Proceedings of the International MultiConference of Engineers and Computer Scientists 2008 Vol I
IMECS 2008, 19-21 March, 2008, Hong Kong

Fast Wavelet-based Macro-block Selection Algorithm for H.264 Video Codec

Shi-Huang Chen, Chung-Hsien Chang, and Shih-Yin Yu

***Abstract* —This paper proposes a fast wavelet-based macro-block mode selection algorithm for H.264/AVC video codec system. The proposed algorithm makes use of the two-dimensional wavelet transform to estimate the sub-band energy of each macro-block in a given video frame. Then the sub-band energy can be applied to the macro-block mode decision as a primary parameter. To prevent the over segmentation of a macro-block, a secondary parameter called PSNR threshold is used in the mode decision procedure. Various experimental results show that the proposed algorithm can effectively make a macro-block mode decision under a given PSNR value. Furthermore, the execution time of the proposed wavelet-based multi-block selection algorithm is faster 1~4 times than that of the traditional algorithm with the similar macro-block mode decision results.**

***Index Terms*— H.264 video codec, macro-block selection, wavelet transform**

I. Introduction

H.264/AVC is the recently developed video coding standard [1] and has been shown that it outperforms previous standards such as MPEG-4 and H.263 in terms of coding efficiency [2]. The basic encoding algorithm of H.264 is very similar to H.26x or MPEG-x except that integer 4×4 discrete cosine transform (DCT) is used instead of the traditional 8×8 DCT. In addition, H.264 adopts several new coding tools, such as quarter-pixel-accuracy motion estimation (ME), multiple reference frames, loop filter, variable block size (VBS), and etc. [1-3]. These new coding tools enable H.264 to achieve higher coding efficiency than previous video coding standards. However, the encoder complexity increases tremendously. In general, most of the computational time is consumed in the mode decision stage where the variable block size ME is employed [4]. Various block size ME can reduce the residual of motion compensation (MC) to improve the compression ratio. In H.264, there are seven different block sizes (16×16, 16×8, 8×16, 8×8, 8×4, 4×8 and 4×4) as shown in Figure 1 that can be used in inter mode coding.

To achieve the highest coding efficiency, the H.264 encoder will test all of the possible modes and then chooses the best one in terms of least rate-distortion (RD) cost. However, such a “try and error” method will cost high computational complexity and limit the use of H.264 encoders in real-time applications [4].

This work was supported in part by the National Science Council under Grant NSC 96-2221-E-214-004 and NSC 96-2221-E-366-011.
Shi-Huang Chen*, Chung-Hsien Chang, and Shih-Yin Yu are with the Department of Computer Science Information Engineering, Shu-Te University, Kaohsiung, 824, Taiwan. (*Corresponding author, Tel: +886-7-6158000 ext. 4607, Fax: +886-7-6158000 ext. 4699, e-mail: shchen@mail.stu.edu.tw)

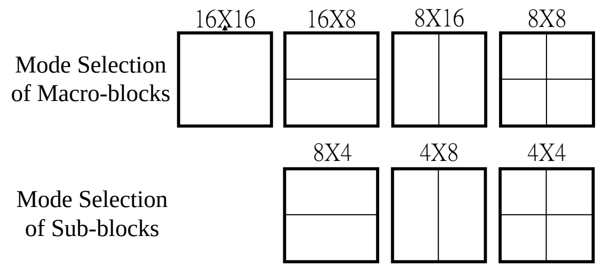


Figure. 1－Micro-block Mode Selection of Different Sizes Used in H.264 Encoder.

Therefore, many approaches have been proposed to reduce the H.264 encoder complexity [5-7]. Yin et al. [5] propose a monotonic error surface based algorithm to optimize mode decision. Their ME algorithm consists of two steps. First, an integer-pixel ME is performed based on the enhanced prediction zonal search (EPZS). Then, using the result of the integer-pixel ME, sub-pixel ME is carried out within some limited areas. For faster macro block mode selection, this method simply examines limited modes, i.e., 16×16, 8×8, and 4×4. In [6], the mean absolute frame difference of the current frame and mean absolute difference of the current MB from the previous frame are used to determine whether the current MB belongs to homogeneous regions or not. Therefore, it only needs to check 16×16, 16×8 and 8×16 modes. L. Yang et al. [7] proposed a variable block size best motion detection (VBBMD) algorithm to check the predicted motion vector (PMV) for different block sizes. If the RD cost with the PMV is less than a given threshold, the ME of this block size can be early stopped. In this algorithm, a set of principal thresholds is determined experimentally based on test sequence.

