

# Optimal Global Threshold based on Two Dimension Otsu for Block Size Decision in Intra Prediction of H.264/AVC Coding

Sawsan Morkos Gharghory  
Computers and Systems Department  
Electronics Research Institute  
Giza, Egypt

**Abstract**—The Advanced Video Coding (H.264/AVC) has proved its ability in finding the tradeoff between the compressed bit rate value and the visual quality of video comparing to the others of traditional coding. One of the most encoder stages consuming time is the intra prediction in which different sizes of a block are exhaustively examined for selecting the suitable block size to the best block mode decision. In this paper, an efficient approach is suggested to select the best block size for the intra prediction adaptively to achieve high compression efficiency. The proposed approach exploits the idea of quad tree decomposition for blocks partitioning based on a predefined threshold value. An optimal global threshold value based on two dimension Otsu technique is suggested for the decision of block division in this work. The proposed technique is carried out on different set of videos resolutions with different quantization parameters using Matlab software. The comparison of the proposed approach with the reference JM18.6 video coding is done in terms of bit rate (BR), time saving and peak signal to noise ratio (PSNR). A tangible acceleration on the running time can be accomplished besides improvement in both of visual quality and bit rate with some of QCIF and CIF videos resolution by the proposed technique. The simulation results demonstrate saving in time by average 42% to 68% with CIF and QCIF videos. Concerning the visual quality in terms of Bjontegaard Delta parameters, the PSNR improved in which its value increased from 0.2 to 1.6, while the value of BR reduced from 0.79 to 15.3 respectively with some videos of resolutions QCIF, CIF and 720p. In addition, the performance of the suggested approach with the high resolution videos achieves minor improvement with some of them while has a slightly degradation with others of them.

**Keywords**—H.264/AVC coding; intra prediction; block size decision; Otsu two dimensions method

## I. INTRODUCTION

H.264/AVC coding [1] has becoming among the pioneers in video coding, it has proved its ability in achieving a preferable performance comparing to the others. The most important parameters for advancing H.264 coding performance are: the intra prediction to a block in spatial domain, the motion estimation for variable block size, reference frames, and finally the rate distortion optimization (RDO) [2-3]. Intra prediction in the spatial domain is one of the most encoder stages consuming time and is considered in this paper. The intra prediction process goal is to reduce the data size needed to the sufficient representation of the concerning block samples of video frame.

Intra prediction is used for editing of video sequence, coding of static image and finally for the significantly changes in the content of scene. The fast intra prediction algorithm was categorized into two types; the block size decision and block mode decision. As follow in the standard H.264/AVC, the determination of block size for the intra prediction was by examining all block sizes exhaustively to the nine modes or the four modes for optimal mode decision which is time consuming. Many fast algorithms have been presented concerning the block size decision for the reduction of computation time. Huang et al. [4] suggested fast block type selection for intra block coding based on the texture complexity of a macro-block (MB). Lin et al. [5] used the ratio of AC and DC coefficient energy to predict the intra prediction block size. Zhang et al. [6] proposed a fast block size decision algorithm based on the relationship between the smoothness of macro-blocks and the block type of intra prediction. The smoothness of a block was assessed by DC and AC coefficients in DCT domain. Wei et al. [7] used the variance of DC coefficients which is obtained from DCT for block size decisions. Many of the above mentioned researches suffered from determining the optimal threshold value required for preferring between different block sizes. So, the setting of optimum threshold value was a very important issue. If the threshold value is set too high, it may be resulted in error of block prediction which consequently leads to a dissipation in coding.

In this paper, the idea of quad tree decomposition based on optimal global threshold for block size decision is suggested for the intra prediction manner. Quad tree decomposition method in partitioning is based on the calculation of intensity variation (homogeneity) value within the macro blocks and its comparison to the global threshold value in order to decide whether performing a new partitioning or not. In this paper, the calculation of optimal global threshold value is based on the two dimension (2-D) Otsu, considering the between classes variance feature for block size decision. The 2-D Otsu method utilizes both of the gray level information of each pixel and its spatial correlation information within the neighborhood. The proposed technique is implemented on sequences of different videos with different quantization parameters (QP) and different resolutions such as QCIF, CIF and high resolution such as 720p and 1080p using Matlab Software. Both Peak signal to noise ratio (PSNR), time saving and bit rate are

considered as criterions for evaluating the efficiency of the proposed technique relative to the standard technique. The simulation results prove the superiority of the suggested approach compared to those of the reference JM18.6 in terms of the aforementioned metrics. The results of simulation achieve time saving by average 48% to 68% with CIF and QCIF videos accompanied with improvement in both of visual quality and bit rate with some of them. Beside, the performance of the suggested approach with the high resolution videos achieves a little improvement with some of them while has a slightly degradation with others of them either of visual quality or bit rate.

The rest of paper is organized as follows. The intra prediction is detailed in Section 2. Code tree decomposition and the suggested global threshold value based two dimension Otsu for block size decision is depicted in details in Section 3. Section 4 presents the proposed algorithm and its flow chart. Finally, the simulation results of the proposed algorithm to different videos sequences with different resolutions and different quantization parameters are reviewed in Section 5. Section 6 concludes the work and results discussion.

## II. INTRA PREDICTION

H.264/AVC performs the intra prediction to the block for reducing the redundancy in the spatial domain in various manners to the previous video coding which used the reconstructed pixels from the adjacent blocks. In intra prediction stage, H.264 code supports the luminance pixels in a macro block (MB) three intra modes representative to different blocks sizes which are: intra 4x4 (I4MB), intra 8x8 (I8MB), and intra 16x16 (I16MB). The luminance (luma) components in block with size 4x4 or 8x8 have nine prediction modes, while the luminance components in block size 16x16 have only four prediction modes [1]. On the other hand, chrominance (chroma) components have only 8 x 8 block size with four prediction modes. Practically, the blocks of small sizes are used for encoding the texture regions, while the large sizes blocks are often used for encoding the smooth (homogenous) regions. The luma sample label and the direction of intra prediction modes of I4MB are illustrated in Fig. 1 and 2, respectively [4].

