ISSN 2229-5518

Review of Mode Decision Algorithms for Video Compression

Umesh W. Kaware, Dr. Sanjay M. Gulhane, Ather Ravish Khan

Abstract— The last decade has seen a quiet revolution in digital video technology. Digital video take a large amount of storage or transmission capacity and so video compression is essential. This survey paper discuss various mode decision algorithms used for intra prediction in video compression. In intra prediction there exists high similarity among neighboring blocks in a video frame. Consequently, a block can be predicted from its neighboring pixels of already coded and reconstructed blocks. It exploits the spatial correlation between the adjacent blocks to reduce spatial redundancies within a picture. Compression performance of H.264/AVC is greatly due to the new inter and intra prediction technique. Rate distortion optimization (RDO) technique is adopted by H.264/AVC to select the best intra prediction modes. It achieves remarkable improvement in compression performance, but the computational complexity of coding increases extensively. Therefore, the development of high performance mode decision algorithm for video compression is one of the most challenging theme.

Index Terms—Bitrate, Intra prediction, Mode decision,PSNR, RDO, Redundancy and Video compression

1 Introduction

Video compression or video encoding is the process of reducing the amount of data required to represent a digital video signal, prior to transmission or storage [1]. Digital video data tends to take up a large amount of storage or transmission capacity and so video encoding is essential for any application in which storage capacity or transmission bandwidth is constrained. Almost all consumer applications for digital video fall into this category, for example: Digital TV broadcasting, Internet video streaming, Mobile video streaming, DVD video, Video calling. Each of these applications includes an encoder, which compresses or encodes an input video signal into a coded bitstream, and a decoder, which decompresses or decodes the coded bitstream to produce an output video signal. Data compression is achieved by removing redundanc information gives only a moderate amount of compression. Lossy compression is necessary to achieve higher compression. Most video compression methods exploit both temporal and spatial redundancy to achieve compression. In the temporal domain, there is usually a high correlation between frames of video that were captured at around the same time. Temporally adjacent frames i.e. successive frames in time order, are often highly correlated, especially if the temporal sampling rate or frame rate is high. In the spatial domain, there is usually a high correlation between pixels that are close to each other, i.e. the values of neighboring samples are often very similar [1].

In order to apply the digital video compression technology effectively, several international standards for video compression have been developed. ITU-T and ISO/IEC have pioneered to contribute in the developments for video compression techniques. In 1990, ITU-T introduced the video compression standard H.261 [2]. This standard was designed for videoconferencing and videotelephony applications over the integrated services digital network on which data rates was specified as multiples of 64 Kbps. Meanwhile, the researchers in ISO/IEC have introduced the first version of the video compression format, MPEG1 in 1992 [3] for multimedia CD-ROM applications with target bit-rate of about 1.5 Mbps. In H.261, only the previous video frame is used as the reference frame for the motion compensated prediction while MPEG-1 allows the future frame to be used as the reference frame for the motion compensated prediction, which can provide better prediction. In 1994, both ITU-T and ISO/IEC jointly developed MPEG-2 [4]. MPEG-2 is primarily targeted at coding high-quality video at 4 –15 Mb/s for Video On Demand, Digital Broadcast Television and Digital Storage Media such as Digital Versatile Disc. It is also used for coding High Definition TV, Cable/Satellite Digital TV, video services over various networks and other high-quality digital video applications. It has more sophisticated motion estimation methods to improve estimation accuracy, different DCT modes, scanning methods, various scalability modes and various profiles and levels, each combination targeted for a different application. With the advancement in video coding, the ITU-T released H.263 standard in 1995 for very low bit-rate applications such as videoconferencing, video e-mailing and video telephony [5]. It provides better picture quality at low bit rates with little additional complexity. Two more versions of H.263 are released as H.263+ in 1998 [6] and H.263++ in 2000 [7]-[9]. H.263+ is an extension of H.263, providing 12 new negotiable modes and additional features. Meanwhile in 1998, ISO/IEC developed MPEG-4 standard [10]. Initially, MPEG-4 was aimed primarily at low bit-rate video communications; however, its scope as a multimedia coding standard was later expanded. It is efficient across a variety of bit-rates ranging from a few kilobits per second to tens of megabits per second. It provides improved coding efficiency over MPEG-2, ability to encode mixed media

- *Umesh W. Kaware is currently working on Mode Decision Algorithms for Video Compression and completed master degree program in digital electronics from S.G.B. Amravati University, India, PH-0919763717414. E-mail: umeshkaw@rediffmail.com*
- *Dr. S.M..Gulhane is doctorate from IIT,Kharagpur, India, PH-0919881832100. E-mail:smgulhane67@rediffmail.com*

ISSN 2229-5518

data (video, audio, speech), error resilience to enable robust transmission, ability to interact with the audio-visual scene generated at the receiver.