This paper proposes a new fast wavelet-based macro-block mode selection algorithm for H.264 encoding system. The proposed algorithm first calculates the sub-band energy by using two-dimensional wavelet transform. The wavelet transform is applied to pre-analysis the macro-block characteristic in a video frame. Then the proposed algorithm will derive a primary parameter for macro-block mode decision from the sub-band energy. To prevent the over segmentation of a macro-block, a secondary parameter called PSNR threshold is used in the mode decision procedure. Various experimental results show that the proposed algorithm can effectively select a macro-block mode decision under a given PSNR value. Furthermore, the decision time of the proposed algorithm is faster 1 to 4 times than that of the traditional multi-block selection method with the similar macro-block mode decision results.

The rest of this paper is organized as fallows. In Section II the traditional algorithm of macro-block mode estimation is given. Section III describes the proposed macro-block mode selection algorithm and Section IV shows the experimental results with comparison with the traditional macro-block mode algorithm. Finally, Selection V concludes this paper.

II. The traditional Algorithm of Macro-block Selection

The macro-block ME/MC is adopted in H.264 to reduce the residual between the original video frame and the predicted frame and that can increase the accuracy of prediction. The macro-block mainly is used in the edge of moving object in the frame and its default size is 16×16. However, in some cases, it is not enough to describe the moving of the macro-block using only one motion vector. Therefore, it is more fit for MC/ME based on smaller block to get the better block matching. In the ME procedure, the cost function or the mean absolute difference (MAD) is used to determine the similarity of the current and reference macro-blocks. The definition of MAD is given by

$$MAD = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left| C_{ij} - R_{ij} \right| \tag{1}$$

where C_{ij} and R_{ij} are the pixels being compared in current macro-block and reference marc-block, N is the side of macro-block. In addition, the peak-signal-to-noise-ratio (PSNR) of 8 bits gray image is given in Eq. 2 where MSE is abbreviation of mean square error. The PSNR can evaluate the quality of the motion compensated image that is created by using motion vectors and macro blocks from reference frame.

$$PSNR = 10 Log_{10} \left[\frac{(255)^2}{MSE} \right] \tag{2}$$

In the traditional algorithm of macro-block selection, it sets a threshold to restrict the PSNR. If the PSNR calculated by a macro-blocks motion estimation using assigned mode is greater than the threshold, one can reserve this mode and perform motion compensation. Otherwise, one needs to further segment the macro-block to reach the PSNR threshold. The flowchart of the traditional macro-block selection algorithm is shown in Figure 2. When a macro-block gets into this flowchart, the algorithm will carry out motion estimation and calculate the PSNR. If the PSNR is greater than threshold value, it will then execute motion compensation using this mode. On the other hand, it needs to segment the macro-block to smaller size and perform motion estimation again. The flowchart of macro-block segmentation is shown in Figure 3. Finally, the traditional algorithm will select the best PSNR and decide the segmentation mode. It goes without saying that such a "try and error" algorithm will cost high computational complexity.

