

FAST ALGORITHM FOR H.264/AVC INTRA PREDICTION BASED ON DISCRETE WAVELET TRANSFORM

Damián Ruíz, Oscar Patiño, David Jiménez, José Manuel Menéndez
Grupo de Aplicación de Telecomunicaciones Visuales
Universidad Politécnica de Madrid, Madrid, España, email: drc@gatv.ssr.upm.es

Email: opg@gatv.ssr.upm.es, djb@gatv.ssr.upm.es, jmm@gatv.ssr.upm.es

Abstract

H.264 or MPEG-4 AVC (Advanced Video Coding) is the new world-wide accepted international Standard for video coding, approved by ITU-T and ISO. New Tools have been added to improve the coding efficiency allowing a save up above of 50%, when is compared with previous standards (H.263, MPEG-2 y MPEG-4). From April 2007 there is a new set of profiles known as “all-Intra”. They were born as a sub-set from the “High” profile and have reached a high impact within broadcast industry where the highest quality video formats are demanded.

The high efficiency of “all-Intra” mode in H.264 is due to Rate Distortion Optimization (RDO) technique. RDO chooses for each macroblock (MB) the best partition mode and directional prediction. However, the computational burden becomes extremely high due to huge number of prediction-prediction modes that should be evaluated.

This article shows a new algorithm for fast partition mode algorithm based on Discrete Wavelet Transform (DWT). It allows reducing the number of candidate modes to those which are strictly defined for each partition.

By using the local 2D-DWT over each MB, information of the homogeneity is obtained. It is got from multiresolution analysis of the transformed coefficients in each sub-band. This way it is easier to classify quickly the optimum partition mode avoiding the exhaustive seek made by RDO.

Keywords

Discrete Wavelet Transform, Intra Prediction, H.264/MPEG-4 AVC, Rate Distortion Optimization, Homogeneity analysis.

INTRODUCTION

H.264/MPEG-4 AVC [1] uses the traditional hybrid predictive-transformational coding scheme, but new techniques that represent the estate-of-art in coding are added, as Multi-Picture Inter-Picture Prediction, Variable Block-Size Motion Compensation, Context-Adaptive Binary Arithmetic Coding, Weighted Prediction and Intra Prediction, among others.

RDO used in H.264 is considered the key tool for Intra coding. It is basic for the selection of the best mode of prediction but its calculation demands a very costly computation. It is based on Lagrange cost function that demands the knowledge of both, distortion rate (D) and bit rate (R) generated during each mode of Intra partition coding (I_Mode).

Next equation (1) shows the mathematical expression for Lagrange function cost, where QP is the quantization parameter applied and λ_{MODE} is the Lagrange multiplier:

$$J(I_{MODE}) = D(I_{MODE}/QP, \lambda_{MODE}) + \lambda_{MODE} R(I_{MODE}/QP, \lambda_{MODE}) \quad (1)$$

Specifically, the exhaustive computation of RDO checks all the possible combinations. Thus it means that RDO has to evaluate every MB of the image 736 times when any profile of the “High” family has been chosen. This has caused that many research works trying to reduce the number of iterations needed in RDO meaningfully have been done. These algorithms are called “Fast Intra”.

Most of them are based on directional information obtained from MBs in order to get a close approximation to the directional predictors defined in the standard. Several analyses in spatial domain are proposed to get that information.

Within spatial domain, Pan [2] deals with a fast algorithm based on histogram analysis of the edge map in the image. It is obtained from the gradient vector calculation for every pixel within the block through Sobel operator. Candidate modes are reduced to those close to the direction of the dominant gradient. Fangwen [3] shows an optimization of Pan’s algorithm based on statistical analysis of neighbours MB. This allows calculating the border map only for the most probable modes. Yue [4] computes RDO for the 16 MB sub-blocks which, in turns, allows to estimate a reduced number of candidates in the 8x8 and 16x16 set of partitions. Chih-Hsien [5] defends a fast algorithm based on the symmetry of adjacent angles of directional predictors in order to reduce the number of candidate modes.

