 2. Apply ACR on the DC component to smoothen the noise followed by IDCT.
step 3. Apply Block wise shape filtering on Ac coefficients after IDCT to eliminate edge artifacts.
step 4. Apply quantization constraint on the outcome of BWSF to prevent over smoothing.

Combined both the outcomes to get enhanced frames.

$$|p_0 - q_0| < \alpha(Index_A) \quad\quad (01)$$

$$|p_1 - p_0| < \beta(Index_B) \quad\quad (02)$$

$$|p_2 - p_0| < \beta(Index_B) \quad\quad (03)$$

$$|q_2 - q_0| < \beta(Index_B) \quad\quad (04)$$

$$|q_1 - q_0| < \beta(Index_B) \quad\quad (05)$$

$$Index_A = Min(Max(0, QP + Offset_A), 51) \quad (06)$$

$$Index_B = Min(Max(0, QP + Offset_B), 51) \quad (07)$$

β And α are dependent on both Quantization Parameter (QP) and encoder selected offset values.

$$p_0' = p_0 - \Delta_0 \quad\quad (08)$$
$$q_0' = q_0 - \Delta_0 \quad\quad (09)$$
$$\Delta_0 = (4(p_0 - q_0) + (p_1 - q_1) + 4) \gg 3 \quad (10)$$
$$p_1' = p_1 - \Delta_1 \quad\quad (11)$$
$$q_1' = q_1 - \Delta_1 \quad\quad (12)$$
$$\Delta_1 = p_2((p_0 + q_0 + 1) \gg 1) - 2p_1 \gg 1 \quad (13)$$

Perceptual Model
Adaption of human perceptual in video compression for perceptual quality results in better coding efficiency by eliminating perceptual redundancy. Perceptual information elimination includes attenuation of components with high frequency, quantization matrices, chroma subsampling, and filtering. However, the Human Visual System (HVS) is exploited to improve coding efficiency without degrading perceptual quality. Human perception describes the abrupt change, motion, shape, orientation, color, appearance, size, and a human face. SAD or SSD used for distortion is criticized in the survey; the structural similarity (SSIM) index based on RDO forms a quality measurement technique to compromise quality and computation efficiency [14]. JND method

was also adopted for video perceptual quality assessments to reduce perceptual distortion D of retrieved video.

Perceptual quality based on metrics are classified as data Metrics to evaluate fidelity without PSNR, MSE, Packet parameter, Picture, Bitstream, and Hybrid metric, respectively
a.　Perceptual quality with Full Reference metric (FR) measures test video degradation with a reference video.
b.　No Reference (NR) measures test video without reference.
c.　Reduced Reference (RR) metric tradeoff between FR and NR metrics.

NR quality evaluation is absolute in monitoring scenarios while challenging without reference. NR provides the advantage of real-time in-service measurement due to low computational complexity [15].

EXPECTED RESULTS
The original and the predicted frames of respective input data, i.e., Video 1, Video 2, and Video 3 is shown in Figure 3(a) and (b), 4(a) and (b), 5(a) and (b).

Quantization Parameter = 30		
Frame 1		
Frame 2		
Quantization Parameter = 40		
Frame 1		
Frame 2		
	(a)	(b)

Quantization Parameter = 30		
Frame 1		
Frame 2		
Quantization Parameter = 40		
Frame 1		
Frame 2		
	(a)	(b)

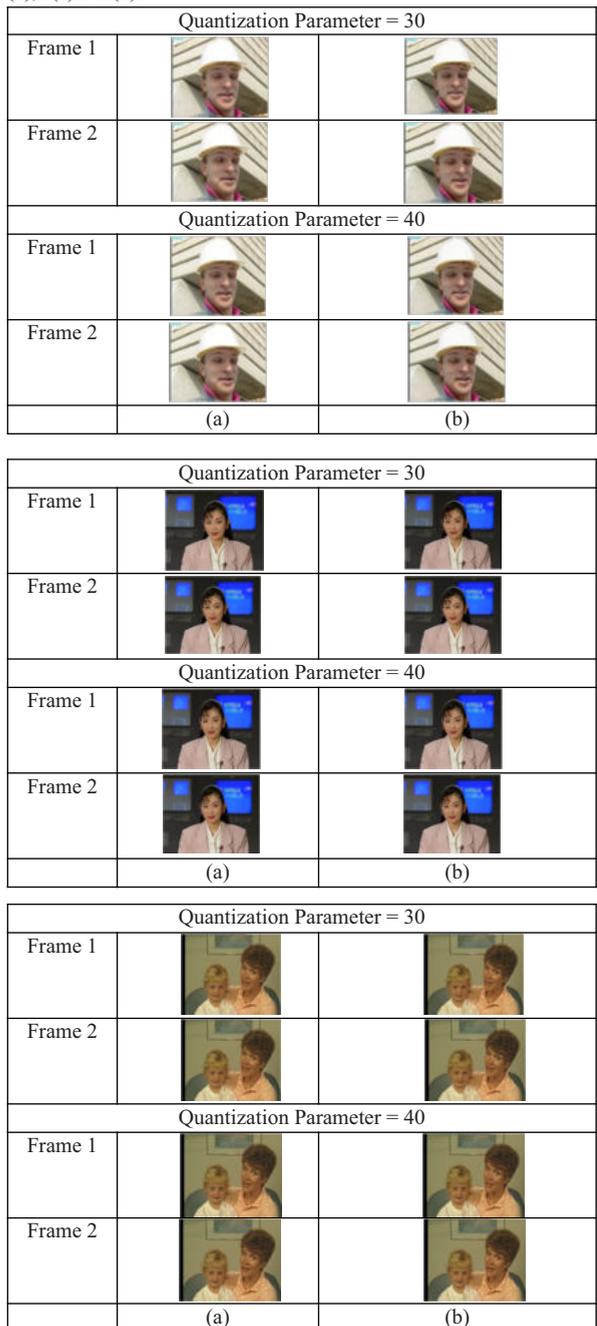| Quantization Parameter = 30		
Frame 1		
Frame 2		
Quantization Parameter = 40		
Frame 1		
Frame 2		
	(a)	(b)