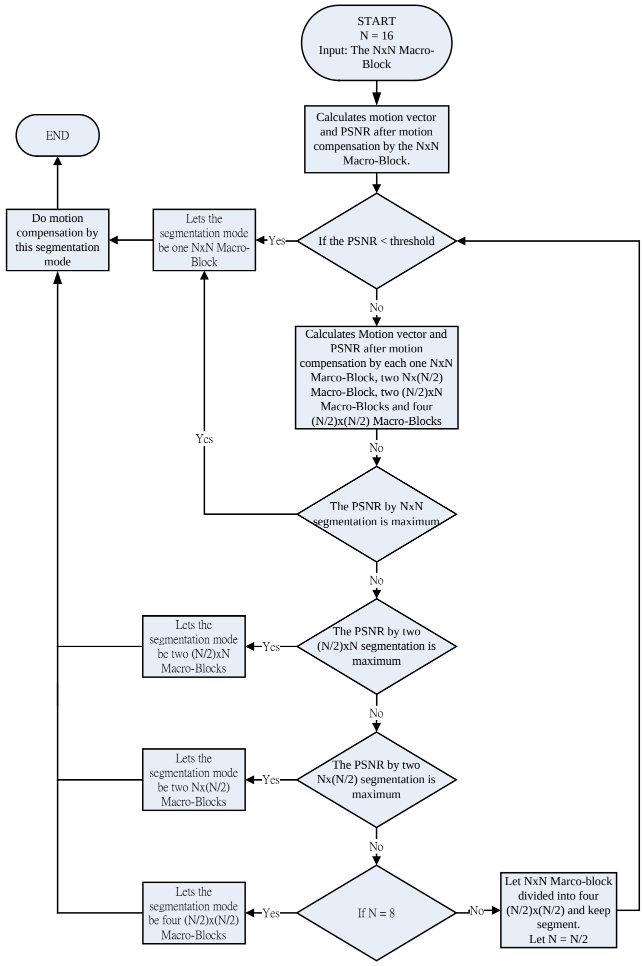


Figure. 2－The flowchart of the traditional algorithm of macro-block selection.

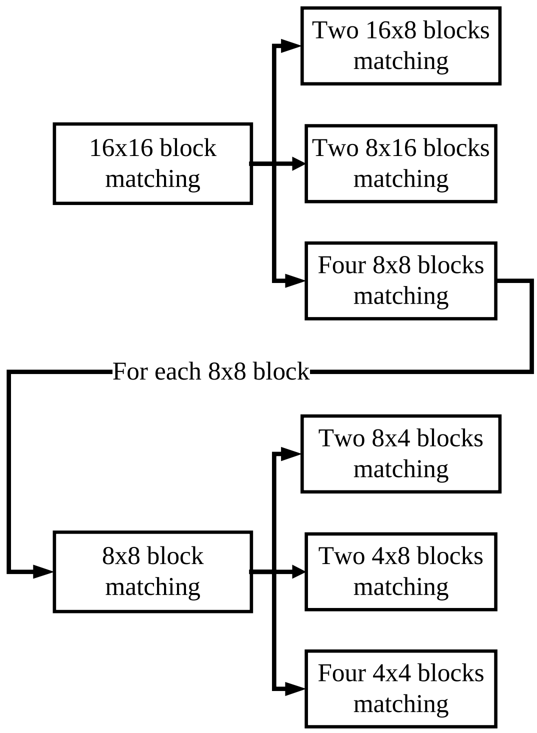


Figure. 3－The flowchart of macro-block segmentation.

III. Macro-Block Mode Selection Algorithm Using Wavelet Transform

To reduce the computational cost of the traditional macro-block selection algorithm, this paper proposes a fast algorithm using wavelet transform for macro-block mode selection. The proposed algorithm uses two-dimensional wavelet transform to pre-decide the segmentation of macro-blocks for reducing the decision time. Using the characteristic of two-dimensional wavelet transform, the proposed algorithm can analyze the sub-band energy of a micro-block. The filter architecture of two-dimensional wavelet transform is shown in Figure 4.

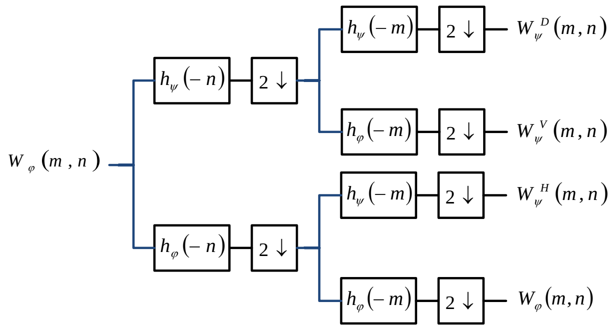


Figure. 4—The filter architecture of two-dimensional wavelet transform.

In figure 4, the index n is the column of image, m is the row of image, W_φ is the sub-band image of low pass approximation, W_Ψ^H is the sub-band image of vertical detail, W_Ψ^V is the sub-band image of horizontal detail, and W_Ψ^D is the sub-band image of diagonal detail. Figure 5 is an example of wavelet transform of a video frame selected from "Susie" sequence. In what follows the wavelet filter function used in this paper is Haar wavelet [8].

