Hybrid Video Encoding for Advanced Video Compression (AVC)

S.Somnathsundar[1] M.E and K.Sathiya[2] M.E
[1]*Department of Electronics and Communication Engineering, Dr.NNCE, India*
somnathsundar87@gmail.com
[2]*Department of Electrical and Electronics Engineering, Dr.NNCE, India*
sathiya.mk@gmail.com

Abstract **— H.264/MPEG-4 AVC is that the latest international video secret writing commonplace. It absolutely was collectively developed by the Video Coding Experts Group (VCEG) of the ITU-T and also the Moving Picture Experts Group (MPEG) of ISO/IEC. It uses state of the art secret writing tools and provides increased secret writing potency for a good vary of applications. The rate-distortion performance of recent video compression schemes is that the results of associate interaction between motion illustration techniques, intra-picture prediction techniques, wave secret writing of variations, and wave secret writing of varied, and rested regions. In this, the compression level and quality is magnified by mistreatment best rework of Karhunen-Loeve Transform (KLT) that is associate optimum decorrelator. Then entropy secret writing, style division matrices (QMs) supported the projected bit allocation theme has been used. Hence Results show that the freshly designed QMs perform higher than the default QMs for H.264/AVC secret writing in terms of each peak signal to noise quantitative relation (PSNR) and structural similarity (SSIM).**

Index terms **-Video secret writing commonplace, increased secret writing potency, Compression level, KLT, division Matrices, PSNR**

I. INTRODUCTION

Digital video communication may be found nowadays in several application situations, such as:

- broadcast, subscription, and pay-per-view services over satellite, cable, and terrestrial transmission channels (e.g., using H.222.0 MPEG-2 systems);
- wire-line and wireless time period informal services(e.g., using H.32x or Session Initiation Protocol (SIP)
- Internet or native space network (LAN) video streaming (using time period Protocol/Internet Protocol (RTP/IP)

H.264 may be a methodology and format for video compression, the method of changing digital video into a format that takes up less capability once it's hold on or transmitted. Video compression or video writing is an important technology for applications like digital TV, DVD-Video, Mobile TV, videoconferencing and net video streaming.

Standardizing video compression makes it doable for product from totally different makers like encoders, decoders and storage media to inter-operate. Associate in nursing encoder converts video into a compressed format Associate in nursingd a decoder convert's compressed video back to an uncompressed format. The fundamental communication drawback could also be display as conveyance supply knowledge with the best fidelity doable inside Associate in Nursing out there bit rate, or it's going to be display as conveyance the supply knowledge mistreatment all-time low bit rate doable whereas maintaining a specified copy fidelity.

However, in sensible video transmission systems the following further problems should be thought of also

- Delay: Delay characteristics square measure influenced by several parameters, as well as process delay, buffering, structural delays of video and channel codec's, and therefore the speed at that knowledge square measure sent.
- Complexity: The quality of the video codec, protocol stacks, and network.

Therefore the Fidelity vary Extensions (FRExt) of the H.264/AVC normal enable the employment of division matrices (QMs) that may be updated at frame level. Though default QMs is specified within the normal, the encoder will specify a custom-built QM for every rework block size and individually for intra and put down prediction, to be used in inverse division scaling by the decoder. Many strategies are projected for the look of QMs for image and video writing; commonest of them being psycho-visual model based mostly techniques and rate-distortion (RD) improvement techniques.