Figure 5: (a) Input Frame; (b) predicted Frame

The main motto of the designed model is to reduce the transmission bit rate of H.264 codec. In Table 3, the bit rate is measured, where the PSNR value is kept constant. Figure 7 shows a graphical representation system performance with constant PSNR values. In table 3, listed system PSNR values by keeping the bit rate constant. The equivalent graph is shown in Figure 8.

Table 2: Performance Analysis of the System for Different QP Value

	QP		PSNR	Bit rate
Foreman	20	Frame 1	47.430	1880.281
		Frame 2	47.424	1777.328
	30	Frame 1	40.869	878.515
		Frame 2	40.867	843.125
	40	Frame 1	31.618	451.093
		Frame 2	31.599	431.890
Akiyo	20	Frame 1	49.729	457.046
		Frame 2	49.742	446.812
	30	Frame 1	41.892	287.046
		Frame 2	41.897	279.421
	40	Frame 1	31.150	205.156
		Frame 2	31.148	199.468
Mother and Daughter	20	Frame 1	49.020	672.359
		Frame 2	49.336	608.140
	30	Frame 1	41.573	320.437
		Frame 2	41.764	290.265
	40	Frame 1	31.293	183.359
		Frame 2	31.278	168.093

Table 3: Average PSNR and Bit Rate of the System (Maintaining Bit rate Constant)

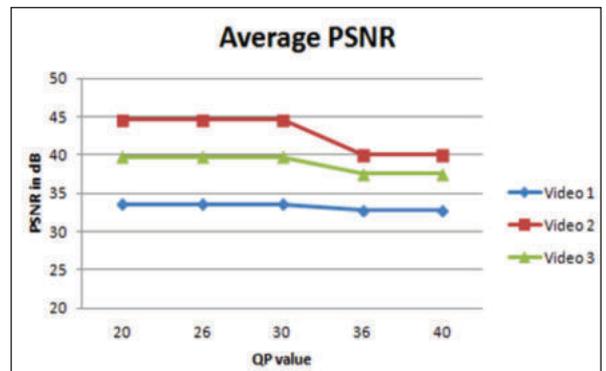| Video	QP	Existing System [25]		Proposed System	
		PSNR (dB)	Bit rate (Kbits)	PSNR (dB)	Bit rate (Kbits)
Video 1 (Foreman)	20	41.52	422.31	42.12	420.21
	26	36.83	180.03	37.23	180.63
	30	33.99	103.02	35.39	101.52
	36	29.96	45.86	31.26	46.26

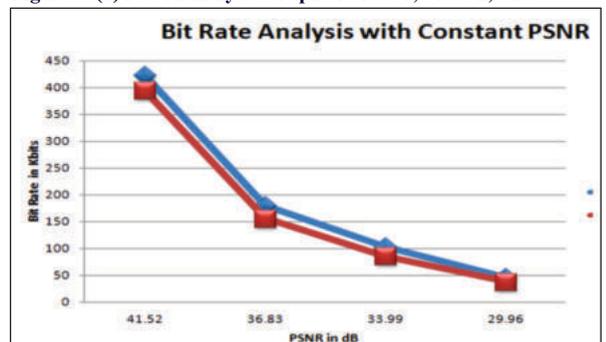**Figure 6: (a) PSNR Analysis Graph of Video 1, Video 2, Video 3**

Figure 7: Average PSNR and Bit Rate Graph of the System (Maintaining PSNR Constant)
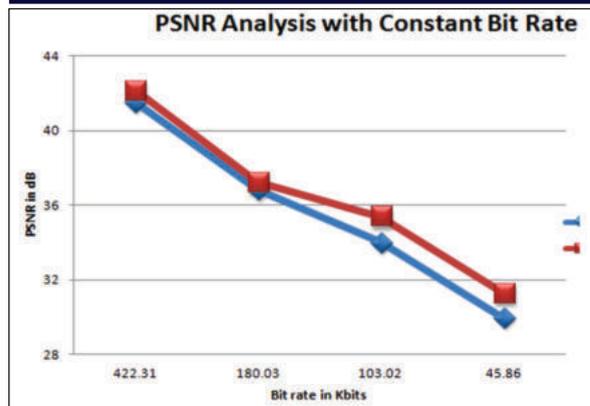
Figure 8: Average PSNR and Bit Rate Graph of the System (Maintaining bit rate Constant)