Upon these developments the ITU-T and ISO established a Joint Video Team to develop a new video compression standard. In 2003, they proposed the H.264 standard , which has also been incorporated into MPEG-4 under the name of Advanced Video Coding (AVC) [11]. The main goals was to double the coding efficiency in comparison to any other existing video coding standards and provision of a "network-friendly" video representation addressing "conversational" (video telephony) and "nonconversational" (storage, broadcast, or streaming) applications [34]. MPEG-4 and H.264 both are concerned with compression of visual data but MPEG-4 emphasizes flexibility whilst H.264's emphasis is on efficiency and reliability. H.264 can deliver better image quality at the same compressed bitrate, or a lower compressed bitrate for the same image quality. Due to its improved compression quality, H.264/AVC is been adopted in many video coding applications. It supports a wide range of prediction options viz intra prediction, inter prediction, multiple prediction block sizes, multiple reference frames and special modes such as Direct and Weighted prediction [15].

Prediction exploits the spatial or the temporal redundancy of a video sequence so that only the difference between actual and predict instead of the whole image data need to be encoded. There are two types of prediction: intra prediction for I-type frame and inter prediction for P-type (Predictive) and B-type (Bidirectional Predictive) frame. In Intra Prediction there exists high similarity among neighboring blocks in a video frame. Consequently, a block can be predicted from its neighboring pixels of already coded and reconstructed blocks. It exploits the spatial correlation between the adjacent blocks to reduce spatial redundancies within a picture whereas the goal of inter prediction is to utilize the temporal redundancy to reduce the data that need to be encoded [35].

Compression performance of H.264/AVC is greatly due to the use of inter and intra prediction techniques, which allows to use variable prediction block sizes and multiple prediction modes. Rate distortion optimization (RDO) technique, which exhaustively examines all intra and inter-prediction modes, is adopted by H.264/AVC to select the best intra and inter prediction modes. It achieves remarkable improvement in compression performance, but the computational complexity of coding increases extensively [24]. The various researchers addressed the issue of reducing the computational complexity and proposed various mode decision algorithms for intra prediction. Improvement in coding complexity is being achieved at the cost of degradation in PSNR and increase in bitrate. Therefore, the development of high performance mode decision algorithm for video compression with high trade-off is one of the most challenging theme.

2 Literature Review

In 2003, a Joint Video Team proposed the H.264 standard [11]. It has shown better coding performance than the existing video coding standards. It supports a wide range of prediction options for intra prediction and inter prediction. The RDO technique for maximizing coding quality and minimizing the bit rate is employed in H.264 but it exhibits extremely high computational complexity. The inter and intra predictions are the dominant components in computation [15]. Various authors have made efforts to reduce down the computational complexity by extracting and exploring the information related to edge, direction, mode features and block features.

The algorithms proposed in references [12]-[25] utilize the edge or directional information. Authors, F. Pan et. al. in [12] proposed a fast mode decision algorithm for H.264/AVC intra prediction based on local edge information. Prior to intra-prediction, an edge map is created and a local edge direction histogram is then established for each sub-block. Based on the distribution of the edge direction histogram, only a small part of intra prediction modes are chosen for RDO calculation. This mode decision scheme increases the speed of intracoding at the cost of loss in peak signal-to-noise ratio (PSNR) and increase in bitrate. In [13], M. de-F. Lopez et. al. describe improved version of Pan's algorithm for reducing computational load and increasing RD performance. The authors suggest reducing the number of gradients computed to obtain the local edge information by changing the gradient operator and evaluating neighboring modes. The proposed algorithm improves Pan's algorithm in computational terms and substantially improves its performance.