II. VIDEO SOURCE CODING BASICS

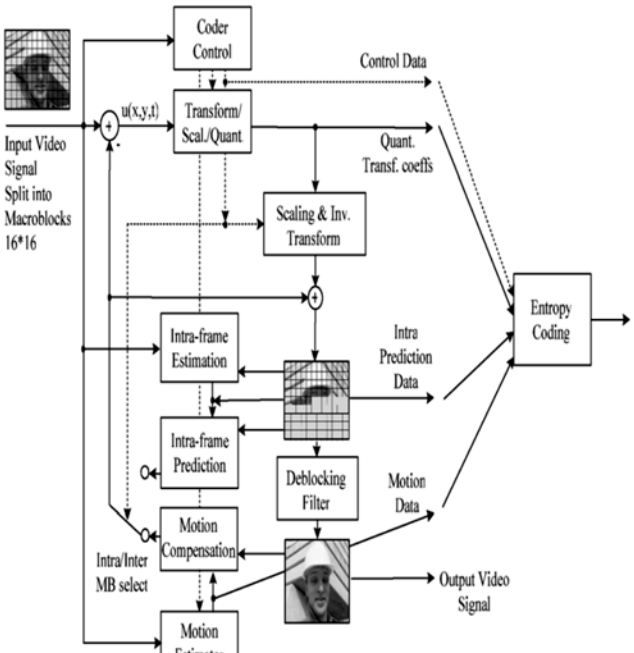
A digital image or a frame of digital video usually consists of 3 rectangular arrays of integer-valued samples, one array for every of the 3 elements of a tri-stimulus color illustration for the spacial space described within the image.

Video secret writing usually uses a color illustration having 3 elements known as Y, Cb, and Cr. part Y is termed luma and represents brightness. The 2 chroma elements Cb and metallic element represent the extent to that the colour deviates from grey toward blue and red, severally.

Techniques for digital compression for several applications will usually be classified as follows.

- **Prediction:** A method by that a collection of prediction values is formed that's accustomed predicts the values of the input samples.
- **Transformation:** A method that's closely associated with prediction, consisting of forming a brand new set of samples from a mixture of input samples, usually employing a linear combination
- **Quantization:** A method by that the exactitude used for the illustration of a sample price (or a gaggle of sample values) is reduced so as to cut back the quantity of knowledge required to code the illustration. The challenge is to attenuate that loss of fidelity in some relevant methodology of activity distortion.
- **Entropy coding:** A method by that discrete-valued supply images are described in an exceedingly manner that takes advantage of the relative chances of the varied potential values of every source symbol.

AA hybrid video coding rule usually returns as follows. Every image is split into blocks. The primary image of a video sequence is often coded in Intra mode. For all remaining photos of a sequence or between random access points, usually inter-picture secret writing modes are used for many blocks. The coding method for Inter prediction (ME) consists of selecting motion knowledge comprising the chosen reference image and MV to be applied for all samples of every block.

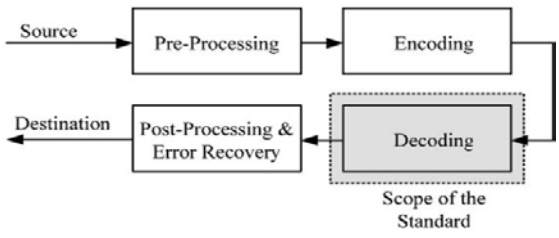


The motion and mode decision data, which are transmitted as side information, square measure employed by the encoder and decoder to get identical Inter prediction signals using MC. The residual of the Intra or Inter prediction that is that the distinction between the first block and its prediction is transformed by a frequency transform. The transform coefficients square measure then scaled, quantized, entropy coded, and transmitted together with the prediction side information.

The encoder duplicates the decoder process so each can generate identical predictions for resulting knowledge. Therefore, the measure transform coefficients square measure created by inverse scaling and square measure then inverse transform to duplicate the decoded prediction residual. The residual is then added to the prediction, and also the results of that addition might then be fed into a deblocking filter to free block-edge discontinuities evoked by the block-wise process. The ultimate image (which is additionally displayed by the decoder) is then stored for the prediction of resulting encoded pictures. The design and operation of an encoder involves the optimization of the many selections to realize the simplest doable trade-off between rate and distortion given the constraints on delay and quality.

III. H.264/AVC VIDEO CODING STANDARD

To address the need of flexibility and customizability to varied applications, the H.264/AVC, style covers a video cryptography layer (VCL), that is intended to with efficiency represent the video content, and a network abstraction layer (NAL), that formats the VCL illustration of the video and provides header data to package that information for network transport



A. H.264/AVC NAL

The NAL is intended to change easy and effective customization of the utilization of the VCL for a broad sort of systems. The total degree of customization of the video

content to suit the requirements of every specific application is outside the scope of the H.264/AVC normal itself; however the look of the NAL anticipates a range of such mappings. Some key building blocks of the NAL style square measure NAL units, parameter sets, and access units.