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				

Fig. 1. Samples Labeling in Luma 4x4 Intra Prediction.

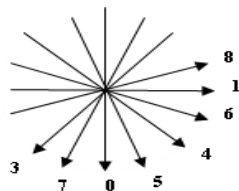


Fig. 2. The Directions of Prediction Modes in I4MB.

The sixteen lowercase letters from (a to p) in the mentioned figure are the pixels to be encoded, while the capital letters from A to M around the block are used for predicting this block. In Fig. 2, the second prediction mode that is dc mode is not shown because it has not direction and is estimated by calculating the average value to the number of luminance pixels from A to L surrounding the block to be coded shown in Fig. 1. Concerning the intra prediction coding of Luma 8x8, the number of its prediction modes besides their directions is similar to that in the Luma 4x4 and is only different in size. On the other hand, the encoding of both of Luma 16x16 and Chroma 8x8 blocks have three prediction modes in addition to the dc mode which can be calculated by the average intensity value to the surrounding coding block pixels [8-13]. Intra prediction modes for Chroma 8x8 and Luma 16x16 blocks are shown in Fig. 3.

In intra prediction stage of H.264/AVC, the encoder calculates and compares the rate of distortion (RD) of all prediction modes to different sizes of the block to be coded. Finding the optimal block size corresponding to the best prediction mode is implemented by minimizing the cost function for rate distortion (RD). Consequently, the intra prediction objective is to achieve an optimal visual quality for each block in the frame under certain bit rate cost and with a given quantization parameter (QP). The number of calculations for the RD before utilizing the suggested method in finding the optimal prediction mode with the suitable block size decision can be obtained by the following equation:

$$N = Mode_{(Chroma_8)} \times (Mode_{(Luma_4)} \times 16 + Mode_{(Luma_8)} \times 4 + Mode_{(Luma_{16})}) \quad (1)$$

Where:  $Modes_{(Chroma_8)}$ ,  $Modes_{(Luma_4)}$ ,  $Modes_{(Luma_8)}$  and  $Modes_{(Luma_{16})}$  are the prediction modes number for the coding blocks Chroma 8 x 8 which equals 4, Luma 4 x 4 with number of modes equaling 9, Luma 8 x 8 in which it has 9 modes and finally Luma 16 x 16 with 4 modes, respectively [14]. As depicted in equation (1), each number of luma prediction modes is multiplied by the number of coding blocks, in which there are sixteen coding blocks of size 4 x 4 and four coding blocks with size 8 x 8 in an MB. Consequently, the number of RDO calculations needs 763 computations in the latest JM version of H.264/AVC encoder for each macro block (MB) which is considered as a heavy computation [14]. With the previous version of JM reference software which is JM 9, the number of RDO computations equals 576 when the selected block size for coding is 4x4. The optimal mode for intra prediction is determined based on minimizing the cost function of RDO process as depicted in equation (2):

$$J = SSD(Mode|QP) + \lambda_{mode} \times R(Mode|QP) \quad (2)$$

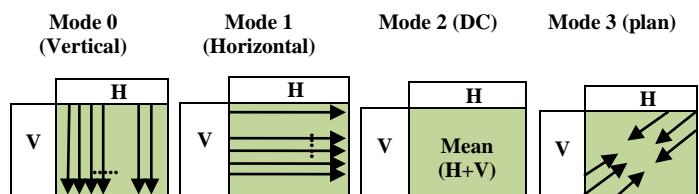


Fig. 3. Directions of Luma 16x16 and Chroma 8x8 Intra Predictions.

Where: SSD is summing the square difference between the original MB and its reconstructed pixels, QP is the parameter of quantization, R is the bit rate of encoding block in which  $\lambda_{mode}$  has a significant control on the rate of distortion and is determined by  $0.85 \times 2^{QP-\frac{12}{3}}$ , and finally J is the cost of the current intra-prediction mode [14]. To reduce the cost function aforementioned, the components of I-frame are partitioned into square sizes blocks and then the best block size corresponding to the optimal prediction mode for each MB in these components will be estimated through iterative computations of RDO.

Concerning the issue of reducing the number of RDO computations to find the suitable block size with its optimal mode, the block size decision based on the concept of code tree decomposition with global threshold based on two dimension Otsu is proposed. Both of the suggested tree decomposition code and two dimension Otsu method for determining the global threshold value is detailed in the following sections.

### III. QUAD TREE DECOMPOSITION

Quad- tree approach is used to split an image into partitions based on the measurement of blocks homogeneity such as coefficient variation or variance. The decomposition code of quad-tree is initialized firstly by partitioning an image into four equal size parts named nodes, each of them represents either a macro block (MB), or it is considered as a leaf. To evaluate if the macro block represented by the node homogeneous or not is through a performed test comparing its intensity variation with the variation of intensity of its parent. If the test is negative the node becomes a MB and it is divided, otherwise the node not divided and became a leaf. The partitioning process is iteratively for all nodes represented by MB till reach to the desired block size in image. Each parent node representing a macro block in the tree has four children called micro blocks (mB) in which each one has a quarter of its parent area. The structure of the tree and the implementation on image are depicted in Fig. (4a), (4b) and (4c). As shown in Fig. (4c), the content of the image or video frame is divided into four macro blocks ( $MB_k$ ), and each  $MB_k$  will be divided into four micro-blocks ( $mB_{ki}$ ) or not based on its homogeneity and then each  $mB_{ki}$  will be iteratively post subdivided or not based on the achievement of this criterion. In this paper, the optimal threshold based variance using two dimensions Otsu method is suggested to determine whether the block is homogenous or not so that the division decision can be taken [15].

The proposed global threshold based on two dimensions Otsu is presented in the following section in details.