Reference [14] presented a fast mode decision algorithm for intra prediction based on integer transform and adaptive threshold. Before the intra prediction, integer transform operations on the original image are executed to find the directions of local textures. The presented algorithm can accelerate the encoding speed with PSNR loss and bit rate increment. J.C. Wang et. al. in [15] proposed a regular spatial domain filtering technique to compute the dominant edge strength (DES) to reduce the possible predictive modes. The proposed algorithm reduces 40% computation with slight PSNR degradation. In [16] authors proposed two fast, efficient direction detection algorithms by computing subblock and pixel direction differences. Both methods effectively estimate the edge direction inside the block to narrow down the predictive modes to reduce the RDO computation. The proposed methods can reduce the encoding time by about 60% at the cost of loss of coding performance. A low complexity fast mode decision algorithm for intra prediction that uses discrete cross differences to reduce the unlikely candidate modes in the RDO calculation is presented in [17]. By using horizontal and vertical differences in different locations, the directions of edges can be precisely detected. This algorithm reduces the encoding time by about 56%, with loss of video quality. Reference [18] explore the hierarchy of H.264 mode decision process and adopt an approach that is in synchrony with the mode decision hierarchy. The authors proposed a variance-based algorithm for block size decision, an improved filter-based algorithm for prediction mode decision using contextual information and a selection algorithm for intra block decision. This algorithm provides significant improvement of computational efficiency with PSNR drop and the bit rate increase by 55%.

In [19], authors introduced a low complexity and fast approach for intra mode decision, based on reducing the number of candidate modes for further RDO calculation and decreas-

ISSN 2229-5518

ing the computational complicacy. The proposed algorithm achieved more than 70% time saving with 0.2dB PSNR decrement and 3.0% bit rate increment. An efficient block type decision algorithm for intra prediction which determines the optimal block type by two steps is proposed in [20]. The first step is based on the fact that the block type of intra prediction is highly dependent on the smoothness of macroblock. The second step is based on the correlation of block type chosen for different chroma modes. The presented algorithm can achieve 72.8% time saving on average with average 0.75% bit rate increase and 0.05dB PSNR degradation. In [21], C. Chen et. al. proposed a hybrid mode decision method for intra prediction based on a more precise direction representation, which exploits the gradation along the prediction direction as well as the differential between the original pixels and their predicted ones and achieved about 0.05 dB PSNR gain and 1.1% bit rate decrease with increase in computational complexity. In reference [22], the authors presented an enhanced rate-distortion cost function which combines the sum of absolute integer-transformed differences (SAITD) and a rate predictor for H.264/AVC intra-4 x4 mode decision. To reduce the computation of the SAITD, the authors develop a fast computation algorithm which successfully uses the property of linear transform and the fixed spatial relationship of predicted pixels in intra-modes. The enhanced cost function achieves better coding performance than the cost function suggested in H.264 reference software in low bitrate applications. Reference [23] suggested an enhanced SATD-based intra mode decision using both DCT and SATD coefficients. The computation time saved is approximately 21% from I4MB/I16MB mode selection and 35% from I4MB/I16MB mode prediction with average 0.081 dB PSNR loss and 0.19 % bit-rate increment.

In reference [24], a simple efficient mode decision algorithm for intra prediction is developed. Firstly, by extracting the texture direction information of the encoding macroblocks and intra prediction modes, the author proposes a new more-refined intra-mode filter. Secondly, this algorithm combines the texture direction difference (TDD) with SATD to choose the intra chroma candidate modes. Thirdly, TDD of the vertical and horizontal direction calculated in the prediction block type decision are utilized to choose the candidate prediction modes. The proposed algorithm cut down about 76.79 % total intra-frame coding time at the expense of about 0.08 dB PSNR degradation and 2.07 % bit rate increase. D. Quan and Y. S. Ho in [25] have categorized intra prediction mode decision based on mode features, block features and edge or directional information. The authors proposed three algorithm one for each type, an unconditional DC mode decision algorithm using mode features, an algorithm using a condition of block boundary to select the DC mode as a representative algorithm using block features and an algorithm using a ratio of variance along the horizontal direction to variance along the vertical direction as a representative algorithm using edge or directional information. The proposed algorithms saved coding time but at the cost of increase in bitrate and increase in PSNR value.