Other algorithms get the directional information from MBs through analysis in the frequency domain. By analysing the transformed coefficients (DCT), Takeshi [6] obtains the energetic information of vertical and horizontal edges to determine the dominant gradient for 4x4 sub-blocks. Zhenyu [7] characterizes the predominant direction of the pixel gradient within the MB through a non-normalized Haar Transform.

The set of described algorithms are based on seeking candidate modes for the directional prediction despite of the possible partition mode. The algorithm that is proposed in this article shows an innovative strategy. The objective consists of obtaining the optimum partition for the MB first, then applying RDO scheme. Thus it is only necessary to check the defined modes for the selected partition while reducing drastically the number of evaluations.

INTRA MODE PREDICTIONS IN H.264/AVC

The new visual coding standard H.264/MPEG-4 AVC unlike previous standards (H.261, H.263, MPEG-1, MPEG-2, MPEG-4) adds the novelty of a spatial prediction stage for the Intraframe mode among the macroblocks to code and the adjacent MBs.

Specifically, the standard defines the next partition-prediction combinations for the Luma component:

- I_16x16: single 16x16 blocks with 4 directional predictors.
- I_8x8: four 8x8 blocks with 9 predictor. It is only defined for “High” profiles by low-pass filtering the pixels used in the prediction.
- I_4x4: sixteen 4x4 sub-blocks with the same 9 directional predictors, this time without previous filtering.

Fig. 1 shows the directional predictors for I_8x8, I_4x4 and 16x16 respectively.

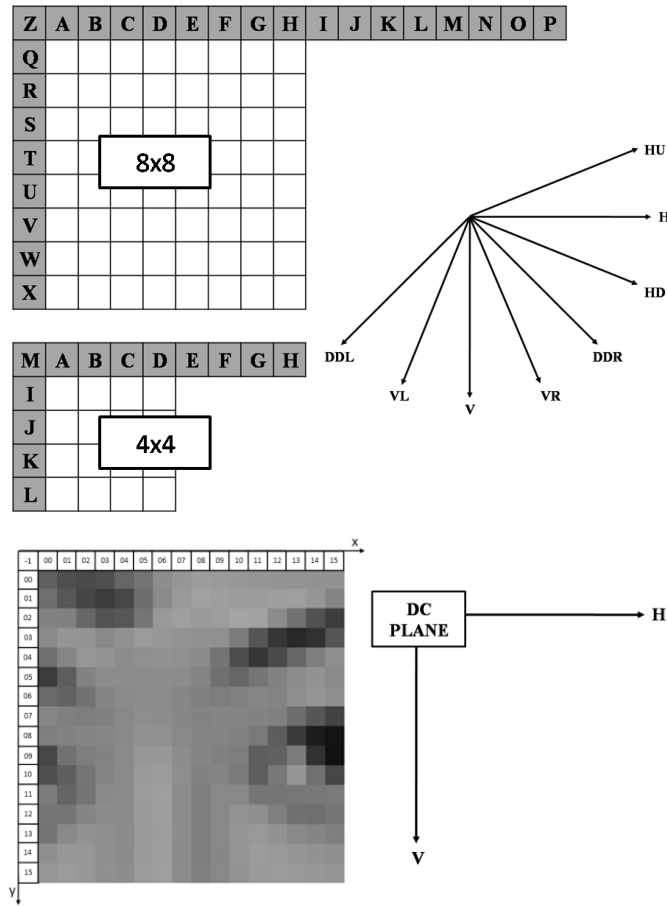


Figure 1. Directional predictors for I_8x8, I_4x4 and 16x16

Table 1 shows the directional predictors aforementioned. Chroma components are never partitioned and always use the same directional predictors as 16x16 Luma MB (I_16x16).