#### A. Two Dimension Otsu Technique

Normally, the Otsu's method calculates the between-classes and within-class variance using the one Dimension (1-D) image histogram. To determine the optimal threshold value, the cost function for minimizing the within-class variance or maximizing the between-classes variance is used. This concept is extended to multilevel threshold as well. It is noted that 1-D Otsu's method is considered as the fastest and the simplest of all methods because the threshold value is mostly determined by the 1-D histogram of the image. However, 1-D histogram does not depict the spatial correlation between the pixels in

image or video frame. For this reason, images having complex boundaries may yield a poor threshold performance. Therefore, Otsu's 1-D method is extended to the 2-D version, where the 2-D histogram of an image is used [16-18]. Also, the 2-D Otsu's criteria has been extended to be implemented to multilevel threshold. The between class variance concept of 2-D Otsu is presented below in details.

Let I represent a gray scale image of size  $M \times N$  with L gray levels  $g = [1, 2, \dots, L]$  and total number of pixels equals  $M \times N$ . Also, suppose  $f(x, y)$  is the gray value of the pixel located at coordinate  $(x, y)$  where  $x \in \{1, 2, \dots, M\}$ ,  $y \in \{1, 2, \dots, N\}$ . Let  $h(x, y)$  represents the local average gray value for a window of size  $w \times w$  with center at the location  $(x, y)$ , and is given by equation 3:

$$h(x, y) = \frac{1}{w \times w} \sum_{m=-\frac{w-1}{2}}^{\frac{w-1}{2}} \sum_{n=-\frac{w-1}{2}}^{\frac{w-1}{2}} f(x + m, y + n) \quad (3)$$

The window size  $w$  is odd and may be equal to  $3 \times 3$  or  $5 \times 5$  and so on. Let  $q_{ij}$  is the occurrence time of the pair  $(i, j)$ , where  $f(x, y) = i$  and  $h(x, y) = j$ . Then, the frequency of occurrence of pair  $(i, j)$  is calculated as follows:

$$P_{ij} = \frac{q_{ij}}{M \times N}, 1 \leq i \leq L, 1 \leq j \leq L, \text{ and } L=256 \quad (4)$$

If the image is segmented into two classes  $C_0$  and  $C_1$  (background and objects) by a threshold at level value  $k$ , in which the background  $C_0$  contains pixels taking the level values from: 1 to  $k$ , and  $C_1$  contains pixels with levels  $[k + 1, \dots, L]$ . If the current threshold value is  $(t, s)$  where  $t$  named the gray threshold and  $s$  is the local average threshold, the probability distribution of the background and object classes  $\omega_0$  and  $\omega_1$  are calculated respectively, through equations 5 and 6. Also, the mean class vectors  $\mu_0$  and  $\mu_1$  for each of them with gray and local average features are calculated by the equations from 7 to 8.

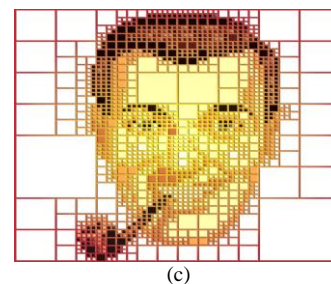
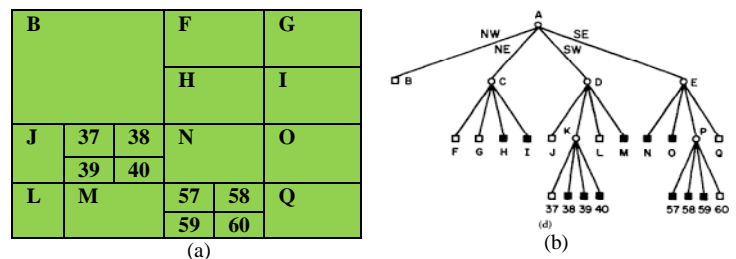


Fig. 4. (a-c) Quad- Tree Decomposition Approach.

$$\omega_0 = P(C_0) = \sum_{i=1}^S \sum_{j=1}^t P_{i,j} \quad (5)$$

$$\omega_1 = P(C_1) = \sum_{i=S+1}^L \sum_{j=t+1}^L P_{i,j} = 1 - \omega(k) \quad (6)$$

$$\mu_0 = (\mu_{00}, \mu_{01})^T = \left[ \frac{\sum_{i=1}^S \sum_{j=1}^t i * P_{i,j}}{\omega_0}, \frac{\sum_{i=1}^S \sum_{j=1}^t j * P_{i,j}}{\omega_0} \right]^T = \frac{\mu(k)}{\omega(k)} \quad (7)$$

$$\mu_1 = (\mu_{10}, \mu_{11})^T = \left[ \frac{\sum_{i=S+1}^L \sum_{j=t+1}^L i * P_{i,j}}{\omega_1}, \frac{\sum_{i=S+1}^L \sum_{j=t+1}^L j * P_{i,j}}{\omega_1} \right]^T = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \quad (8)$$

Where:  $\omega(k) = \sum_{i=1}^k P_i$   $\mu(k) = \sum_{i=1}^k i * P_i$

Both of  $\omega(k)$  and  $\mu(k)$  are the zero-order and the first-order cumulative moments of the histogram up to the  $k^{\text{th}}$  level. The total mean vector  $\mu_T$  is given by equation 9:

$$\mu_T = (\mu_{T0}, \mu_{T1})^T = \left[ \sum_{i=1}^L \sum_{j=1}^L i * P_{i,j}, \sum_{i=1}^L \sum_{j=1}^L j * P_{i,j} \right]^T \quad (9)$$

Where:  $\omega_0 + \omega_1 \approx 1$  and  $\omega_0 \mu_0 + \omega_1 \mu_1 \approx \mu_T$  and  $\omega_1 \approx 1 - \omega(k)$