CONCLUSION

The H.264/AVC standardized by the ITU-T VCEG and ISO/IEC MPEG standard is now representing advances in video technology concerning coding efficiency, enhancement, and flexibility to use virtually over a wide variety of network types. The conventional block-based design for inter and intra prediction summarize few features for future scope such as enhanced prediction capability, use of deblocking filter and perceptual model, use of enhanced entropy coding methods when all these features of the design are combined to provide high-quality compression and better PSNR ratio.

REFERENCES:

[1]. Thomas Wiegand, Gary J. Sullivan, Gisle Bjøntegaard and Ajay Luthra, "Overview of The H. 264/AVC Video Coding Standard", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 13, No. 7, pp. 560-57, 2003.

[2]. Chun-Wei Ku, Chao-Chung Cheng, Guo-Shiuan Yu, Min-Chi Tsai, and Tian-Sheuan Chang, "A High-Definition H. 264/AVC Intra-Frame Codec IP for Digital Video and Still Camera Applications", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 16, No. 8, pp. 917-928, 2006.

[3]. Ajay Luthra and Pankaj N.Topiwala, "Overview of the H.264/AVC Video Coding Standard", Proceedings Volume 5203, Applications of Digital Image Processing XXVI, 2003.

[4]. Lalita. Jiwani, A. Siva Krishna and Mohammed Owais, "Adaptive Energy Weighted Direction Method for Intra-Prediction Video Compression," Region 10 Conference (Tencon), IEEE, 2016.

[5]. Jiwani, Lalit K, "Least Square Intra-Prediction Mode Prediction Based Video Compression," 2018 Conference on Information and Communication Technology (CICT), IEEE, 2018.

[6]. Jiefeng Guo, Zhixin Yang, Jianwei Zheng, and Donghui Guo, "Enhanced Pipelined Architecture of H.264/AVC Intra Prediction, Signal Processing, 2016.

[7]. K Sridhar, Syed Abdul Sattar, and M. Chandra Mohan, "Review on Data Hiding In Motion Vectors and Intra-Prediction Modes for Video Compression," Proceedings of The International Conference on Data Engineering and Communication Technology, Springer, Singapore, 2017.

[8]. Gayathri Venugopal, Philipp Merkle, Detlev Marpe and Thomas Wiegand, "Fast Template Matching For Intra Prediction," 2017 International Conference on Image Processing (ICIP), IEEE, 2017.

[9]. Abderrahmane Elyousfi, Hamza Hamout, and Asma El Hachimi, "An Efficient Intra Block Size Decision for H. 264/AVC Encoding Optimization", 13th International Conference of Computer Systems and Applications (AICCSA), IEEE, 2016.

[10]. Yenewondim Biadgie, Jung-Ju Choi, and Kyung-Ah Sohan, "Improving The Intra-Prediction of H. 264 and H. 265 Video Coding Standards Using Adaptive Weighted Least Squares Based Predictor", International Conference On Mobile and Wireless Technology, Springer, Singapore, 2017.

[11]. Xin Jin, King Ngi Ngan, and Guangxi Zhu, "Combined Inter-Intra Prediction for High Definition Video Coding," In Picture Coding Symposium, 2007.

[12]. Keshaveni, N., S. Ramachandra, and K. S. Gurumurthy, "Implementation of context-adaptive variable length coder for H. 264 video encoder", International Journal of Recent Trends in Engineering, Vol. 2, No. 5, pp. 341, 2003.

[13]. Mirza, Gulistan Raja, and Muhammad Javed and M. Javed, "In-Loop Deblocking Filter for H. 264/Avc Video", Proceedings Of The 5th seas International Conference On Signal Processing, Madrid, Spain, 2006.

[14]. Wang, Shiqi, Abdul Rehman, Zhou Wang, Siwei Ma, and Wen Gao, "Perceptual video coding based on SSIM-inspired divisive normalization," IEEE Transactions on Image Processing, Vol. 22, No. 4, pp. 1418-1429, 2012.

[15]. Karthikeyan, Ramasamy, Gopalakrishnan Sainarayanan, and Subramaniam Nachimuthu Deepa, "Perceptual Video Quality Assessment in H. 264 Video Coding Standard using Objective Modeling", SpringerPlus, Vol. 3, No.1, pp. 174, 2014.

[16]. Luo, Zhengyi, Li Song, Shibao Zheng, and Nam Ling, "H. 264/advanced video control Perceptual Optimization Coding Based on JND-Directed Coefficient Suppression", IEEE Transactions On Circuits and Systems for Video Technology, Vol. 23, No. 6, pp. 935-948, 2013.

[17]. Kenneth Andersson, "Combined Intra Inter Prediction Coding Mode," ITU-T SG16/Q6 (VCEG), Doc. VCEGAD11, Hangzhou, China,