IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 4, No 9, July 2010
ISSN (Online): 1694-0784
ISSN (Print): 1694-0814

A Fast Mode Selection Algorithm Using Texture Analysis for H.264/AVC

Sourabh Rungta[1], Kshitij Verma[2], Anupam Shukla[3]

[1]*Department of Computer Science*
RCET, Bhilai, India

[2]*ABV-IIITM*
Gwalior, India, 474010

[3]*Department of ICT*
ABV-IIITM, Gwalior, India

Abstract

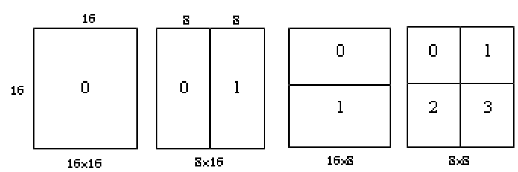
H.264/AVC, the newest international video coding standard achieves higher compression efficiency as compared to all other existing coding standard such as MPEG-4 and H.263. However this efficiency comes with a dramatic increase in computational complexity due to several advanced techniques, such as inter mode and intra mode prediction with variable size motion compensation. It adopts rate distortion optimization (RDO), while maximizing visual quality and minimizing the required bit rate. In this paper, we propose a fast inter mode selection algorithm. The aim is to reduce the number of modes in intra and inter mode prediction. Experimental results show that this algorithm reduces the total encoding time with little loss of bit rate and visual quality.

Keywords*-H.264/AVC, rate distortion optimization (RDO), inter-mode prediction, intra-mode prediction.*

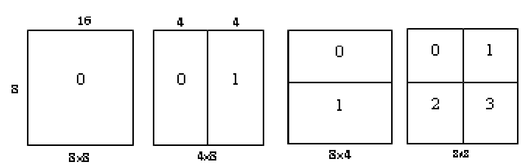
1. Introduction

H.264/AVC is a transform based, block-based, motion compensated video format jointly developed by the International standards organization (ISO) Moving Picture Expert Group (MPEG), and Telecommunication standardization Sector (ITU-T). H.264/AVC is composed of many advanced video coding techniques. The main goal of this standardization effort has been enhanced compression, provision of network-friendly video representation addressing both ‘conversational’ (video conferencing, video telephony) and ‘non-conversational’ (storage, broadcast or streaming) applications. H.264/AVC encoder has 5-10 times computation complexity compared to MPEG-2. This makes it a must to reduce the computational complexity of algorithm in H.264/AVC.

For Inter-frame prediction, the H.264 encoder performs motion estimation of blocks with variable sizes, such as 16x16, 16x8, 8x16, and 8x8, for each 16x16 macroblock[9][4]. In the case of 8x8 block, the macroblock is divided into four 8x8 blocks and each of these 8x8 blocks can be encoded into 8x8,8x4,4x8, or 4x4 blocks, as shown in fig.1.



Macroblock partitions: 16x16, 8x16, 16x8 and 8x8



Submacroblock partitions: 8x8, 4x8, 8x4 and 4x4.

Fig.1.Macroblock and Submacroblock partitions.

H.264/AVC supports various techniques among which intra-mode and inter-mode prediction mostly contribute to the coding efficiency. The Rate Distortion Optimization (RDO) mechanism is used to select the best coding mode among all of the block modes[6][7].However, the computational complexity of the RDO technique is extremely high since the encoder has to encode the target block by searching all possible modes for the best mode. Therefore there is an obvious need for reducing the amount of modes that are evaluated in order to speed up the encoding and hence to reduce the complexity of the encoder.

2. Intra-And Inter Prediction for H.264/AVC

H.264/AVC defines two coding modes which are intra-frame coding and inter-frame coding. Intra-frame coding supports two classes which are Intra 4x4(*I4MB*) and Intra 16x16(*I16MB*). Intra-prediction is based on the observation that adjacent macroblocks tend to have similar properties. When the subsequent frames of the video sequence have comparably large difference among them, Intra coding [1] is selected. For prediction of 4x4 luminance blocks, the 9 directional modes consist of a DC prediction (Mode 2) and 8 directional modes: labeled 0,1,3,4,5,6,7, and 8 as shown in Fig. 2, the block (values of pixels “a” to “p”) is to be predicted using A to Q. The pixels “A” to “Q” from neighboring blocks have already been encoded and may be used for prediction.

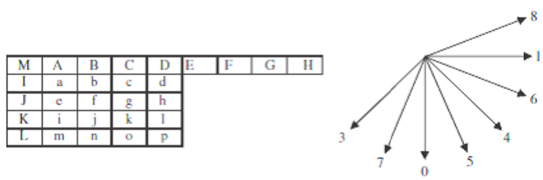


Fig.2.Labelling and direction of intra-prediction modes.