Partition	Mode		Directional Predictor
16x16	0	V	Vertical
	1	H	Horizontal
	2	DC	Mean Value
	3	D	Planar Mode
8x8 & 4x4	3	V	Vertical
	1	H	Horizontal
	2	DC	Mean Value
	3	DDL	Diagonal Down Left
	4	DDR	Diagonal

Partition	Mode		Directional Predictor
			Down Right
	5	VR	Vertical Right
	6	HD	Horizontal Down
	7	VL	Vertical Left
	8	HU	Horizontal Up

Table 1. Prediction Modes for H.264/AVC

OPTIMUM PARTITION BASED ON LOCAL 2D-DWT

A. 2D-DWT

The transformation based on wavelets presents a wide range of applications in the field of images and video sequences, from border detection to pattern recognition, and especially in the image compression field where it is the basis of JPEG 2000 standard [8].

Currently, it has established itself as an alternative to the Discrete Cosine Transformation to decorrelate images, as presenting no block distortion like DCT does. Another advantage is its intrinsic spatial scalability.

The bi-dimensional Discrete Wavelet Transform (2D-DWT) can be implemented as one dimensional low and high filtering, followed by a new low and high filtering in the other dimension.

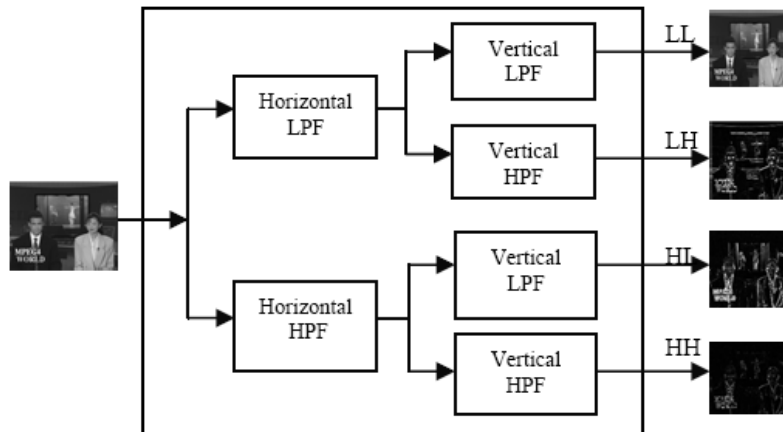


Figure 2. Single decomposition level of 2D-DWT

Every scale or decomposition level λ provides 4 sub-bands μ , containing different directional information of the image. LH gets horizontal information, HL gets vertical information and HH gets diagonal information. Additionally, a continue component LL is obtained from the decomposition. LL sub-band can be iterated applying a new 2D-DWT to obtain a new decomposition level λ_2 .

Classic presentation of 2D-DWT consists of placing, left to right and up to down, the first four sub-bands LL^1 , LH^1 , HL^1 , and HH^1 , and then replacing on each new scale the LL^x sub-band for the new four sub-bands.

B. ALGORITHM

I_16x16 modes provide good performance for low spatial complexity blocks which show a global homogeneity in the MB as a whole. Otherwise, the MBs that are not considered as homogeneous obtain a lower residue if they are coded with smaller partitions, 8x8 or 4x4. The selection is made based on the overall complexity of the MBs.

It is easy to consider that using I_4x4 mode can be faster and provides a better adjustment to the gradient for any MB, so it would minimize the cost function. This statement is not true, as there are other factors that have to be taken in account as signaling or the statistical coding.

The “homogeneity level” of the MB is the threshold used within the algorithm that presents a 3 steps scheme:

- 1) *Three level 2D-DWT*. 2D-DWT is applied using three levels of decomposition by MB. The orthogonal Haar filter of length 2 is used. A hierarchical storage structure has been used in order to analyse the value of LL2 sub-band.
- 2) *16x16 homogeneity analysis*. This analysis is made by evaluating the 2x2 sub-bands of the third scale λ^3 in the transformed domain: LL^3 , LH^3 , HL^3 , and HH^3 . Every one of the 4 transformed coefficients of each sub-band provides directional information from the 8x8 ($2^\lambda \times 2^\lambda$) blocks of the original MB. The same interpretation is applied through HL^3 , LH^3 and LL^3 sub-bands and the corresponding blocks within MB.