Figure. 5—Sub-band images: W_φ, W_Ψ^H, W_Ψ^V, W_Ψ^D of two-dimensional of wavelet transform.

Next, one can calculate the sum of characteristic energy in the sub-band image of vertical detail, horizontal detail and the diagonal detail, respectively. These characteristic wavelet sub-band energy functions are given in Eqs. (3)~(5)

$$EnergyV = \sum_{i=0}^{\frac{N}{2}-1}\sum_{j=0}^{\frac{N}{2}-1}\left(W_\psi^V(i,j)\right)^2 \quad (3)$$

$$EnergyH = \sum_{i=0}^{\frac{N}{2}-1}\sum_{j=0}^{\frac{N}{2}-1}\left(W_\psi^H(i,j)\right)^2 \quad (4)$$

$$EnergyD = \sum_{i=0}^{\frac{N}{2}-1}\sum_{j=0}^{\frac{N}{2}-1}\left(W_\psi^H(i,j)\right)^2 \quad (5)$$

where *EnergyV* is the sum of characteristic energy in the sub-band image of vertical detail, *EnergyH* is the sum of characteristic energy in the sub-band image of horizontal detail and *EnergyD* is the sum of characteristic energy in the sub-band image of diagonal detail. If the *EnergyV* is the maximum of the characteristic energy, more horizontal information will be contained in the image. Similarly, if the *EnergyH* is the maximum, more vertical information will be contained in the image. Finally, if *EnergyD* has the maximum value, more diagonal information will be contained in the image. These sub-band energies, called the primary parameter in this paper, derived from the two-dimensional wavelet transform are applied to predict the macro-blocks selection.

The flowchart of the proposed algorithm is shown in Figure 6. It follows from this flowchart that the proposed algorithm will analyze macro-block characteristic of video frame using wavelet transform. It will select a sub-band whose energy is maximal from *EnergyH*, *EnergyV*, and *EnergyD* and select a corresponding macro-block mode. If the current selection has met the requirement of the second parameter, i.e., a PSNR value defined by user, the proposed algorithm will stop. Otherwise it will keep on the segmentation procedure.

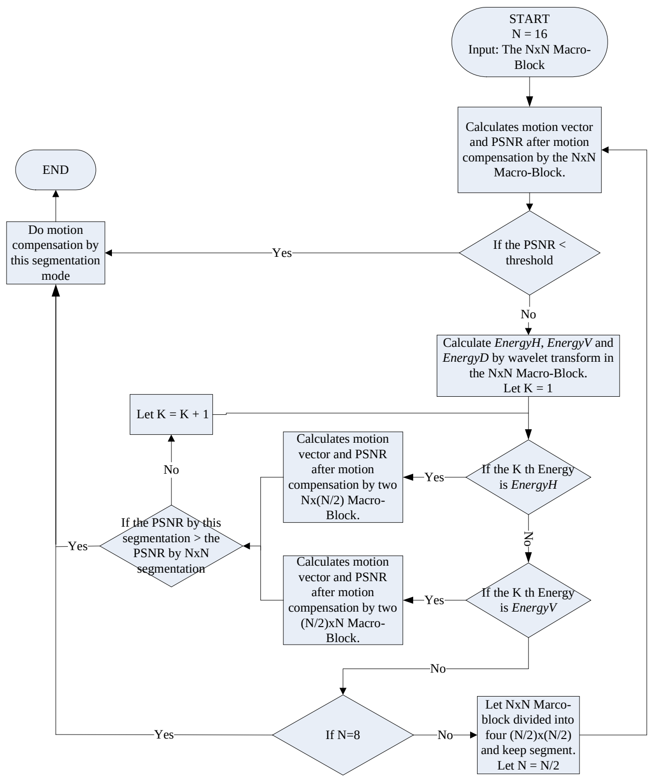
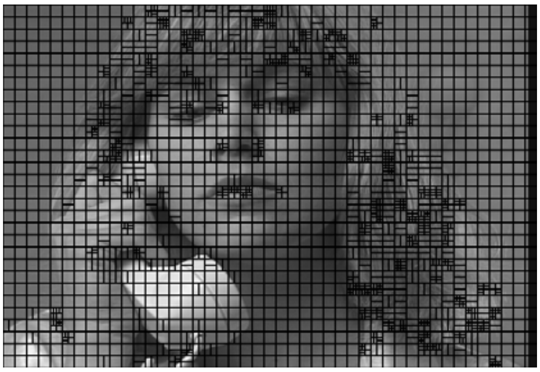
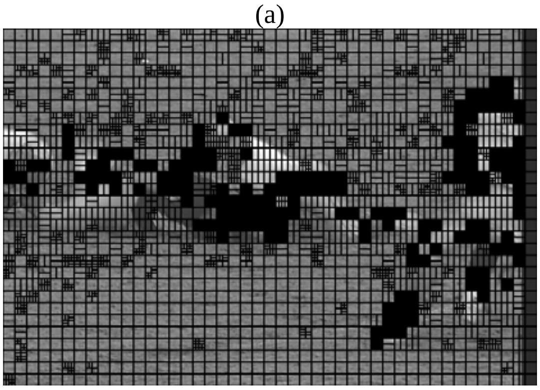