The two classes variance are calculated as described in equations 10 to 11 as follows:

$$\sigma_0^2(k) = \sum_{i=1}^k (i - \mu_0)^2 P_{i,j} / \omega_0 \quad (10)$$

$$\sigma_1^2(k) = \sum_{i=k+1}^L (i - \mu_1)^2 P_{i,j} / \omega_1 \quad (11)$$

The within class variance is given by equation 12:

$$\sigma_W^2 = \omega_0 \mu_0^2 + \omega_1 \mu_1^2 \quad (12)$$

The between class variance  $\sigma_b^2$  is computed by:

$$\begin{aligned} \sigma_b^2 &= \omega_0 ((\mu_0 - \mu_T)^2) + \omega_1 ((\mu_1 - \mu_T)^2) \\ \sigma_b^2 &= \omega_0 ((\mu_{00} - \mu_{T0})^2 + (\mu_{01} - \mu_{T1})^2) + \omega_1 ((\mu_{10} - \mu_{T0})^2 + (\mu_{11} - \mu_{T1})^2) \end{aligned} \quad (13)$$

For simplification, by substituting in equation 13 using equation 5 and 6, the between classes variance  $\sigma_b^2$  is given by equation the equation 14:

$$\sigma_b^2(k) = \frac{(\mu(k) - \mu_T \omega(k))^2}{\omega(k)(1 - \omega(k))} \quad (14)$$

The between class variance  $\sigma_b^2$  is calculated by the mean of class subtracted from the total mean level of the original frame multiplied by probability of occurrence of this class divided by the multiplication of probability of this class occurrence and the probability of the other class occurrence.

The maximization of the between class variance value gives the optimal threshold for image partition as depicted in equation 15:

$$t_{opt}, S_{opt} = \arg \max_{t,S} \{ \sigma_b^2(t, S) \} \text{ in which } 1 \leq t, S \leq L \quad (15)$$

The optimal threshold value  $k^*$  is calculated by the maximizing  $\sigma_b^2(k)$ , which is termed by  $\max_{1 \leq k \leq L} \sigma_b^2(k)$ . Fig. 5 (d to f) shows the implementation of the two dimension Otsu method to the I-frame from sequences of different test videos with different resolutions such as CIF and 1080p as shown in Fig. 5 (a to c).



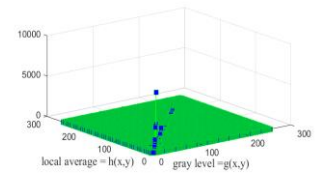
(a) I-frame of News video with resolution CIF (352x288)



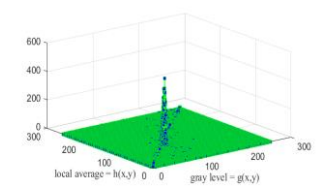
(b) I-frame of football video with resolution CIF (352x240)



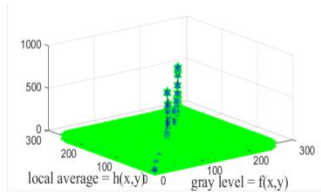
(c) I-frame of Highway video with resolution CIF (352x192)



(d) Two dimension Otsu for I-frame of News video



(e) Two dimension Otsu for I-frame of football video



(f) Two dimension Otsu for I-frame of Highway video

Fig. 5. (a-c) Sequences of Videos with different Resolutions and (d-f) the Two Dimension Otsu for the Test Videos

Variance-based optimal threshold using 2-D Otsu method is suggested in this paper to partition the I-frame into blocks with different sizes which is presented in details in the following section.

#### IV. THE PROPOSED TECHNIQUE

This paper suggests the determination of optimal threshold value based on automatic algorithm for partitioning the frame under consideration into different blocks sizes. The proposed algorithm is employed for block size decision in intra prediction stage in which it utilizes the idea of quad-tree decomposition technique based on a global threshold value determined by two dimension Otsu method. An input frame will be divided into blocks either of  $4 \times 4$  or  $16 \times 16$  pixels according to the smoothness or the complication of the macro block through utilizing the statistical measurements of homogeneity feature of block. Each large square block in the frame will be divided if the homogeneity feature measured by the difference between maximum and minimum intensity values within the block exceeds a threshold value  $T$ , otherwise, the block is not divided. Thus, the value of threshold plays a very crucial role in which the independent selection to the threshold value without considering input frame statistical features may lead to bad compression characteristics. This problem is addressed in this paper by utilizing the aforementioned 2-D Otsu method to calculate the optimal threshold value which consequently leads to accurate decision about the blocks division into sub blocks. Once the frame is optimally decomposed into suitable blocks, the intra prediction to the block with final size either  $4 \times 4$  or  $16 \times 16$  are implemented using the corresponding prediction modes. After that, the predicted block is passing by the following processes: the DCT transformation, the quantization to the difference



between the original block and the predicted block, the entropy coding on this quantized difference and finally the inverse process to both the quantization and DCT transformation. The flowchart of the proposed technique is shown in Fig. 6. Also The Pseudo code concerning the block size decision in H.264/AVC intra prediction coding is described as follows.

The Pseudo code of the proposed approach

1) Determine the optimal global threshold to the I frame from the input video sequence using the variance-based two dimension Otsu method.

2) Then, I frame is divided into equally blocks of sizes 16x16, decompose each one of these blocks into sub-blocks of size 4x4 using the idea of Quad tree decomposition based on the optimal threshold value T calculated from the previous step.

3) Check if the difference between maximum and minimum intensity values of the square block exceeds the optimal threshold value T, it means the block is non-homogeneous and it should be divided, otherwise the block is not divided.

4) Finally the whole frame is becoming blocks with 4x4 or 16x16 sizes.

5) The resulted block is encoded using 4x4 intra prediction if its size 4x4, which it is well suited for coding the region of significant details, otherwise it is encoded by 16x16 intra prediction which is suitable for the smoothness region [1].

6) After that, the difference between the original and the predicted block called the residual coefficients block are calculated, transformed by DCT, quantized, coded by the entropy encoding that yields the bits of the quantized residual coefficients of block and finished by the bit streams for all I-frame blocks.

