Elysium Journal Engineering Research & Management

P-ISSN: 2347-4408
E-ISSN: 2347-4734

IMPLEMENTATION OF HIGHLY SECURED H.264/AVC BIT-STREAMS USING CORRELATED MOTION (CM) ESTIMATION FOR HIGH VIDEO QUALITY OF SERVICE (QOS)

G.Divya[1], A.Rajeswari[2]
[1]ME Scholar, Department of Applied Electronics, Dhanalakshmi Srinivasan College of Engineering and Technology, Chennai.
E-mail: dhivya711@gmail.com
[2]Assistant Professor, Dhanalakshmi Srinivasan College of Engineering and Technology, Chennai.

Abstract— **Quality of Service (QoS) is the important criteria in the video processing in H.264/AVC Bit-streams. In this paper, the encryption based bit streams transmitted to improve the security of the bit streams. So the proposed system mainly considered reduction of the de-blocking filter effect in the Mean Square Estimation (MSE). The Correlation Motion (CM) based MSE estimation method derives elaborate distribution models for the transform coefficients in H.264/AVC bit-streams. The MSE between the input and decrypted frames split into two terms. They are quantization error and deblocking effect in H.264/AVC. The quantization parameters and the deblocking effect in MPEG-2 I-frames are estimated with an overall accuracy of 99.9 % and the MSE is estimated with an overall average error of 0.3 dB. The MATLAB simulation results confirms the effectiveness of proposed method over the existing PSNR estimation.**

Index Terms—**Correlation Motion (CM), Deblocking, Discrete Cosine Transform (DCT), Mean Square Estimation (MSE), Motion Estimation, Peak Signal to Noise Ratio (PSNR), Quality of Service (QoS).**

1. INTRODUCTION

Wireless network infrastructure is the key technology for better video communication from one place to another. It uses the greater portion of electromagnetic spectrum for fast transfer. Hence, the internet service providers and mobile network provider updates their strategy in order to provide the best experience to the users based on traffic. With the rapid increase in the demand for high quality video services, Quality of Service (QoS) has become an important issue; thus, intensive study of accurate quality assessment methods on video is essential. The Video Quality Expert Group (VQEG) specifies several guidelines to measure video quality. Traditionally, quality assessment methods are classified into three depending on the availability of original reference data. Full Reference (FR) methods measures the quality by referencing the video data. Reduced-Reference (RR) methods are similar to the full-reference except that the original data are limitedly transmitted or partially used at receiver sides and No-Reference (NR) methods makes Visual Quality Assessment (VQA) without the original data, which are very practical in many applications.

RR methods requires less information than the full-reference methods where the visual quality of delivered video can be estimated by the limited amount of features that reflect the original video. Edge in the image is considered an important feature in the RR method. Human eyes are known to be sensitive to edge features for image perception. In MPEG-7, there is a descriptor for edge distribution in the image. This edge histogram descriptor proposed for MPEG-7 consists only the local edge distribution in the image. That is, since it is important to keep the size of the histogram as small as possible for the efficient storage of the metadata, the normative edge histogram for MPEG-7 is designed to contain only local edge distribution with 80 bins. These 80 histogram bins are the only standardized semantics for the MPEG-7 edge histogram descriptor. However, with the local histogram bins only, it is not sufficient to represent global features of the edge distribution. Note that to improve the retrieval performance, we need global edge distribution as well. In this paper, we generate the semi-global and global edge histograms from the local histogram bins. Then, the global, semi global, and local histogram bins are used to evaluate the similarity between images.

There are two quality evaluation methods in the video quality assessment. They are subjective and objective quality evaluation. Subjective quality evaluation is based on the ratings of pole of human viewers. But, it is not feasible in real time due to complexity. Objective quality assessment methods estimates the video quality without using human observers. Majority of video traffic are based on HTTP / TCP based one due to the reliable delivery and without distortion. The Peak Signal to Noise Ratio (PSNR) mostly used to quantify the encoding the video streams on sender due to low saturation and high PSNR regions.