Figure. 6—The flowchart of Proposed Multi-block Selection Algorithm using Wavelet Transform

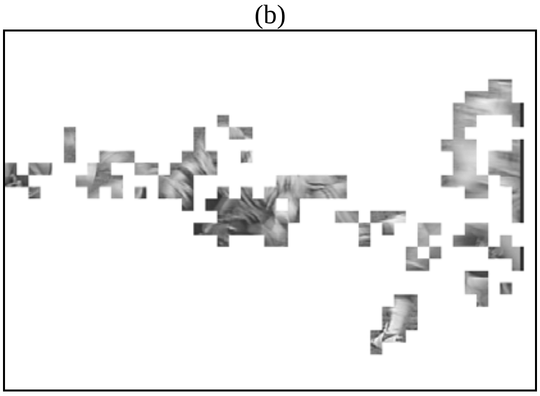
Figure 7 illustrates two example of the proposed algorithm on video frames selected from “Susie” and “football” sequences. In Figure 7(a), one can find that the areas of eyes, hair, and the megaphone are more complex than other regions. Therefore, these areas have more detailed segmentations. The similar results can also be found in Figure 7(b). Meanwhile the black blocks shown in Figure 7(b) mean that they cannot finally meet the given PSNR threshold value via the proposed algorithm and have to be processed with INTRA-MODE. Figure 7(c) shows the contents of these INTRA-MODE blocks.



(a)



(b)



(c)

Figure. 7－TWO EXAMPLE OF THE PROPOSED ALGORITHM ON VIDEO FRAMES SELECTED FROM “SUSIE” AND “FOOTBALL” SEQUENCES.

VI. EXPERIMENTAL RESULT

In this paper, three CIF video sequences, namely “Susie”, “Football”, and “Mall” are selected to demonstrate the results of the proposed algorithm. Each video sequence contains 30 frames and only gray frames are considered in this paper. All of the experimental results are performed on a PC with Pentium-4 2.4GHz CPU and Matlab simulation program. The first video frame is encoded as I frame, and the remaining frames are encoded as P frame. To decrease the variable factor of experiments, both the proposed algorithm and the traditional algorithm use full search method with ±16 search range and half-pixel accuracy in ME.

Figures 8, 9, 10 are the experimental results of the proposed wavelet-based macro-block selection algorithm compared with the traditional algorithm on “Susie”, “Football”, and “Mall”, respectively. The evaluation items include PSNR, execution time, intra mode rate, and mode selection similarity. One can find that the proposed algorithm can achieve similar macro-block mode selection result as well as intra mode rate to the traditional algorithm. Although the PSNR values of the proposed algorithm are somewhat lower than those of the traditional algorithm, it is worth to note that the execution time of the proposed algorithm is faster than that of the traditional algorithm. Table 1 lists the summary of the experimental results. It follows from Table 1 that the proposed algorithm can reduces computational cost by 1 to 4 times.

V. CONCLUSIONS

In this paper, a new wavelet-based macro-block selection algorithm for H.264/AVC video coding system is proposed. The first step of the proposed algorithm is to make use of two-dimensional wavelet transform to estimate the sub-band energy of each macro-block. Then by the use of the primary parameter, i.e., sub-band energy, and the secondary parameter, i.e., PSNR threshold, the proposed algorithm can achieve macro-block mode selection results. It follows from various experimental results that the execution time of the proposed algorithm is faster 1~4 times than that of the traditional multi-block selection method with the similar macro-block mode decision results.