7) On the other path of the quantization process, the inverse quantization and the inverse transformation is applied to the residual coefficients block. Finally the re-constructed block is built from adding the predicted block to that block of the residual coefficients.

8) The result of encoding process is frame bits streams and re-constructed frame after encoding.

The metrics used to evaluate the efficiency of the proposed method comparing to the standard H.264 video coding are PSNR, run time saving and bite rate of the compression process and they are calculated in the equations 16-18.

$$\Delta PSNR = \frac{PSNR(Proposed) - PSNR(JM)}{PSNR(JM)} \times 100\% \quad (16)$$

$$\Delta Time = \frac{Time(Proposed) - Time(JM)}{Time(JM)} \times 100\% \quad (17)$$

$$\Delta Bitrate = \frac{Bitrate(Proposed) - Bitrate(jm)}{Bitrate(jm)} \times 100\% \quad (18)$$

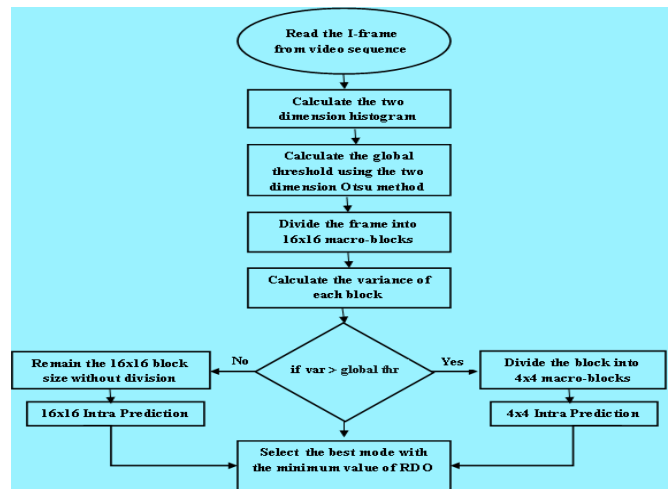


Fig. 6. The Flowchart of the Proposed Technique with the Mode Selection using Minimum ORD.

Also, the Sum of Absolute Differences (SAD) is used as other metric for measuring the error between the original frame and the predicted frame in intra prediction stage. Each of PSNR and SAD is depicted in the following equations.

$$PSNR = 10 \log \frac{255^2}{MSE} \quad (19)$$

$$MSE = \frac{SSD}{H \times W} \quad (20)$$

$$SSD = \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} (p(x, y) - p_r(x, y))^2 \quad (21)$$

$$SAD = \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} |p(x, y) - p_r(x, y)| \quad (22)$$

Where: H and W are the height and the width of the frame, each of  $p(x, y)$  and  $p_r(x, y)$  represents the original and reconstructed frame pixels at location x and y, as well as both of MSE and SSD are Mean Square Error and Sum of Square Differences between the predicted and the original blocks respectively.

## V. SIMULATION RESULTS

The proposed block size decision in intra prediction based on global threshold using 2-D Otsu method is employed for different sets of test videos sequences with different resolutions. The performance of the proposed algorithm is evaluated and compared numerically with the standard version of H.264/AVC (JM18.6) in terms of Bjontegaard Delta bit rate, Bjontegaard Delta PSNR and run time [19]. Different set of videos with different resolutions at frame rate of 30Hz are used for testing and are depicted in Table I as follows:

TABLE I. SET OF VIDEOS SEQUENCES WITH DIFFERENT RESOLUTIONS

Videos Resolution	Tested Video Sequences
1080p	Station2 and Pedestrian (1920x1080)
720p	Johnny and kristenandsara (1280x720)
576p	Harbour and Soccer (704x576p)
CIF	Stefan, Container, Hall and Mother-daughter (352x288)
QCIF	Claire, Akiyo, Hall and News (176x144)

The setting parameters for the proposed technique are:

- 1) A group of III.....I frames from the test videos for luminance components of the frames, while the intra prediction based block size 8 is applied on the chrominance components of I frames.
- 2) Entropy coding of Context Adaptive Variable Length Coding (CAVLC),
- 3) Quantization parameters (QP) are setting at 20, 24, 28 and 32.

The performance of the proposed algorithm is evaluated through its comparison to the standard reference video coding JM 18.6 and the results of simulation in terms of different metrics are presented to demonstrate the suggested approach efficiency. The simulation results and BD curves of the proposed method using the Bjontegaard Delta (BD) parameters to the set of test videos with resolutions QCIF, CIF, 576p, 720p and 1080p comparing to that of the reference coding JM18.6 are visually shown in Fig. 7 (a to d), Fig. 8 (a to d) and Fig. 9 (a to f), respectively. Additionally, the numerical results to the performance of the proposed algorithm in terms of BD measurements of both bit rate (BDBR) and peak signal to noise ratio (BDPSNR) are depicted in Tables II and III for the video sequence with different resolutions including the high definition (HD) video resolution 1080p. While the setting parameters for JM18.6 are: enabling for both of RD-Optimization, RD-quantization and slice mode equals zero. Also, QP equals 20, 24, 28 and 32 and CAVLC are used with the reference code JM18.6.

Where: BD\_PSNR and BD\_ Bit Rate are the average PSNR difference and average percentage of time saving between the reference JM18.6 video coding and the proposed technique, respectively.