For regions with less spatial details(i.e. flat regions), H.264/AVC supports 16x16 intra-coding; in which one of four prediction modes[2] (DC, vertical, horizontal, and planer) is chosen for the prediction of the entire luminance component of the macroblocks as shown in Fig. 3.

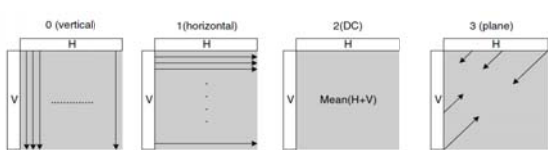


Fig.3.Intra 16x16 prediction modes.

Inter-frame selection supports the following modes: SKIP, 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, and 4x4. The mode decision is made by choosing the mode having minimum RDO[10] cost.

J(s, c, MODE|QP, λmode) = SSD(s, c, MODE|QP) + λ_{MODE} R(s, c, MODE|QP) (1)

Where *QP* is the quantization parameter of the macroblock, *J(s, c, MODE|QP,* λ_{mode}*)* is the mode cost, $\lambda_{mode} = 0.85x2^{(QP-12)/3}$ is the Lagrangian multiplier[3], and *MODE* indicates a mode chosen from the set of macroblock modes: {*SKIP*, 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4, Intra 4x4, Intra 16x16}, *R(.)* represents the rate i.e. the number of bits associated with chosen *MODE*, *SSD* denotes the sum of squared differences between the original block s and its reconstructed signal *c*, computed as,

SSD(s, c, MODE|QP) = [x, y]-Cy[x, y, MODE |QP])^2 [x, y]-Cu[x, y, MODE |QP]) ^{2}+ [x, y]-Cv[x, y, MODE |QP])^2 (2)

Where c_y*(x, y, MODE|QP)* and s_y*(x, y)* represent the reconstructed and original luminance values; c_u, c_v and s_u, s_v indicates the corresponding chrominance values.

3. The Proposed Improvement of Fast Mode Selection Algorithm

The proposed algorithm of mode decision analyzes texture of an image and the motion characteristics for choosing a group of effective predicting modes to reduce the searching modes.

The texture characteristics of the current macroblock, the definition of variable V_{ar} which is complexity level of the texture for the current macroblock can be found as:

$$Var = \sum_{x=0}^{A-1}\sum_{y=0}^{B-1}\left| f(x,y) - \frac{1}{A*B}\sum_{x=0}^{A-1}\sum_{y=0}^{B-1} f(x,y)\right| \qquad (3)$$

In the above equation, *f(x, y)* is a pixel of *(x, y)*, *A* and *B* are used to calculate the block size of variable V_{ar}. A macroblock with well-proportioned texture, V_{ar} is very small. If V_{ar} is less than threshold ψ_{inter} which is a given threshold in the experiment shows that the current macroblock belongs to flat background. For which 16x16 inter-frame coding mode is selected for motion estimation and motion compensation rest of all other coding mode. If $\psi_{inter1} < V_{ar} < \psi_{inter2}$, then it shows that the macroblock belongs to complex area. So the Inter-frame coding mode such as 16x16, 16x8, 8x16, 8x8 is selected. Otherwise if $V_{ar} > \psi_{inter2}$ the current macroblock belongs to motion area, and small inter-frame mode[8] such as 8x8, 8x4, 4x8, 4x4 are selected.

The two parameters of texture direction which are horizontal texture direction D_{hr} and vertical texture direction D_{ver} are also calculated. D_{hr} and D_{ver} are calculated as follows:

$$D_{hr} = \sum_{x=1}^{A-1}\sum_{y=0}^{B-1}\left[f(x,y)-f(x-1,y)\right]^2 \quad (4)$$

$$D_{ver} = \sum_{x=0}^{A-1}\sum_{y=1}^{B-1}\left[f(x,y)-f(x,y-1)\right]^2 \quad (5)$$

The steps of implementation of the proposed algorithm are as follows:

1. First we compute the variable through equation (3).

2. If $V_{ar} < \psi_{inter1}$, it shows that the current macro block belongs to background area. So mode *SKIP* is selected .Then equation $V_{ar} < \psi_{inter2}$ is checked, if it is satisfied, it indicates that the current macroblock belongs to complex background area. If not satisfied then the current macroblock is motion area.