1) NAL Units: The coded video information is organized into NAL units, every of that is effectively a packet that contains number of bytes.

2) ***Parameter Sets:*** A parameter set contains vital header data that may apply to an oversized variety of VCL NAL units.

3) Access Units: The set of VCL and non-VCL NAL units that's related to one decoded image is spoken as an access unit.

B. H.264/AVC VCL

As all told previous ITU-T and ISO/IEC JTC one video standards since H.261 the VCL style follows the supposed block-based hybrid video cryptography approach.

Macroblocks, Slices, and Slice Groups:

A coded video sequence in H.264/AVC consists of a sequence of coded footage. every image is divided into fixed size macro blocks that every contains an oblong image space of 16 X 16 samples for the luma part and also the corresponding 8 X 8 sample regions for every of the 2 chroma elements. Macro blocks square measure the fundamental building blocks that the secret writing method is specified. Every color part of the residual is divided into blocks; every block is transformed using an integer transform, and also the transform coefficients quantized and entropy coded. The macro blocks of the image square measure organized into slices, which represent regions of a given image that may be decoded severally. Every slice could be a sequence of macro blocks that's processed within the order of a formation scan.

Slice Types

There square measure five basic slice Types.

- **I slice:** Slice coded while not relevance the other frame
- **P slice:** Slice using motion compensated prediction from one reference frame.
- **B slice:** Coded slice predicted using bidirectional motion compensation.
- **SP slice:** Inter-coded slice used for switch between coded bit streams.
- **SI slice:** Intra-coded slice used for switch between coded bit streams (H.264)

Macroblock prediction

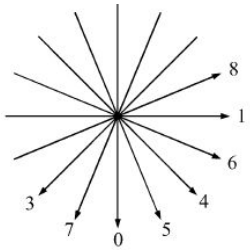
The prediction sources for 3 macroblocks, an I Macroblock, a P Macroblock and a B Macroblock. An I Macroblock is predicted using intra prediction from near samples within the current frame. A P Macroblock is predicted from samples in an exceedingly previously-coded frame which can be before or once this image in show order, i.e. a 'past' or a 'future' frame. Different rectangular sections (partitions) in an exceedingly P MB is also predicted from different reference frames. Every partition in an exceedingly B Macroblock (B MB) is predicted from samples in one or two previously-coded frames, for instance, one 'past' and one 'future'.

Intra Prediction

An intra (I) macroblock is coded while not touching on any information outside this slice. I macroblocks might occur in any slice kind. Each macroblock in an I slice is an I macroblock. I macroblocks square measure coded using intra prediction, i.e. prediction from antecedently coded information within the same slice.

Intra 4*4 prediction using samples.

M	A	B	C	D	E	F	G	H
I	a	b	c	d				
J	e	f	g	h				
K	i	j	k	l				
L	m	n	o	p				

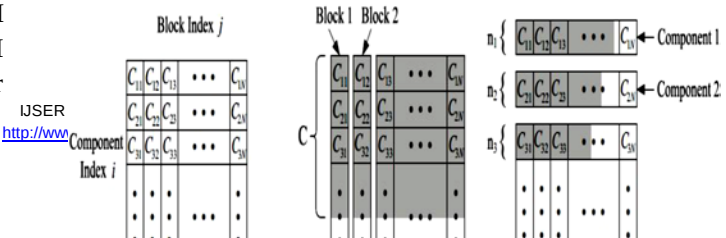


Inter prediction

Inter prediction is that the method of predicting a block of luma and chroma samples from an image that has antecedently been coded and transmitted, a reference image. This involves choosing a prediction region, generating a prediction block and subtracting this from the first block of samples to create a residual that's then coded and transmitted

Transform and Scaling

H.264/AVC uses DCT transform cryptography of the prediction residual. However, in H.264/AVC, the transformation is applied to 4 X 4 blocks (instead of the larger 8 X 8 blocks utilized in previous standards),and rather than providing a theoretical inverse KLT formula to be approximated by every implementer among specified tolerances, a dissociable integer transform with similar properties to a 4 X 4 KLT is employed. The transform cryptography method is analogous to it in previous standards, however since the inverse transform is defined by terribly straightforward actual number operations, inverse-transform mis matches square measure avoided and cryptography complexness is decreased . The smaller transform needs less computation and a smaller process word length.