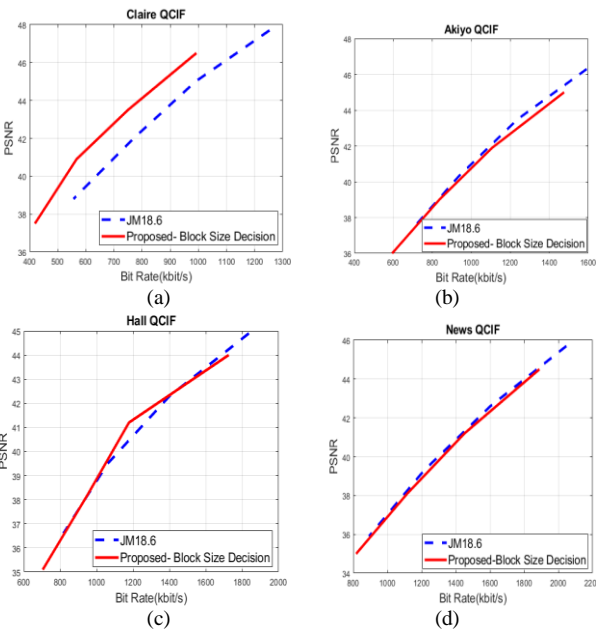


Fig. 7. (a-d) BD Curves for QCIF Test Video Sequences with the Proposed Approach and JM18.6.

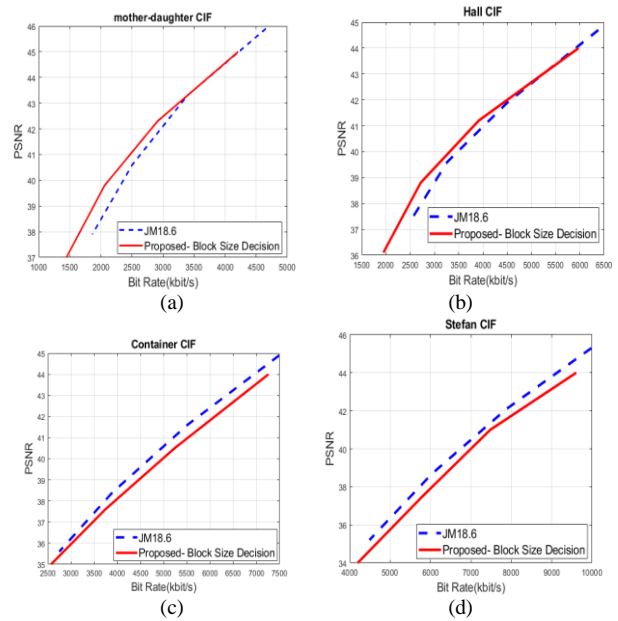


Fig. 8. (a-d) BD Curves for CIF Videos Sequences to the Proposed Approach and JM18.6.

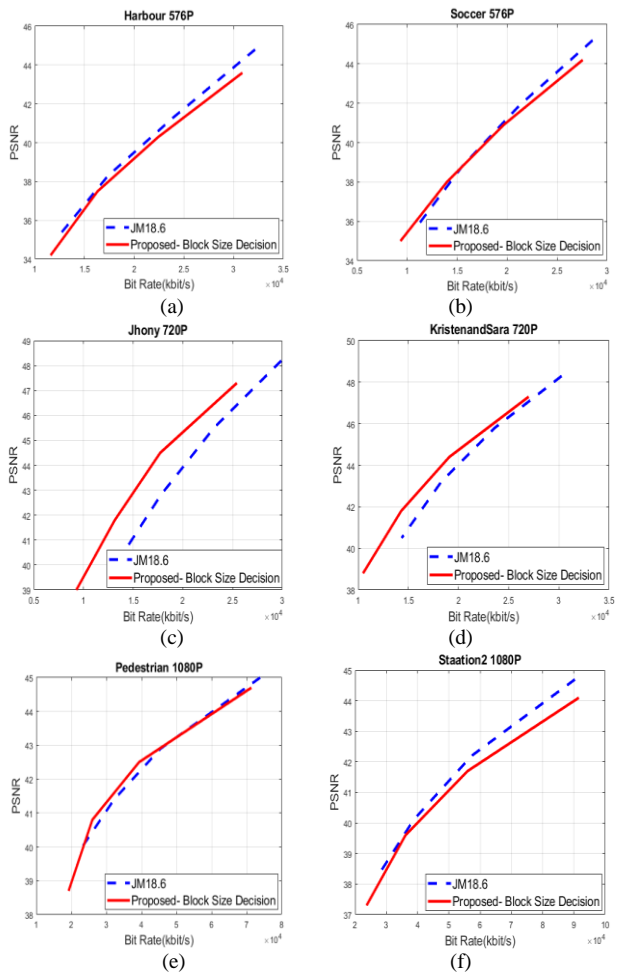


Fig. 9. (a-f) BD Curves for Test Video Sequences of 576p, 720p and 1080p Resolutions to the Proposed Approach and JM18.6.

The sign (-) means value decreasing while the sign (+) means increasing in value. As demonstrated in Table III, the proposed method achieve best simulation results for the video sequences with QCIF resolution in which the values of BDPSNR increased from 0.2 to 1.6 for both of videos sequence Claire and hall, and the values of BDBR to aforementioned videos decreased by percentage values range from 0.79 to 14.9. In contrast, with the videos sequences Akiyo and News QCIF resolution; the BD-PSNR values slightly decreased by percentage values ranging from 0.17 to 0.24 and the BDBR increased from 1.5 to 2.3, respectively.

Concerning the simulation results for the videos sequences with CIF resolution, the values of BDPSNR increased for both of Mother-daughter and Hall videos from 0.49 to 0.58 respectively, while slightly decreased for both of the videos sequence Stefan and Container from 0.47 to 0.6, respectively. On the other hand, the values of BDBR decreased from 5.5 to

6.2, respectively for both of the videos sequences Mother-daughter and Hall, while BDBR value increased from 5.1 to 5.3 for the videos sequences Stefan and Container. The percentage values in the time saving of the proposed algorithm for the videos sequences with QCIF and CIF resolutions was from 0.42 to 0.68.