REFERENCES

[1] Advanced Video Coding for Generic Audiovisual Services, ISO/IEC 14496-10:2005(E) ITU-T Rec. H.264(E), Mar. 2005.

[2] J. Ostermann, J. Bormans, P. List, and D. Marpe et al., “Video Coding with H.264/AVC: Tools, Performance, and Complexity,” IEEE Circuits Syst. Mag., vol. 4, no. 1, pp. 7–28, 2004.

[3] T. Wiegand and G. J. Sullivan, “Overview of the H.264/AVC Video Coding Standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 13, pp. 560-576, July 2003.

[4] Hanli Wang, Sam Kwong, Chi-Wah Kok, “An Efficient Mode Decision Algorithm for H.264/AVC Encoding Optimization,” IEEE Transactions on Multimedia, vol 9, no. 4, pp. 882-888, June 2007.

[5] P. Yin, H. C. Tourapis, A. M. Tourapis, and J. Boyce, “Fast mode decision and motion estimation for JVT/H.264,” in Proc. IEEE ICIP’03, Sept. 2003, vol. 3, pp. 853–856.

[6] X. Jing and L. P. Chau, “Fast approach for H.264 inter mode decision,” Electron. Lett., vol. 40, no. 17, pp. 1051–1052, Aug. 2004.

[7] L. Yang, K. Yu, J. Li, and S. Li, “An effective variable block-size early termination algorithm for H.264 video coding,” IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 6, pp. 784–788, Jun. 2005.

[8] C. Sidney Burrus, Ramesh A. Gopinath and Haitao Guo, *Introduction to Wavelets and Wavelet Transforms,* Upper Saddle River, NJ: Prentice-Hall, 1998.