The NR matrices operates in pixel domain and bit stream domain. In pixel based approaches, the video encoded fully which leads to high complexity. Hence, it is not considered in applications. On the other hand, the bit stream based approach based on the statistical properties of the transformed coefficients. Here the comparison between the quantized coefficients with the original transform coefficients constitutes the distortion. The bit stream based video encode is extended with the use of deblocking filter. The discontinuities in homogeneous regions leads to blocking artifact in highly compressed videos. This blocking effect is efficiently removed with the help of deblocking filter.

In this paper, the efficient Correlation Motion (CM) based MSE on highly secured H.264/AVC Bit-Streams is proposed. The system uses the CM model of the Bit stream network to measure the MSE between the input and decrypted frame. The comparison results confirm the effectiveness of the proposed method over existing approaches using the NR measurement.

P-ISSN: 2347-4408
E-ISSN: 2347-4734

The rest of the paper is organized as follows. Section II presents a description about the previous research which is relevant to the highly secured H.264 / AVC bit streams and the possible solutions. Section III involves the detailed description about the proposed method. Section IV presents the performance analysis. This paper concludes in Section V.

2. RELATED WORK

This section deals with the works related to the highly secured H.264/AVC Bit-Streams, estimation technique implemented to enhance the performance of CR network. *Panayides et al* [1] described the efficient video communication for H.264 / AVC medical video over the mobile network. The performance of the system was validated using OPNET simulations of mobile access control. They also provided the guidelines for the design of diagnostic based encoding of High Efficiency Video Coding standard. The diagnostic based HEVC yielded the high computational complexity. *Schroeder et al* [2] presented the No Reference (NR) based Peak Signal to Noise Ratio (PSNR) estimation based on average bit rate and mean quantization parameter. The low complexity of this approach is most suitable in real time applications. They also investigated the effect of encoding techniques on the PSNR estimation. The prediction of weak textured patches from the noisy image was difficult task. Xinhao et al [3] proposed the novel patch based noise level estimation algorithm to predict the weak textures from noisy image based on the gradients and statistics. The noise level of the image was estimated using Principle Component Analysis (PCA) technique. This model was not suitable where the partial information of original image was available. Reduced Reference (RR) was the alternative tool for the applications with partial information. *Rehman et al* [4] presented the RR Image Quality Assurance (IQA) model which estimated the structural similarity index and the features were extracted using multistate multi orientation divisive normalization transform. *Chaofeng et al* [5] developed the NR IQA with the use of General Regression Neural Network (GRNN). The functional relationship between the features such as entropy of phase congruency image, distorted image and the gradient of the distorted image with subjective mean opinion scores. The experimental results confirmed the closeness to human subjective judgment.

The computational model was needed to predict the human perceived quality of distorted image without using the knowledge of reference. *Peng et al* [6] presented the general purpose NR-IQA based on visual code books. The complex statistics of image captured by using the Gabor filter based feature extraction. The tracking of moving objects was the necessary task in the video bit streams. Sabirin et al [7] presented spatio-temporal graph-based method of detecting and tracking moving objects in video bit streams. The temporal connections between the sub graphs in two consecutive frames were computed to remove the false positive blocks. The best approximation of real shape and position of objects given by varying the size of rectangular region of interests. The power factor was considered in design of high accuracy H.264 / AVC. Zhan et al [8] presented the complexity reduction model which is formed by decomposing the decoder into number of decoding modules. The power level achievement in Intel and ARM chip was discussed. The optimal quality near end users to be enhanced for high accuracy. *Staelens et al* [9] presented novel no-reference bitstream-based objective video quality metric which was constructed by genetic programming-based symbolic regression. The suitable scheme was required for transparent encryption in real time applications. Deng et al [10] presented the Block Based Encryption Scheme (BBES) for transparent encryption. The complexity and compression of this model was very low.