Regarding the high resolutions videos, the simulation results of BDPSNR of the proposed technique as depicted in Table II slightly decreased by percentage values from 0.005 to 0.3 for the videos sequence with resolutions 576p and 1080p. While BDPSNR increased from 0.96 to 1.5 for the videos sequences with 720p resolution. The results of BDBR increased by percentage values from 0.4 to 3.5 for both of the videos sequences with resolutions 576p and Station2 video with 1080p resolution. While, BDBR decreased from 1.8 to 15.3 for 720p video sequences and Pedestrian video with 1080p.

TABLE. II. THE COMPARISON OF THE PROPOSED APPROACH IN TERMS OF BDPSNR AND BDBR TO HIGH RESOLUTIONS VIDEOS: 576P,720P AND 1080P

Reso-lution	Video Test	QP	PSNR (dB)			Bit rate (Kbits)		
			JM 18.6	Proposed	BD PSNR	JM 18.6	Proposed	BD BR%
576P	Harbour	20	<b>44.8</b>	43.6	-0.35	<b>32165</b>	30880	3.59
		24	<b>41.3</b>	40.3		<b>23937</b>	22436	
		28	<b>38.3</b>	37.5		<b>17368</b>	16321	
		32	<b>35.2</b>	34.2		<b>12405</b>	11588	
576P	Soccer	20	<b>45.2</b>	44.2	-0.005	<b>28554</b>	27562	0.45
		24	<b>41.9</b>	40.9		<b>21255</b>	19660	
		28	<b>38.8</b>	38		<b>15520</b>	13980	
		32	<b>35.7</b>	35		<b>10927</b>	9337	
720P	Johnny	20	<b>48.2</b>	47.3	1.51	<b>29901</b>	25428	-15.3
		24	<b>45.5</b>	44.5		<b>23115</b>	17715	
		28	<b>42.8</b>	41.8		<b>17769</b>	13145	
		32	<b>40.4</b>	39		<b>13870</b>	9284	
720P	Kristen And Sara	20	<b>48.3</b>	47.3	0.96	<b>30278</b>	26991	-8.27
		24	<b>45.8</b>	44.4		<b>23606</b>	19083	
		28	<b>43.2</b>	41.8		<b>18272</b>	14285	
		32	<b>40.5</b>	38.8		<b>14326</b>	10484	
1080P	Station2	20	<b>44.7</b>	44.1	-0.3	90346	91590	0.4
		24	<b>42.2</b>	41.7		57316	56010	
		28	<b>40.1</b>	39.6		38707	36151	
		32	<b>38.3</b>	37.3		27509	23681	
1080P	Pedestrian	20	<b>45</b>	<b>44.7</b>	-0.07	73870	71339	-1.8
		24	<b>42.9</b>	<b>42.5</b>		45419	39282	
		28	<b>41.4</b>	<b>40.8</b>		32186	25991	
		32	<b>40</b>	<b>38.7</b>		23104	19217	

TABLE. III. THE COMPARISON OF THE PROPOSED APPROACH WITH JM18.6 IN TERMS OF BDPSNR AND BDBR TO QCIF, CIF RESOLUTIONS VIDEOS AND THE RUNNING TIME

Resolution	Sequence	QP	PSNR (dB)			Bit rate (Kbits)			Running Time (Sec)		
			JM 18.6	Proposed	BDPSNR	JM 18.6	Proposed	BDBR %	JM 18.6	Proposed	Average Time Saving %
QCIF	Claire	20	47.7	46.4	1.6124	<b>1253</b>	993	-14.953	<b>47</b>	11.5	% -56
		24	45	43.5		<b>990</b>	750		<b>34</b>	11	
		28	42	40.9		<b>770</b>	567		<b>26</b>	10.8	
		32	38.8	37.5		<b>556</b>	419		<b>20</b>	10.5	
QCIF	Hall	20	44.9	44	0.2041	<b>1837</b>	1726	-0.7940	<b>71</b>	13.3	% -68
		24	42.3	41.2		<b>1401</b>	1178		<b>55</b>	12.5	
		28	39.5	38.2		<b>1061</b>	950		<b>41.5</b>	12	
		32	36.6	35.1		<b>817</b>	704		<b>31.3</b>	11.5	
QCIF	Akiyo	20	46.3	45	-0.2409	<b>1592</b>	1476	2.3738	<b>56</b>	12.8	% -57
		24	43.5	41.9		<b>1238</b>	1107		<b>42</b>	12.3	
		28	40.3	39		<b>935</b>	836		<b>31</b>	11.8	
		32	37.3	36		<b>692</b>	594		<b>23</b>	11.6	
QCIF	News	20	45.7	44.5	-0.17	<b>2047</b>	1889	1.5	<b>82</b>	16	% -68
		24	42.6	41.2		<b>1602</b>	1451		<b>64</b>	13	
		28	39.3	38.1		<b>1215</b>	1116		<b>48</b>	12.8	
		32	35.9	35		<b>892</b>	817		<b>35</b>	12.7	
CIF	Mother-daughter	20	45.9	45	0.5811	<b>4656</b>	4209	-6.2416	<b>156</b>	44	% -42
		24	43.3	42.3		<b>3392</b>	2912		<b>113</b>	42	
		28	40.6	39.8		<b>2503</b>	2059		<b>85</b>	41	
		32	37.9	37		<b>1862</b>	1444		<b>69</b>	40	
CIF	Hall	20	44.7	44	0.4973	<b>6382</b>	5972	-5.5747	<b>228</b>	44	% -61
		24	41.9	41.2		<b>4506</b>	3914		<b>165</b>	42	
		28	39.6	38.8		<b>3263</b>	2716		<b>124</b>	40	
		32	37.3	36.1		<b>2497</b>	1946		<b>96</b>	37	
CIF	Stefan	20	45.3	44	-0.6	<b>9982</b>	9605	5.1	<b>369</b>	61	% -68
		24	42.1	41		<b>7882</b>	7478		<b>293</b>	58	
		28	38.7	37.5		<b>6040</b>	5800		<b>216</b>	54	
		32	35.2	34		<b>4487</b>	4196		<b>159</b>	52	
CIF	Container	20	44.9	44	-0.47	<b>7494</b>	7260	5.3	<b>304</b>	47	% -66
		24	41.6	40.5		<b>5509</b>	5235		<b>227</b>	44	
		28	38.4	37.6		<b>3891</b>	3738		<b>161</b>	41	
		32	35.6	35		<b>2744</b>	2568		<b>114</b>	40	