The algorithms proposed in references [26]-[28] utilize the information on block features. Authors in [26] proposed fast feature-based intra/inter coding mode selection scheme. First, three features are extracted from a macroblock to form a feature space. Then, the feature space is partitioned into three regions, where the risk is calculated using the rate-distortion (RD) performance loss due to wrong mode decision. The authors demonstrated that approximately 20%–32% of the total encoding time can be saved with little degradation in the rate-distortion performance. In reference [27], fast sum of absolute transformed difference based 4x4 intra-mode decision scheme is presented. The scheme reduces the candidate of the prediction modes based on the correlation between neighboring blocks and the sum of absolute transformed difference between the original block and the intra-predicted block. The proposed scheme reduces about 91% of mode decision time and 70% of total encoding time with little degradation of coding performance. A fast block size and mode decision algorithm for intra prediction is presented in [28], demonstrated that maxima of 79% and 77% average time savings with negligible loss in PSNR and bitrate. For fast block size decision, intra16x16 is selectively searched based on homogeneity of inner-8x8 block of macroblock. For fast mode decision, the search of intra16x16 and intra4x4 prediction modes is restricted by the similarity of the reference pixels.

The algorithms in [29]-[33] explore the mode features. In reference [29], the authors suggested two simple algorithms for efficient mode decision by exploiting the input modes and to further improve the conversion performance, two refinement methods are also explained, one is to look at the neighboring of the selected modes and the other to look at horizontal, vertical and DC modes. The proposed algorithms achieve more than 50% complexity reduction while showing a PSNR penalty of 0.1dB and an excess bitrate of 1-6%. H. Zeng et. al. in [30] proposed fast hierarchical intra mode decision (HIMD) to speed up the mode decision process by reducing the number of modes and achieves a reduction of 85.75% computational complexity on average, while incurring 0.164 dB loss in PSNR and 2.336% increment on the total bit rate. Reference [31] presented a feature-based mode decision algorithm based on the fact that a good prediction mode usually has a small residue block. The sum of the absolute transformed coefficients and deviation information of residue block are used to measure the distortion of prediction mode and the smoothness of residue block, respectively. The proposed method reduces about 60% of total encoding time of intra coding at the cost of loss in coding performance. Authors in [32], presents two fast Intra prediction mode decision algorithms for Intra coding. The proposed algorithms select one representative prediction mode or a few representative prediction modes among all intra predictions using the existing DCT and quantization schemes of H.264/AVC, and only the representative modes are used for the RDO process. The proposed algorithms achieve encoding time savings of 53% and 63%, with bitrate increment of 0.88% and 0.93% and loss of PSNR. A new fast mode decision method for H.264/AVC encoder is introduced in [33], based on the idea that when one of the prediction modes achieves good RDO, its orthogonal prediction mode will not perform well. The proposed method has 60% lower encoding time and achieves better PSNR and bitrate when compared with the other methods.

3 Limitations and Scope

Compression performance of H.264/AVC is due to the new inter and intra prediction technique, which allows to use variable prediction block sizes and multiple prediction modes. RDO technique, which exhaustively examines all inter and intra prediction modes, is adopted by H.264/AVC to select the best inter and intra prediction modes. It achieves remarkable improvement in compression performance. However, this technique has extremely high computational load.

At the present stage, low-complexity intra prediction algorithms can be mainly divided into two categories. One is to reduce the computational complexity of each RD cost calculation. The other is to reduce the number of the prediction modes, which can obtain a considerable reduction of computational complexity at the expense of visual quality and compression efficiency. In order to reduce the number of the prediction modes, two kinds of method have been proposed. Firstly, the appropriate smoothness factor of the encoding macroblocks is defined to remove the prediction block types with low possibility. Secondly, the fast prediction mode decision is proposed.

The algorithms proposed in the literature have limitations in coding performance and time consumption. The existing algorithms show the trade-off between the encoding time and the coding performance.

4 Conclusion

In this paper we have discussed various mode decision algorithms for intra prediction in video compression. The improvement in coding complexity is being achieved at the cost of degradation in PSNR and increase in bitrate. The efforts can be made further for improving the mode decision algorithm for video compression that provides better trade-off between various performance parameters.

Acknowledgment

The authors wish to thank Dr. S. M. Gulhane for his support, guidance and motivation.

References

[1] Iain E. Richardson, " The H.264 Advanced Video Compression Standard", *John Wiley and Sons Ltd.*, 2010.

[2] CCITT, "Video Codec for Audiovisual services at p × 64 kbit/s", *CCITT Recommendation H.261,* 1990.