Pourazad et al [11] proposed the design methodologies of High Efficiency Video Coding (HEVC). They also explained the processes such as prediction, frame estimation, encoding and decoding schemes were available in the HEVC. Several fast algorithms developed for reduction of energy consumption in HEVC system. *Sadhvi et al* [12] presented the improved 8 point Discrete Cosine Transform (DCT) on multitude of compression standard for reduction of complexity and the energy consumption in HEVC system. The sharing architecture for complexity reduction was needed in HV.264 bit streams. Yu-Lin et al [13] presented the hardware sharing architecture for efficient multi-standard deblocking strategy. The existing deblocking strategy was reorganized according to H.264 / VP 8. Then the multi sharing architecture was developed for reduction of complexity and energy consumption. The integration of several CPU cores and GPU accelerators called heterogeneous platforms. The parallel operation of CPU+GPU for high performance analyzed. *De Souz et al* [14] aimed the exploitation of multi parallelism level in HEVC deblocking filter by following processes: The implementation of highly optimized CPU parallel operation, the implementation of GPU for HEVC deblocking filter and implementation of hybrid, load balanced CPU+GPU. The code structure unit to be reduced in HEVC based deblocking filter. *Liquan et al* [15] presented the Coding Unit (CU) size utilization algorithm for HEVC Model (HM) for complexity reduction. The CU depth range was calculated and rarely used depth features were skipped.

The video content acquired and distributed in digital format. *Bestagini et al* [16] presented the video codec and analysis of video codec done on the basis of foot prints. The algorithm based approach in video codec used for lossy compression. The experimental setup was constructed to validate the results. But, the suitable model was needed to estimate the perceived video in codecs. *Joskowicz et al* [17] proposed the better estimation model for perceived video. The model was operated on low bit ranges. The error estimation also performed against the International Telecommunication Units (ITU) standards. The parametric model has four parameters for codec. Two of them were related to video movement, the third parameter related to display format and the fourth parameter related to codec performance. The quality estimation of proposed model computed with standard model. The research works concentrated on the video quality estimation for complexity and energy consumption. Chikkerur et al [18] presented the classification based model for FR and RR model where the classification based on natural visual characteristics and sub classification based on natural features. The media layer objective model was analyzed based on classification scheme. The

P-ISSN: 2347-4408
E-ISSN: 2347-4734

classification scheme based on the feature extraction. *Shengbin et al* [19] presented the NR ScalableVideo Coding (SVC) bitstream extraction problem and the solution to an approximately optimal extraction. The packet's priority assignment was done using the greedy algorithm. The Quantization-Distortion model required for the video quality estimation. *Wang et al* [20] presented the QD model for quality estimation of HV.264 / SVC. The CM model based MSE estimation was proposed in this paper.

3. PROPOSED METHOD

The No Reference (NR) based model of HV.264 / AVC video streams is proposed in this paper. The input video is passed through the several processes such as transformation, quantization and deblocking filtering for efficient encryption and decryption. The block diagram of proposed HV.264 / AVC is shown in Fig. 1.

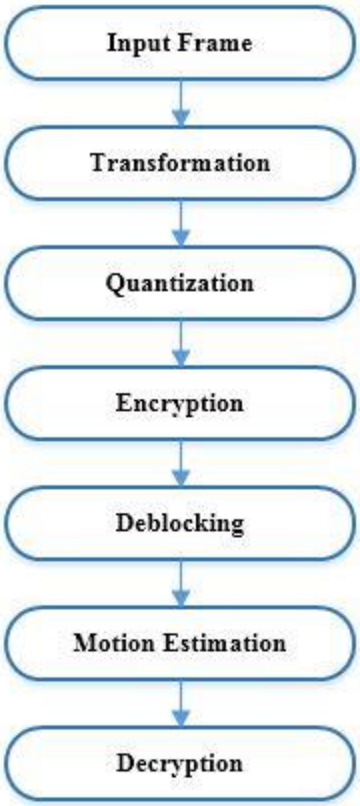


Fig. 1 Block diagram of proposed method

The proposed method implemented using various processes. The input video for analysis formed and the video frames are transformed using Discrete Cosine Transformation (DCT) technique and encrypted. Then overall performance also evaluated over existing method. The flow for proposed method consist of successive steps. They are encryption of video, quantization, inverse DCT, decryption and Mean Square Error (MSE) estimation.

The input video is arranged in number of frames initially. Then the spatial domain of input frames are converted to frequency domain using DCT. The smooth regions are identified in adaptive quantization. The missing pixels in the quantization methods are recovered by deblocking filter. The good estimator of random variables of DCT coefficients is to minimize the mean square error of the frame levels. The correlation between the video bits and neighboring bits are calculated in the motion estimation process. The shifted position of row and columns in encryption process are move back to original position in decryption process.