a. Coefficient selection

Consider a zero-mean stationary continuous-time random process X(t).Using the K-L expansion in the time interval [0,T], the process can be decomposed as

$$X(t) = \sum_{i=1}^{M} \varphi i(t)$$

where $\varphi i(t)$'s are normalized eigen functions and Ci's are uncorrelated random variables with $E[Ci]=0$ and $E[C_i^2]= \lambda i$ .Hence, the random process can be represented by a random vector$\{C_1, C_2,..., C_M\}$and the total average energy of the process is $\sigma^2=\sum_{i=1}^{M} \lambda i$.

Let $x_1(t_1)$, $x_2(t_2)$,..... $x_N(t_N)$ be N independent sample functions of X(t). Consider the product of these N independent sample functions

$$y(t_1, t_2,...., t_N) = x_1(t_1), x_2(t_2),..... x_N(t_N)$$

Yang and Gibson showed that for large N, the number of occurrences of λi in the high energy terms of the energy of $y(t_1, t_2,...., t_N)$ is proportional to λi and given by

$n i = \frac{\lambda i}{\sigma^2} N$, i =1, 2,...,M

and the number of high energy coefficients μ in the product is

$$\mu = \exp[-N \sum_{i=1}^{M} \frac{n_i}{N} \log \frac{n_i}{N}] = e^{NH(S)}$$

where $H(S) = - \sum_{i=1}^{M} \frac{\lambda i}{\sigma^2} \log \frac{\lambda i}{\sigma^2}$ is the spectral entropy in discreteform.

b. Bit Allocation

Consider the decomposition. As before, let the M components {Ci, i = 1, 2, . . . , M} be independent with

E[Ci] = 0 and E[C^2i] = λi. Without loss of generality, we can assume that λ1 ≥ λ2 ≥ . . . ≥ λM.

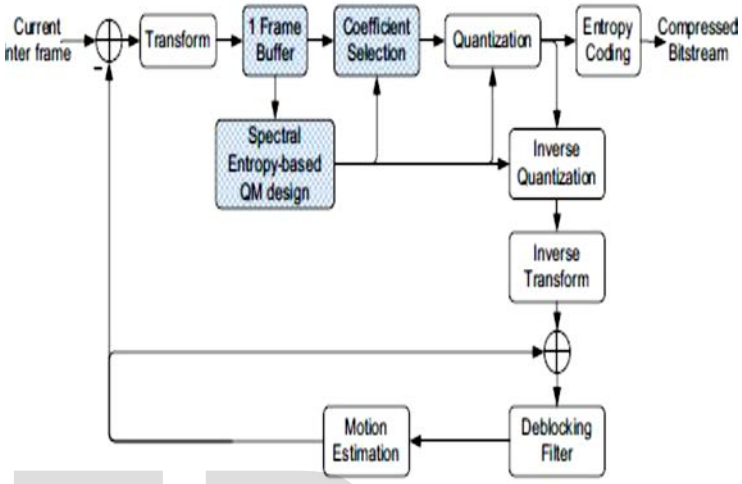
c. Quantization Matrix Design

Since the number of bits $b_i^{(s)}$ allocated for the ith component $C i$ is indicative of the number of steps in the uniform quantizer of $C i$, entropy constrained uniform scalar quantizers optimal for the component distribution can be used

Considering $C i$ to be uniformly distributed within its range, the ith element of the QM can be designed as follows,

$$QM(i)= \text{range}(C i)/(2^{b_i^{(s)}}), \quad i=1, 2,...,M$$

Since the number of bits $b^{(S)}i$ allocated for the ith component C_i is indicative of the number of steps in the uniform quantizer of C_i, entropy constrained uniform scalar quantizers optimal for the component distribution can be used.



Block diagram of the H.264/AVC encoder

Considering Ci to be uniformly distributed within its range, the $_i$th element of the QM can be designed as follows,

$$QM(i) = \text{range}(C_i)/(2^{b(S)i}), \quad i= 1, 2, . . . , M.$$

Alternatively, quantizers can be designed to be optimal for Laplacian distributed random variables which are commonly used to model DCT coefficients of image and video.