If the analysis results in a homogeneous MB, a 16x16 partition (I_16x16) is selected as the optimum one and the algorithm finishes. Otherwise, the third step is executed.

- 3) *8x8 homogeneity analysis*. In order to obtain the local homogeneity for every one of the 4 blocks (B_i) within the MB, the second level of decomposition is evaluated. The λ_2 4x4 sub-bands, LL^2 , LH^2 , HV^2 , and HH^2 are analysed. Thus, every transformed coefficient of this scale provides frequency information of a 4x4 area ($2^\lambda \times 2^\lambda$), that corresponds with the sub-blocks that conform each B_i . Only in the case that the four blocks B_i were labeled as homogeneous, the MB is processed with an optimum partition of I_8x8; otherwise the MB would be processed as I_4x4.

Figure 4 shows the global scheme of the architecture of the proposed algorithm (detailed for 3 levels).

One key element of the algorithm is the metric that performs the homogeneity evaluation. For the first two steps dealing with 16x16 and 8x8 MBs, the directional homogeneity index Γ has been applied to every sub-band μ belonging to decomposition levels 2 and 3.

Γ index is calculated as the quotient between the energy of each transformed coefficient $C_{\mu\lambda}(k,l)$ and the variance of that sub-band $\sigma_{\mu\lambda}^2$ as is shown in Eq. (2):

$$\Gamma(\mu,\lambda) = C_{\mu\lambda}(k,l) / \sigma_{\mu\lambda}^2 \quad (2)$$

The mentioned index is calculated for the 4 sub-bands μ . Then it is compared with the threshold Th_{LL} , Th_{LH} , Th_{HL} and Th_{HH} . Each threshold is calculated empirically for the scales 2 and 3.

The global homogeneity condition implies that the index $\Gamma(\mu,\lambda)$ does not exceed the corresponding threshold for the sub-bands.