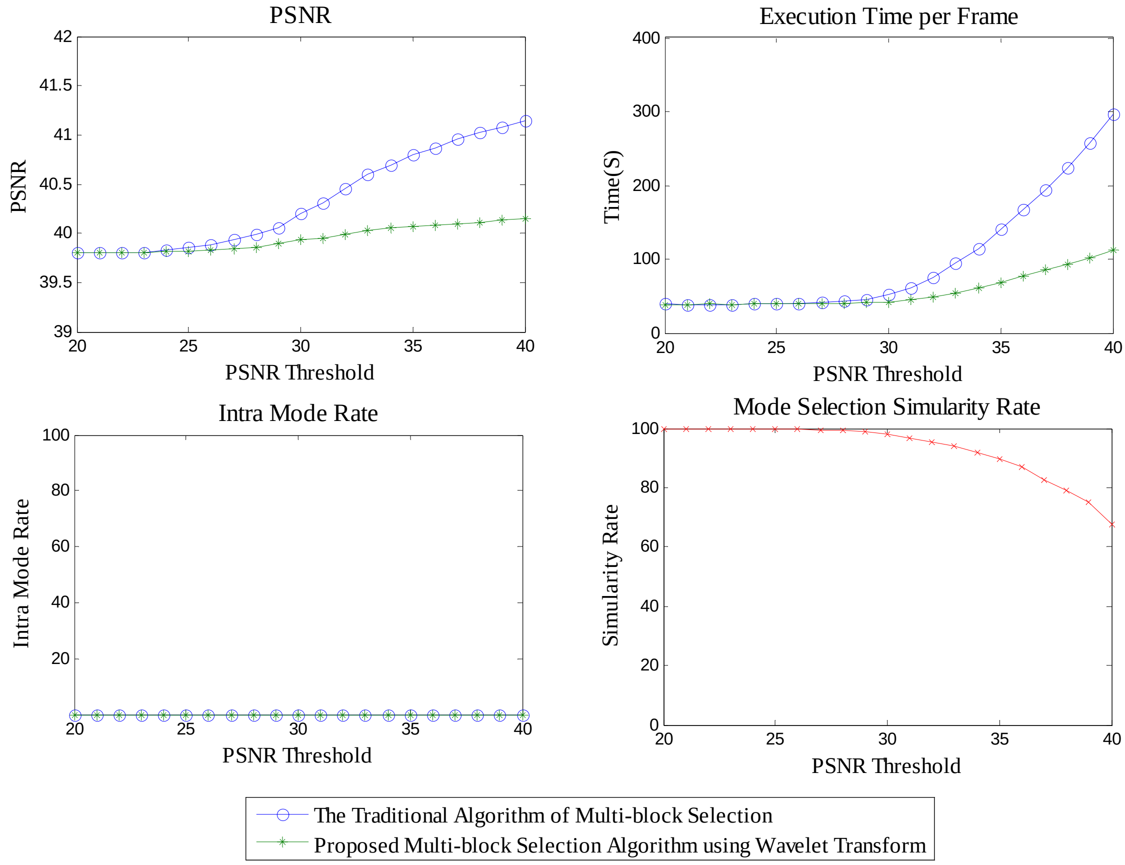


Figure. 8 – The experimental results of the video sequence: “SUSIE”.

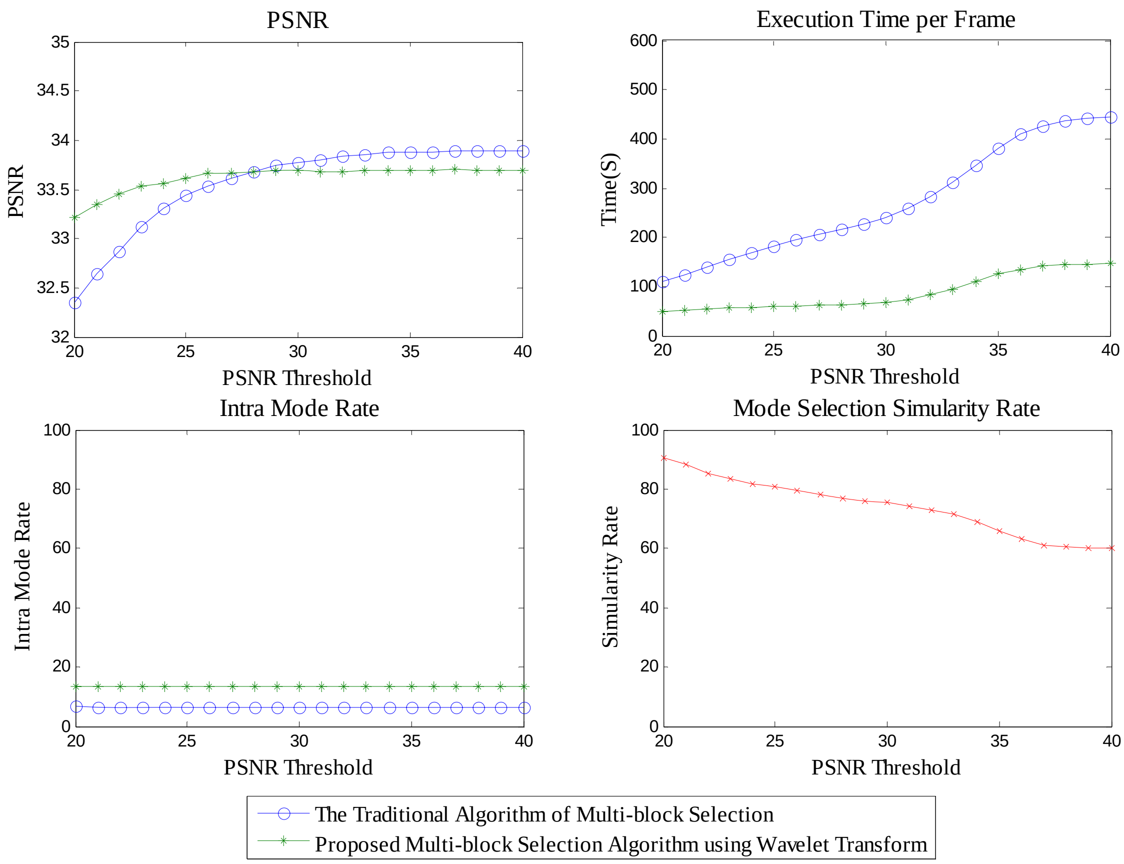


Figure. 9 – The experimental results of the video sequence: “FOOTBALL”.