The flow of HV.264 / AVC bitstream encryption with deblocking is shown in Fig. 2.

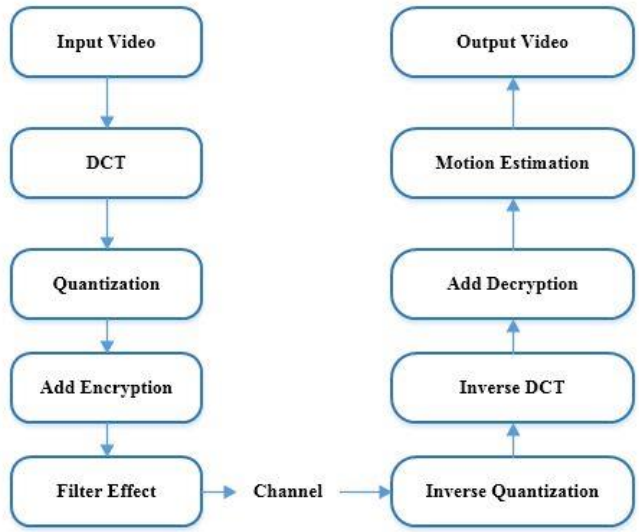
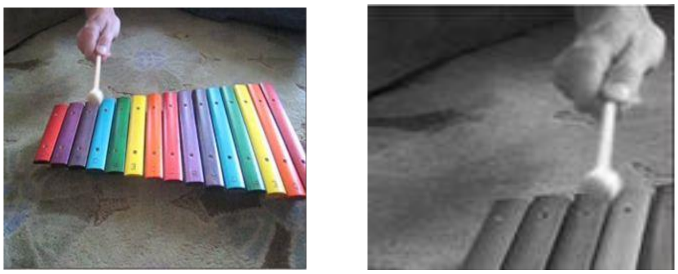


Fig. 2 Flow of proposed method

3.1 INPUT FRAME

The initial process in the video encryption is frame conversion. The input video is converted to I and P frames in this stage. In general, the video stream refers to display of moving pictures in the monitor by progression of images with changing positions. A single frame of video is split into thin lines called scan lines. The conversion of video into frames using two methods called interlacing and progressive method. Interlacing method split the video into two fields odd and even. The odd and even fields transmitted in alternate fashion and displayed in the device. On the other hand, progressive method displays the entire scan lines i.e first to last in a single frames. Hence, large bandwidth is required for P frames. After conversion of video into frames the frames are transferred to encryption process.

P-ISSN: 2347-4408
E-ISSN: 2347-4734



(a) (b)

Fig. 3 Frame Conversion

Fig. 3(a). Fig. 3(b) describes the input original video to the frame conversion process and the output resize frame after the conversion process.

3.2 TRANSFORMATION

The I frames of video be the intra coded with no reference to other frames. P frames of video are predictively coded using neigh boring frames I or P. the video coding is split into three processes such as transform coding, macro block and interpolation. The frames of video represents two dimensional array of RGB triplets. This triplets are transformed to YCrCb triplets using transform coding. The Y value represents luminance and Cr /Cb value represents the chrominance levels. Then the frames converted to macro blocks which corresponds to 16 x16 pixel area of original image. Finally, the interpolation process produces the forward prediction vector for P frame and backward prediction vector for I frame.

The frames are converted to frequency components by the suitable transformation technique called Discrete Cosine transformation (DCT). Let us consider the list of numbers be x of length n. The transformed coefficients of s(x) is given by

$$s(u) = \sqrt{\frac{2}{n}}\, c(u) \sum_{x=0}^{n-1} s(x) \cos\left(\frac{(2x+1)u\pi}{2n}\right), u = 0 \dots . n \qquad (2)$$

Where $c(u) = \begin{Bmatrix} 2^{-0.5} & u = 0 \\ 1 & otherwise \end{Bmatrix}$

The element of list *s(u)* is the dot product of input and base vector. The one dimensional transform is used to process the 1 D signals such as speech signals. There is a need of 2 D transform to process the images. Hence the 2D transform of input values given by following equation.