## VI. CONCLUSION

In this paper, an efficient technique for block size decision to the intra prediction stage of H.264 video coding is suggested. An optimal global threshold based on the 2-D Otsu method is suggested for block partitioning. The proposed method considers the between classes variance as feature for the block size decision. The suggested technique achieved acceptable results comparing to the reference JM18.6 video coding. By the proposed approach, a tangible acceleration on the code running time was accomplished with improvement in both of PSNR and bit rate with some of CIF and QCIF videos. While with some others CIF and QCIF videos, the running time acceleration was accompanied with slightly degradation in the visual quality and increasing in bit rate. Also with high video resolutions, the increasing of PSNR may be accompanied occasionally with decreasing in bit rate while it may be accompanied with increasing in bit rate with others. The results of simulation demonstrate saving in time by average 42% to 68% with QCIF and CIF videos. Besides, the visual quality in terms of BDPSNR improved in which its value increased from 0.2 to 1.6, while the value of BDBR reduced from 0.79 to 15.3 respectively with some videos of resolutions QCIF, CIF and 720p. While with high videos resolutions there is some improvement in visual quality but on expense of running time and increasing in bit rate.

## REFERENCES

- [1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, document JVT-G050.doc, ITU-T Rec. H.264 and ISO/IEC 14496-10 AVC, 2003.
- [2] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of H.264 video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.
- [3] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74–90, Nov. 1998.
- [4] Yi-Hsin Huang; Tao-Sheng Ou; Chen, H.H.; , "Fast decision of block size, prediction mode, and intra block for H.264 intra prediction," *Circuits and Systems for Video Technology*, *IEEE Transactions on* , vol.20, no.8, pp.1122-1132, Aug. 2010.
- [5] Y. K. Lin and T. S. Chang, "Fast block type decision algorithm for intra prediction in H.264 frext," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2005, vol. 1, pp. 585–588.
- [6] Tianruo Zhang, Guifen Tian, Satoshi Goto, "A frequency-based fast block type decision algorithm for intra prediction in H.264/AVC high profile," *Circuits and Systems*, 2008. APCCAS 2008. *IEEE Asia Pacific Conference on* , vol., no., pp.1292-1295, Nov. 30 2008-Dec. 3 2008.
- [7] Yung-Chiang Wei, Chao-Hsuing Tseng, "Transformed domain block size and intra mode decision for advanced video coding," *Computer Communication Control and Automation (3CA)*, 2010 International Symposium on , vol.1, no., pp.221-224, 5-7 May 2010.
- [8] J. Ostermann et al., "Video coding with H.264/AVC: Tools, performance, and complexity," *IEEE Circuits Syst. Mag.*, vol. 4, no. 1, pp. 7–28, Jan./Mar. 2004.
- [9] S.-K. Kwon, A. Tamhankar, and K. R. Rao, "Overview of H.264/MPEG-4 part 10," *J. Vis. Commun. Image Represent.*, vol. 17, no. 2, pp. 186–216, Apr. 2006.
- [10] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.
- [11] Sara Hamdy, AbdelhalimZekry and Wael A. Mohamed , "Development of the best prediction matrix mode algorithm for-intra prediction in h.264 encoder", *International Journal of Engineering & Technology*, 7 (4) (2018) 3847-3853.
- [12] Manjanaik.Na, Parameshachari B Db Hanumanthappa S Nc, Reshma Banud, "Intra Frame Coding In Advanced Video Coding Standard (H.264) to Obtain Consistent PSNR and Reduce Bit Rate for Diagonal Down Left Mode Using Gaussian Pulse" *IOP Conf. Series: Materials Science and Engineering* 225 (2017) 012209, doi:10.1088/1757-899X/225/1/012209.
- [13] Kyungmin Lim, Seongwan Kim, Jaeho Lee, Daehyun Pak and Sangyoung Lee, "Fast Block Size and Mode Decision Algorithm for Intra Prediction in H.264/AVC", *IEEE Transactions on Consumer Electronics*, Vol. 58, No. 2, May 2012.
- [14] Lih-Jen Kau, Member and Jia-Wei Leng, " A Gradient Intensity-Adapted Algorithm With Adaptive Selection Strategy for the Fast Decision of H.264/AVC Intra-Prediction Modes", *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, VOL. 25, NO. 6, pp: 944-957, JUNE 2015.
- [15] Carlos Mera and Maria Trujillo, "Using Dispersion Measures for Detering Block-Size In-Motion Estimation", *DYNA*, Vol. 79, Num. 171 , pp. 97-104 , (2012).
- [16] N. Otsu, "A threshold selection method from gray level histograms", *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS*, VOL. 9, No. 1, pp. 62-66, JANUARY 1979.
- [17] Wuli Wang, Liming Duan and Yong Wang, "Fast Image Segmentation Using Two-Dimensional Otsu Based on Estimation of Distribution Algorithm", *Journal of Electrical and Computer Engineering*, Article ID 1735176, PP.1-12 , September 2017,Hindawi, <https://doi.org/10.1155/2017/1735176>.
- [18] Fangyan Nie , Yonglin Wang , Meisen Pan , Guanghan Peng and Pingfeng Zhang, "Two-dimensional extension of variance-based thresholding for image segmentation", (2013) 24:485–501, © Springer Science+Business Media, DOI 10.1007/s11045-012-0174-7
- [19] G. Bjontegaard, "Calculation of average PSNR differences between RD curves," presented at the 13th VCEG-M33 Meeting, Austin, TX, Apr. 2001.