3. D_{hr} and D_{ver} is calculated using eqn. (4) and (5) respectively.

4. Mode selection should chosen according to the estimation of D_{hr} and D_{ver} ,if D_{her}>D_{ver}, then mode 16x16, 16x8 should be selected, otherwise mode 16x16, 8x16 should be the best choice. If D_{hr}>D_{ver} and is in the motion area then mode 8x8, 8x4, 4x4, INTRA 4x4, INTRA 8x8 should be selected, otherwise mode 8x8, 4x8, 4x4, INTRA 4x4, INTRA 8x8 should be selected.

5. Rate Distortion Cost of all modes in the coding mode are calculated, and the mode which has minimal Rate Distortion Cost is selected.

4. Expected Results

H.264/AVC reference software JM 11[7] is applied to validate the efficiency of the proposed fast algorithm. The international standard sequences, which are QCIF video sequences i.e. Foreman, Salesman, Container and Coastguard, are used for simulations.

The test condition [1] is shown in Table I. For simulation we used recommended sequences with various quantization parameters, i.e. QP = 28, QP = 32, QP = 36, QP = 40.

The Δ Bit, Δ Time and Δ PSNR were used for the performance evaluation.

T_{ref} is the coding times used by JM11 [7] encoder and $T_{proposed}$ be the time taken by the proposed algorithm. The ΔTime% is defined as:

$$\Delta Time\% = \frac{Timeproposed - Timeref}{Timeref} \times 100$$

Also the ΔBit value is defined as:

$$\Delta Bit\% = \frac{Bitproposed - Bitref}{Bitref} \times 100$$

$$\Delta PSNR = PSNR_{proposed} - PSNR_{ref}$$

Table I. Simulation Parameter

GOP	IPPP
Codec	JM11
MvSearch Range	16
QP	28,32,36,40
Transform	Hadamard Transform
ProfileIDC	66,30
Coding Method	CAVLC
RDO	Enabled
Size	QCIF

The experimental results are shown in Table II. From Table we can see that the proposed algorithm reduces the encoding time by 81.91% for slow motion videos and 64.2% for fast motion videos.

Table II. Simulation Results for IPPP Type Sequences

Video Sequences	***ΔTime %***	***ΔPSNR(dB)***	***ΔRate %***
Container_qcif.yuv	-82.70	-0.02	0.32
Salesman_qcif.yuv	-79.19	-0.01	0.83
Foreman_qcif.yuv	-66.45	-0.10	1.30
Coastguard_qcif.yuv	-59.63	-0.08	1.34

5. Conclusion

The algorithm of mode decision is key to image compression. The full search algorithm is too complex. A fast mode selection algorithm for H.264/AVC is proposed, which is based on image texture analysis with a simple algorithm to decrease the computational complexity. The experiment results show that the fast mode selection algorithm can increase the calculating speed of H.264/AVC coding greatly, which is helpful for real time application of H.264/AVC coding standard.

References

[1]. Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra, Overview of the H.264/AVC Video Coding Standard.

[2]. Iain E.G. Richardson. *H.264 and MPEG-4 Video Compression*, Wiley 2004.

[3]. Jeyun Lee and Byeungwoo Jeon. Fast Mode Decision for H.264 with Variable Motion Block Sizes. Springer -Verlag, LNCS 2869, pp 723-730, 2003.

[4]. Song Hak Ri and Joern Ostermann, Fast mode decision for H.264/AVC using mode prediction. Springer Berlin / Heidelberg, Volume 4179, pp. 354-363,2006.

[5]. Gary Sullivan, Recommended Simulation Common Conditions for H.26L Coding Efficiency Experiments on Low Resolution Progressive Scan Source Material, VCEG-N81, 14th meeting: Santa Barbara, USA. 2001.

[6]. Zhi Zhou and Ming-Ting Sun. Fast Macroblock Inter Mode Decision and Motion Estimation for H.264/MPEG-4 AVC. Proceedings of IEEE International Conference on Image Processing, ICIP 2004, Singapore, pages 789-792.

[7]. JM 11: Available at http://iphome.hhi.de/suehring/tml/download/

[8]. C. Kim, H. Shih, C. Kuo, Feature-based intra prediction mode decision for H.264, in: Proceedings of ICIP 2004, Singapore, pp. 769–772, 2004.

[9]. B. Meng, O. Au, C. Wang, H. Lam, Efficient intra-prediction mode selection for 4x4 blocks in H.264, in: Proceedings of ICME 2003,Baltimore, USA, pp. III521–III524, 2003.

[10]. Y. Cheng, Z.Y. Wang, K. Dai, J.J. Guo, Analysis of inter frame coding without intra modes in H.264/AVC, in: Proceedings of the Seventh Eurographics Symposium on Multimedia, Nanjing, China, pp. 77–86, 2004.