$$s(u,v) = \sqrt{\frac{2}{n}}\, c(u)c(v) \sum_{x=0}^{n-1} s(x,y) \cos\left(\frac{(2x+1)u\pi}{2n}\right) \cos\left(\frac{(2y+1)v\pi}{2m}\right),$$
$$u = 0 \dots . n \; v = 0 \dots . m \qquad (3)$$

Where $c(u) = \begin{Bmatrix} 2^{-0.5} & u = 0 \\ 1 & otherwise \end{Bmatrix}$

$c(v) = \begin{Bmatrix} 2^{-0.5} & v = 0 \\ 1 & otherwise \end{Bmatrix}$

The application of DCT to the image converts the image into block of transformed coefficients.

3.3 QUANTIZATION

Quantization is the process of reducing the possible values of quantity thereby reducing the number of bits represent it. The functions for quantize the image are developed and the level of compression by quantization also calculated in this process. In video encoding process, the DCT coefficients are quantized using the weight which represents the frequencies to respective coefficient. The input videos are converted to frames, then they are transformed and they are quantized to bit streams. The higher spatial frequencies are less visible to human eye. Hence, the large quantization factors are chosen for high frequencies.

The overall energy resides in the particular section of array of transformed digital image. Adaptive quantization methods are suitable for prior knowledge of smooth regions by using minimum quantization levels in these areas. Hence the smaller number of bits required for representation of quantization levels.

3.4 ENCRYPTION

The sign bits for the pixel positions are encrypted for secure video transmission. The Advanced Encryption Standard (AES) algorithm is used for encryption and decryption of video. AES is the iterative block cipher in which one dimensional array of plain text converted to the state matrix. Transformation round key is generated from cipher key for each round. The objective of video encryption is to prevent the unauthorized receivers in the process. The general scheme for encryption is to apply the transformation $E_{k1}$to video stream S called plain text. Then the *C* is called cipher text is given by following equation:

$$C = E_{k1}(S) \qquad (1)$$

Where *k1* is the encryption key. The proposed AES algorithm is a selective encryption algorithm which operates on the coefficients of Discrete Cosine Transform (DCT). Hence, the encrypted cipher text transformed with DCT in the next process.

3.5 DEBLOCKING

The artifacts are originated in video coding systems by the prediction and decoding strategies called blocking artifacts. The deblocking filter is required to remove the blocking artifacts and improve the video quality. The global filtering strength is modulated to the specific characteristics of video sequence. The amount of filtering is increased or decreased by adjusting the value of pair of quantization dependent parameters. Based on the key values, the filtering parameters are modified. The coefficient values are EX-ORed with 128 and complemented for key value is 1. The

P-ISSN: 2347-4408
E-ISSN: 2347-4734

coefficients EX-ORed with 64 and then complemented for the key value is 0.

Block-based DCT coefficients followed the Laplacian distributions for which the optimal distribution parameters are estimated using quantized coefficients. Due to the DCT and quantization the some pixel are missed due to that transformation. The de-blocking filter is used to remove the missed pixel and it is recovered by de blocking filters.

3.6 MOTION ESTIMATION

The macro blocks (16 x 16) are decomposed into 4x4 blocks in the motion estimation process. The motion vector is calculated for each block. The motion vector prediction and the motion vector differences are calculated to minimize the Mean Square Error (MSE) of the original and decrypted video. During the deblocking, some pixels are altered to new position. The pixel intensity correlation with the original intensity is calculated. Finally, the motion of video bit is found by measuring the correlation between the neighboring bits.

A good estimator is required to minimize the MSE between the pixel intensity correlation calculations. The frame level of overall MSE is given by following equation:

$$MSE^{(n)} = \frac{1}{N_R N_C} \sum_{i=0}^{N_R-1} \sum_{j=0}^{N_C-1} \left\{ f^{(n)}(i,j) - \hat{f}_{DF}^{(n)}(i,j) \right\}^2 \qquad (4)$$

Where N_R, N_C are the number of frames of reference and Correlation between the estimators.