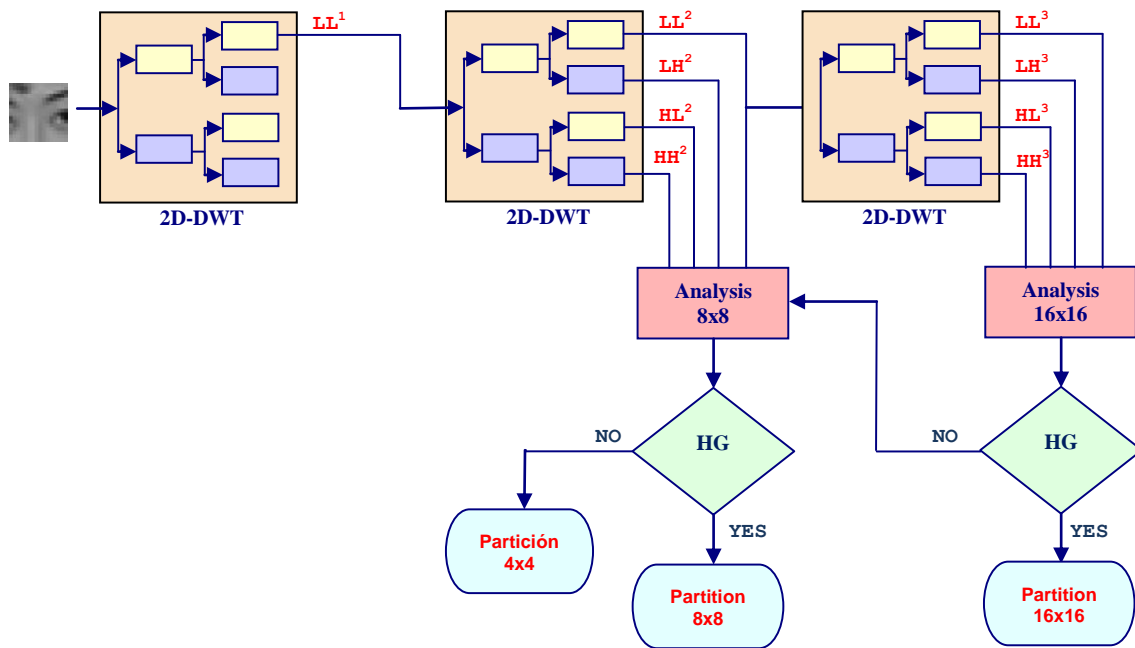


FIGURE 3. FAST INTRA ALGORITHM ARCHITECTURE FOR H.264/AVC

RESULTS

To evaluate the proposed algorithm efficiency three different 720p high definition video sequences with a wide range of spatial complexity have been tested: “*EBU Test Pattern 7*”, “*Shield*” and “*New Mobile-Calendar*” (Figure 4). It can be seen that sequences are highly detailed and have several areas with meaningful high frequency content. Otherwise, they also present homogeneous (low frequency) areas.

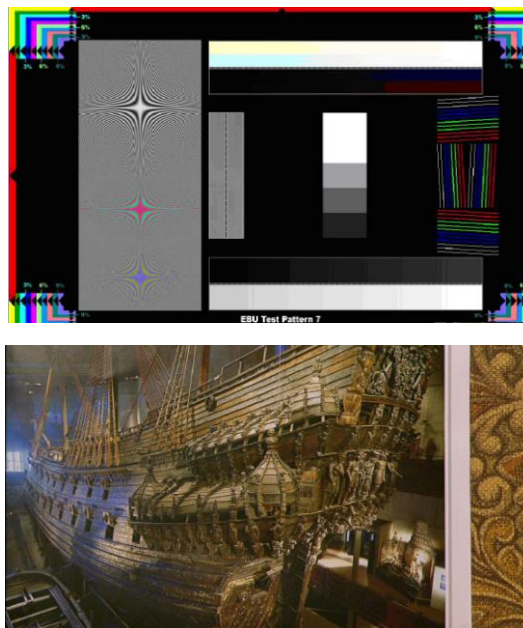




Figure 3. Sequences "EBU Test Pattern 7", "Shields" & "New Mobile Calendar"

The first of them, "EBU Test Pattern 7", contains patterns with very different resolution and spatial orientation, which provides a good environment to test the algorithm performance on the MB at those sites. The partitions are obtained by applying the proposed algorithm (shown in Figs. 5, 6 & 7), where can be observed how the partition map is accurately adjusted to the image topology setting I_4x4 partitions where complex textures are present, for example on the text displayed at the bottom of the image, the percentage details at the corners and high frequencies data in the left window. In the homogeneous regions can be seen that the algorithm fits perfectly, choosing 16x16 partition even on areas with high detail level but low energy, as can be observed at the bottom of the high frequencies window.

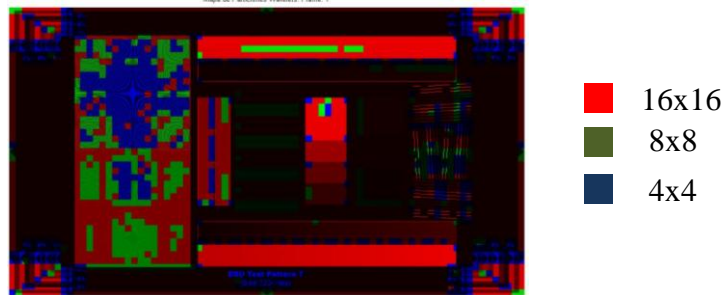


Figure 4. 2D-DWT partition map for "EBU Test Pattern 7"

On the same way, it can be seen that the 8x8 sized partitions are related to intermediate critical areas where I_8x8 is a good balanced decision between I_16x16 and I_4x4.