Deblocking Filter

Block edges square measure usually foreseen by MCP with less accuracy than interior samples, and block transforms additionally manufacture block edge discontinuities. Block is usually thought-about to be one in all the foremost visible artifacts with the current compression ways. For this reason, H.264/AVC defines an adaptative deblocking filter.

The filter reduces blockiness whereas essentially holding the sharpness of truth edges within the scene. Consequently, the subjective quality is significantly improved. The filter usually reduces bit rate by 5%–10% for identical objective quality because the non-filtered video, and improves subjective quality even additional.

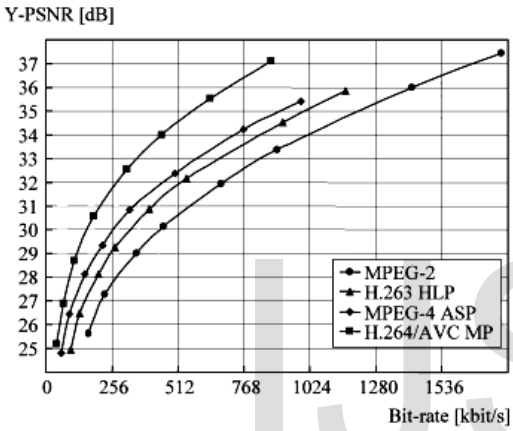
The improved performance of the projected methodology will be attributed to 2 reasons. First, the quantization matrix adapts to the relative energies of the

transform coefficients within the current frame. Second, the spectral entropy-based coefficient selection mechanism exploits latency to look at all the coefficients within the current frame and chooses people who square measure additional vital for holding fidelity.

Experimental Results

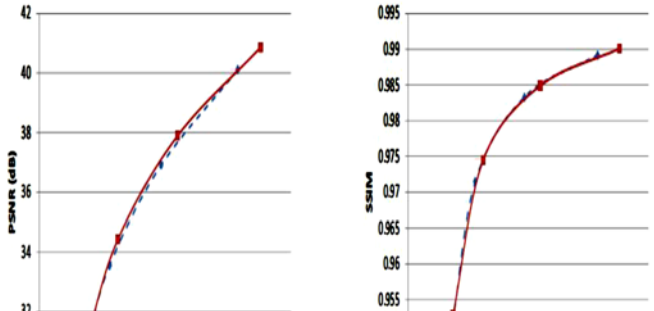
The results of an experiment targeting video streaming application. The live of fidelity is luma peak signal-to-noise ratio (PSNR), that is that the most generally used such objective video quality live.

$$PSNR = 10\log_{10}(\frac{255^2}{MSE})$$



PSNR versus bit rate curves for the sequence "Tempete." For this sequence and for all others within the check set, H.264/AVC considerably outperforms the opposite codecs. Hence the planned bit allocation themes to design inter 4 × 4 QM for the luminosity part within the H.264/AVC encoder. For every inter frame, the residual transform coefficients of all the inter 4 × 4 luminosity blocks square measure 1st buffered i.e. M = 16 and N = $\frac{Width \times height}{16}$.

The energies of the transform parts square measure calculable because the empirical variances of the buffered coefficients. The coefficients selected using the spectral entropy-based coefficient selection is quantity using the designed QMs and at last entropy coded. Here the status of the JM17.0 encoder was used with IPPP. . . group of images (GOP) structure and an intraperiod of 15. The RD curves were obtained by secret writing the check video sequences at four quantization parameters (QP): 20, 25, 30, and 35



The distortion of the luma part versus the typical bits per frame for the encoder using the default QMs (reference method) and freshly designed QMs (proposed method) for the quick content motion with camera pan and high quantity of background detail. Curves square measure provided using each PSNR and SSIM as distortion metrics. It are often seen that the planned QMs perform in addition as or higher than the opposite in terms of each PSNR and SSIM. The PSNR improvement determined is with none loss in sensory activity quality as indicated by SSIM. The improved performance of the planned methodology is often attributed to 2 reasons. First, the division matrix adapts to the relative energies of the transform coefficients within the current frame. Second, the spectral entropy-based coefficient selection mechanism exploits latency to look at all the coefficients within the current frame and chooses people who square measure additional necessary for holding fidelity.