ISBN: 978-988-98671-8-8

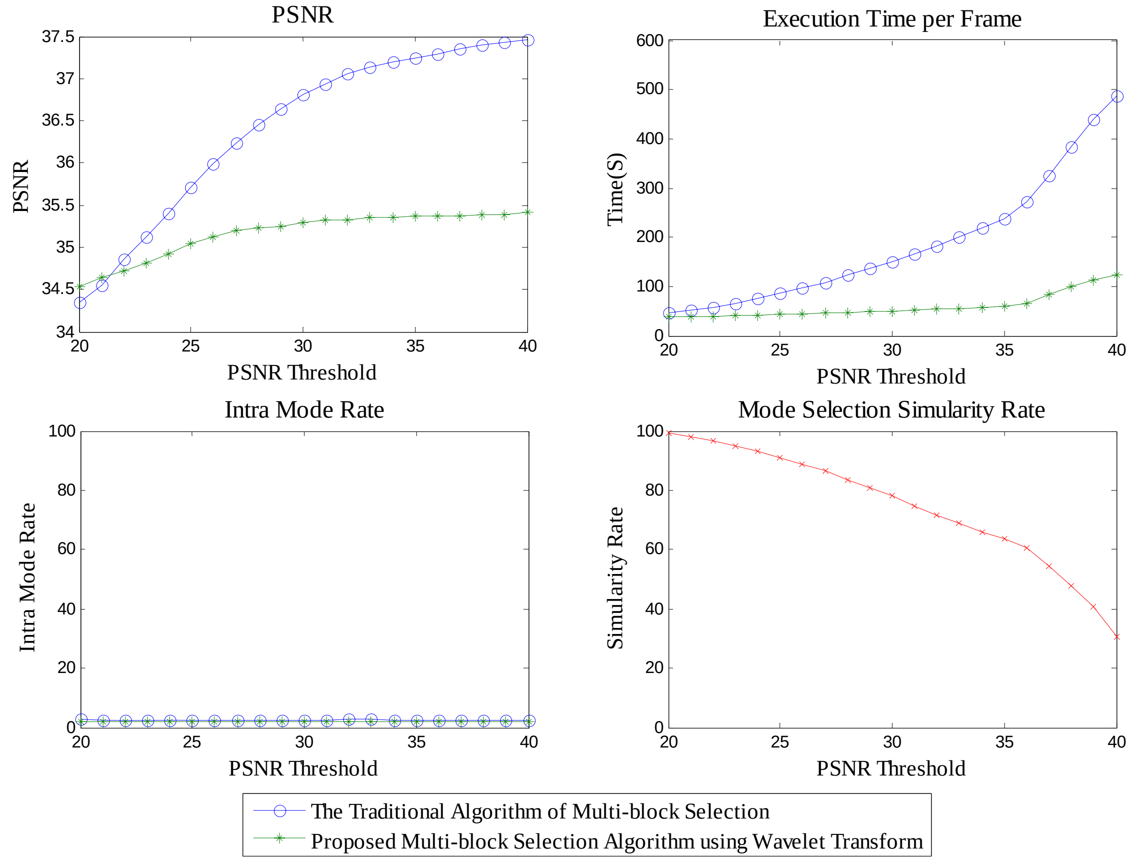


Figure. 10 – The experimental results of the video sequence: “MALL”.

Table. 1 The summary of the experimental results of the proposed and the traditional algorithms.

Video Sequence	Threshold (dB)	The Traditional Algorithm of Multi-block Selection		Proposed Multi-block Selection Algorithm using Wavelet Transform		PSNR Diff.	Speed up Ratio
		PSNR(dB)	Execution Time (Second)	PSNR(dB)	Execution Time (Second)		
SUSIE	20	39.801	38.969	39.801	38.898	0	1.001
	30	40.196	51.445	39.941	42.422	-0.255	1.212
	40	41.147	296.79	40.149	112.04	-0.998	2.648
FOOTBALL	20	32.352	109.03	33.212	49.672	0.860	2.194
	30	33.744	241.48	33.686	68.216	-0.058	3.539
	40	33.895	444.64	33.69	146.23	-0.205	3.040
MALL	20	34.349	47.258	34.542	37.956	0.193	1.245
	30	36.814	151.27	35.294	50.401	-1.520	3.001
	40	37.457	487.56	35.41	123.76	-2.047	3.939