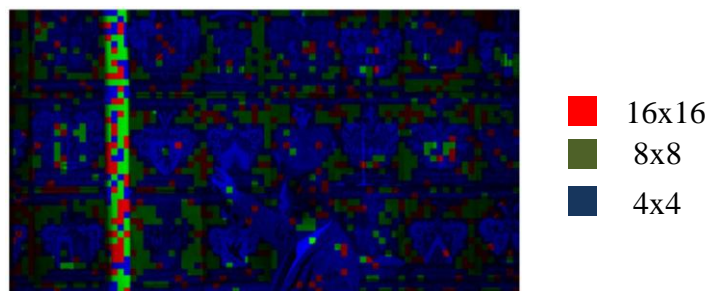


Figure 5. 2D-DWT partition map for "Shields"

Qualitatively, the partition map shows that I_4x4 is extensively used and fits with the high complex areas. Other less complex areas shows I_8x8 or I_16x16 partitions. The 8x8 and 16x16 partitions are distributed by certain randomness in small regions where it can be considered that MB's area contains a global homogeneity at MB and Block levels. The good behavior of the algorithm is ratified both by natural and synthetic sequences.

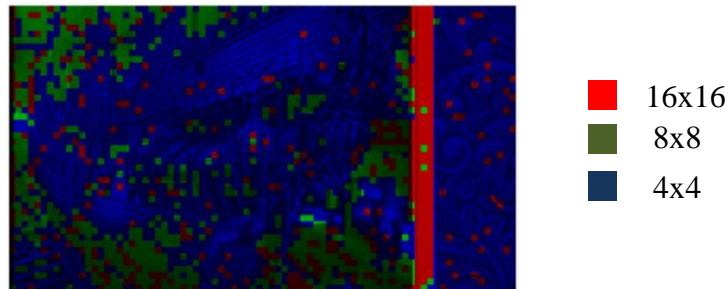


Figure 6. 2D-DWT partition map for “New Mobile Calendar”

DISCUSSION

This article aims to propose a fast method to get the partitions of MBs for Intraframe mode predictors in H.264/AVC. The described algorithm is based on the application of 2D-DWT locally in the MBs that allows obtaining the optimum partition a priori. By getting a fast estimation of the partition mode the algorithm avoids having to evaluate all the modes as needed in RDO.

The experimental results that have been obtained get well correlated results both for synthetic and natural video sequences. The intrinsic features and complexity of the image is extremely related with the content. However, the homogeneity index has provided successful results for all the sequences used as test.

REFERENCES

- [1] ITU-T Rec. H.264 and Final Draft International Standard of Joint Video Specification (ISO/IEC 14496-10 AVC) Joint Video Team. 2003, Doc. JVT-G050.
- [2] F. Pan, X. Lin, S. Rahardja, et al, “Fast Intra Mode Decision Algorithm for H.264/AVC Video Coding”, 2004 ICIP.
- [3] F. Fu; X. Lin; L. Xu, “Fast intra prediction algorithm in H.264-AVC”, 7th ICIP, 2004, Vol. 2, Pages:1191-94
- [4] Yue Yu, Wang, Limin, “A fast intra mode selection method for H.264 high profile”, ICASSP 2008. Page(s):681 – 684.
- [5] C.Hsia; J.Chiang; et al; “Fast Intra Prediction Mode Decision Algorithm for H.264/AVC Video Coding Standard”, 3rd IHH-MSP 2007, Vol. 2, Pages:535-38.
- [6] T. Tsukuba†, I. Nagayoshi‡, T. Hanamura‡ and H. Tominaga, “H.264 Fast Intra-Prediction mode decision based on Frequency characteristic”.
- [7] Zhenyu Wei; Hongliang Li; King Ngi Ngan;” An Efficient Intra Mode Selection Algorithm For H.264 Based On Fast Edge Classification”, International Symposium on Circuits and Systems, 2007,Pages:3630–33.
- [8] Michael W. Marcellin, Michael J. Gormish, Ali Bilgin, Martin P. Boliek, “An Overview of JPEG2000”, Proc of IEEE Data Compression Conference, pp. 523-541, 2000.