3.7 DECRYPTION

The reverse process of encryption called decryption. In this process, the authorized user with key *k2* decrypt the video by the transformation.

$$D_{k2} = E_{k1}^{-1} \qquad (5)$$

The decryption process is given by the following equation:

$$D_{k2}(C) = E_{k1}^{-1}(C) = E_{k1}^{-1}\big(E_{k1}(S)\big) = S \qquad (6)$$

The sign bit is changed based on the key and the changes are applied to the Inverse DCT coefficients in decoding process. The shifted positions in the row and column of image moved back to the original position if the key is matched.

Fig. 4 shows the final decrypted HV.264 / AVC bit streams into video. The CM based model for HV.264 / AVC is proposed and the performance measures confirms the effectiveness in encryption of video over the existing methods.

Fig. 4 Decrypted Video

4. PERFORMANCE ANALYSIS

This section presents the performance analysis of the proposed Correlation Model (CM) based video encryption. The proposed estimation scheme used Motion Estimation (ME) for MSE measurements. The proposed scheme is also compared with the Reduced Reference (RR) based PSNR estimation on the ME performance. The evaluation of the motion estimation for various frame index values done and also compared with the PSNR estimation.

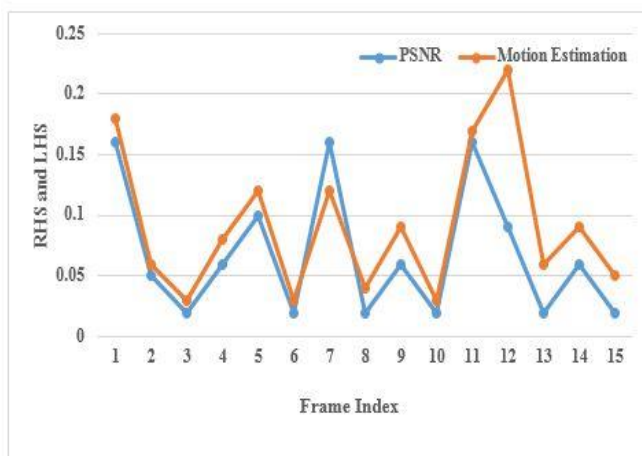


Fig. 5 Motion Estimation vs. PSNR Estimation

Fig. 5 describes the comparison between the existing PSNR and proposed motion estimation is carried out by Right Hand Side (RHS) and Left Hand Side (LHS) estimation. The estimation done for each frame index shows the efficiency of the proposed work in the encryption.

P-ISSN: 2347-4408
E-ISSN: 2347-4734

5. CONCLUSION

Secure and efficient transfer considered as the important criteria in the video processing in H.264/AVC Bit-streams. In this paper, the Advanced Encryption Standard (DES) and Correlation Model based bit streams transmitted the security of the transmitted bit streams are improved. So the proposed system mainly considered reduction of the de-blocking filter effect in the Mean Square Estimation (MSE). The quantization parameters and the deblocking effect in MPEG-2 I-frames are estimated with an overall accuracy of 99.9 % and the MSE is estimated with an overall average error of 0.3dB. The MATLAB simulation results proved the effectiveness of proposed method over the existing PSNR estimation. The comparative analysis between Advanced Encryption Standard (AES) and DES also carried out in future to confirms the effective secure transmission of video signals over the existing methodologies.

REFERENCES

[1] A. Panayides, Z. C. Antoniou, Y. Mylonas, M. S. Pattichis, A. Pitsillides, and C. S. Pattichis, "High-Resolution, Low-Delay, and Error-Resilient Medical Ultrasound Video Communication Using H.264/AVC Over Mobile WiMAX Networks," *IEEE Journal of Biomedical and Health Informatics,* vol. 17, pp. 619-628, 2013.

[2] D. Schroeder, A. El Essaili, E. Steinbach, D. Staehle, and M. Shehada, "Low-Complexity No-Reference PSNR Estimation for H.264/AVC Encoded Video," in *20th International Packet Video Workshop (PV), 2013* 2013, pp. 1-6.

[3] L. Xinhao, M. Tanaka, and M. Okutomi, "Noise level estimation using weak textured patches of a single noisy image," in *19th IEEE International Conference on Image Processing (ICIP), 2012* 2012, pp. 665-668.