Therefore it adapts to the particular coefficient values that require to be quantized. To attain more sensory activity quality improvement, sensory activity weight of the coefficients is often incorporated into the QM style by employing a weighted distortion metric. To boot, the planned QM style methodology are often used for customizing H.264 QMs for alternative rework sizes and chrominance parts. The Karhunen-Loève rework (KLT), that is an optimum decorrelator, and therefore the discrete cosine transform (DCT), which has performance near that of a KLT once applied to highly correlated auto regressive sources.

IV. CONCLUSION

The new video customary referred to as H.264/AVC presents an upscale assortment of progressive video committal to writing capabilities which will give practical video broadcast or communication with degrees of capability that way surpass those of previous standards.H.264/AVC represents variety of advances in customary video committal to writing technology, in terms of each coding efficiency enhancement and flexibility for effective use over a broad variety of network types and application domains. therefore

summarize a number of the vital variations once compare with the opposite customary once used well along, the options of the new style give or so a 50%bit rate savings for equivalent sensory activity quality relative to the performance of previous standards.

References

[1] M. Narroschke, J. Ostermann, "Adaptive prediction error coding in spatial and frequency domain for H.264/AVC," ITU-T Q.6/GS16,doc. VCEG-AB06, Bangkok, Thailand, January. 2006

[2] Malavika Bhaskaranand and Jerry D. Gibson,Department of Electrical and Computer Engineering "Spectral Entropy-Based Quantization Matrices for H.264/AVC Video Coding"

[3] O. Nemčie, M. Vranješ, S. Rimac-Drlje "Comparison of H.264/AVC and MPEG-4 Part 2 Coded Video" 49th International Symposium ELMAR07, 12-14 September 2007, Zadar, Croatia

[4] Iain Richardson, "An Overview of H.264 Advanced video coding".2007whitepaper. http://www.vcodex.com/files/H.264_overview.pdf (retrieved March 02, 2009).

[5] Ostermann, J., Bormans, J., List, P., Marpe, D., Narroschke, M., Pereira, F., Stockhammer, T., Wedi, T. "Video coding with H.264/AVC: tools, performance, and complexity. IEEE circuits and system magazine , Vol 4,issue 1 , pp. 7-28, 2004

[6] Tian, X.H. Le, T.M. Ho, B.L. Lian, Y., " A CABAC Encoder Design of H.264/AVC with RDO Support", 18th IEEE/IFIP International Workshop on Rapid System Prototyping, 2007. RSP. Pp 167-173, 2007

[7] Detlev Marpe, Heliko Schwarz,Thomas Wiegand."context-based adaptive binary alithmetic coding in the H.264/AVC video compression standard", IEEE transactions on circuits and systems for video techology,vol. 13, no. 7, july 2003

[8] Yuanzhi Zou, Tiejun Huang, Wen Gao, Longshe Huo. Nov, "H.264 video encryption scheme adaptive to DRM". IEEE Transactions on Consumer Electronics, pp. 1289 – 1297, 2006

[9] www.vcodex.com

Author's Profile

S.Somnath sundar was born in Mudukulathur city, Tamilnadu State, India country in 20 May, 1987. Received the B.E., Electronics and Communication Engineering in Syed Ammal Engineering College, Anna University and M.E., Applied Electronics in V.L.B.Janakiammal College of Engineering and Tech, Anna University in 2008 & 2011 respectively. Currently, his job carrier is a Assistant Professor Electronics and Communications Engineering Department, Faculty of Dr.Navalar Nedunchezhiyan College of Enginerring, Tholudur-606303,Tamilnadu,India.

K.Sathiya was born in Virudhachalam city, Tamilnadu State, India country in 29 August, 1988. Received the B.E., Electrical and Electronics Engineering in Kamban Engineering College, Anna University and M.E., Applied Electronics in V.L.B.Janakiammal College of Engineering and Tech, Anna University in 2009 & 2011 respectively. Currently, his job carrier is a Assistant Professor Electrical and Electronics Engineering Department, Faculty of Dr.Navalar Nedunchezhiyan College of Enginerring, Tholudur-606303, Tamilnadu, India.