[4] A. Rehman and W. Zhou, "Reduced-Reference Image Quality Assessment by Structural Similarity Estimation," *IEEE Transactions on Image Processing,* vol. 21, pp. 3378-3389, 2012.

[5] L. Chaofeng, A. C. Bovik, and W. Xiaojun, "Blind Image Quality Assessment Using a General Regression Neural Network," *IEEE Transactions on Neural Networks,* vol. 22, pp. 793-799, 2011.

[6] Y. Peng and D. Doermann, "No-Reference Image Quality Assessment Using Visual Codebooks," *IEEE Transactions on Image Processing,* vol. 21, pp. 3129-3138, 2012.

[7] H. Sabirin and K. Munchurl, "Moving Object Detection and Tracking Using a Spatio-Temporal Graph in H.264/AVC Bitstreams for Video Surveillance," *IEEE Transactions on Multimedia,* vol. 14, pp. 657-668, 2012.

[8] M. Zhan, H. Hao, and W. Yao, "On Complexity Modeling of H.264/AVC Video Decoding and Its Application for Energy Efficient Decoding," *IEEE Transactions on Multimedia,* vol. 13, pp. 1240-1255, 2011.

[9] N. Staelens, D. Deschrijver, E. Vladislavleva, B. Vermeulen, T. Dhaene, and P. Demeester, "Constructing a No-Reference H.264/AVC Bitstream-Based Video Quality Metric Using Genetic Programming-Based Symbolic Regression," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 23, pp. 1322-1333, 2013.

[10] R. Deng, X. Ding, Y. Wu, and Z. Wei, "Efficient block-based transparent encryption for H.264/SVC bitstreams," *Multimedia Systems,* vol. 20, pp. 165-178, 2014/03/01 2014.

[11] M. T. Pourazad, C. Doutre, M. Azimi, and P. Nasiopoulos, "HEVC: The New Gold Standard for Video Compression: How Does HEVC Compare with H.264/AVC?," *Consumer Electronics Magazine, IEEE,* vol. 1, pp. 36-46, 2012.

[12] U. Sadhvi Potluri, A. Madanayake, R. J. Cintra, F. M. Bayer, S. Kulasekera, and A. Edirisuriya, "Improved 8-Point Approximate DCT for Image and Video Compression Requiring Only 14 Additions," *IEEE Transactions on Circuits and Systems I: Regular Papers,* vol. 61, pp. 1727-1740, 2014.

[13] C. Yu-Lin and W. Chung-Bin, "A hardware sharing architecture of deblocking filter for VP8 and H.264/AVC video coding," in *IEEE International Symposium on Circuits and Systems (ISCAS), 2012* 2012, pp. 1915-1918.

[14] D. F. De Souza, N. Roma, and L. Sousa, "Cooperative CPU+GPU deblocking filter parallelization for high performance HEVC video codecs," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014* 2014, pp. 4993-4997.

[15] S. Liquan, L. Zhi, Z. Xinpeng, Z. Wenqiang, and Z. Zhaoyang, "An Effective CU Size Decision Method for HEVC Encoders," *IEEE Transactions on Multimedia,* vol. 15, pp. 465-470, 2013.

[16] P. Bestagini, A. Allam, S. Milani, M. Tagliasacchi, and S. Tubaro, "Video codec identification," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012*, 2012, pp. 2257-2260.

[17] J. Joskowicz and J. C. Ardao, "A parametric model for perceptual video quality estimation," *Telecommunication Systems,* vol. 49, pp. 49-62, 2012/01/01 2012.

[18] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, "Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison," *IEEE Transactions on Broadcasting,* vol. 57, pp. 165-182, 2011.

[19] M. Shengbin, S. Jun, D. Yizhou, and G. Zongming, "An efficient method for no-reference H.264/SVC bitstream extraction," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014*, 2014, pp. 1981-1985.

[20] R.-J. Wang, Y.-T. Jiang, J.-T. Fang, and P.-C. Chang, "Quality Estimation for H.264/SVC Inter-layer Residual Prediction in Spatial Scalability," in *Advances in Image and Video Technology*. vol. 7088, Y.-S. Ho, Ed., ed: Springer Berlin Heidelberg, 2012, pp. 252-261.