[3] ISO/IEC, "MPEG-1: Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s", *Video,* Geneva, 1993.

[4] ISO/IEC JTC1 and ITU-T, "MPEG-2/H.262: Generic Coding of Moving Pictures and Associated Audio Information - Part 2", *Video,* November 1994.

[5] ITU-T. Video Coding for Low Bit-rate Communication, ITU-T Recommendation H.263, March 1996.

[6] ITU-T. Video Coding for Low Bit-rate Communication, ITU-T Recommendation H.263, version 2 (H.263+), January 1998.

[7] ITU-T Recommendation H.263 - Annex U, "Video Coding for low bit rate communication - Annex U: Enhanced reference picture selection mode," Nov.2000.

[8] ITU-T Recommendation H.263 - Annex V, "Video Coding for low bit rate communication - Annex V: Data-partitioned slice mode," Nov. 2000.

[9] ITU-T Recommendation H.263 - Annex W, "Video Coding for low bit rate communication - Annex W: Additional supplemental enhancement information specification," Nov. 2000.

[10] ISO/IEC JTC1/SC29/WG11. MPEG-4 Video Verification Model: Version 18.0,WG11 Document N3908, Pisa, January 2001.

[11] ITU-T and ISO/IEC JTC1. Advanced Video Coding for Generic Audiovisual Services, ITU-T Recommendation H.264 ISO/IEC 14496-10 AVC, 2003.

[12] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu and S. Wu, "Fast Mode Decision Algorithm for Intraprediction in H.264/AVC Video Coding", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 15, No. 7, pp. 813-822,July 2005.

[13] M. de-F. Lopez, D. O. Quiros, Jose C, P. Alcolado and F. D-de-Maria, "An improved fast mode decision algorithm for intraprediction in H.264/AVC video coding", *Signal Processing: Image Communication* 25, pp.709–716, Elsevier 2010.

[14] R. Su, G. Liu and T. Zhang, "Fast mode decision algorithm for intra prediction in H.264/AVC with integer transform and adaptive threshold", *SIViP* 1, pp.11–27, Springer 2007.

[15] J. C. Wang, J. F. Wang, J. F. Yang and J.T. Chen, "A Fast Mode Decision Algorithm and Its VLSI Design for H.264/AVC Intra-Prediction", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 17, no. 10, pp. 1414-1422,October 2007.

[16] A.C. Tsai, J.F. Wang, J. F. Yang and W. G. Lin, "Effective Subblock-Based and Pixel-Based Fast Direction Detections for H.264 Intra Prediction", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 18, No. 7, pp.975-982, July 2008.

[17] K. Bharanitharan, B.D. Liu, J.F. Yang and W.C. Tsai, "A Low Complexity Detection of Discrete Cross Differences for Fast H.264/AVC Intra Prediction", *IEEE Transactions on Multimedia,* Vol. 10, No. 7, pp.1250-1260,Nov. 2008.

[18] Y.H. Huang, T.S. Ou and H. H. Chen, "Fast Decision of Block Size, Prediction Mode, and Intra Block for H.264 Intra Prediction", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 20, No. 8, pp.1122-1132, Aug. 2010.

[19] X.Q. Su, L. Ji and X. Li, "A fast and low complexity approach for H.264/AVC intra mode decision", *Multimed Tools Appl* 52, pp. 65–76, Springer, 2011.

[20] P. Wang, H. Huang and Z. Tan, "A fast two-step block type decision algorithm for intra prediction in H.264/AVC high profile", *Multimed Tools Appl* 60,pp.139–160,Springer,2012.

[21] C. Chen, J. Chen, K. Ouyang, T. Xia and J. Zhou, "A hybrid fast mode decision method for H.264/AVC intra prediction", *Multimed Tools Appl* 62,pp. 719–731, Springer,2013.

[22] C.H. Tseng, H.M. Wang and J.F. Yang, "Enhanced Intra-4X4 Mode Decision for H.264/AVC Coders", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 16, No. 8, pp. 1027 - 1032,August 2006.

[23] Y. Lin, Y.M. Lee and C.D. Wu, "Efficient Algorithm for H.264/AVC Intra Frame Video Coding", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 20, No. 10, pp.1367-1372,October 2010.

[24] Y. Kuo, J. Yang and J. Chen, "An efficient mode decision algorithm for H.264/AVC intra prediction", *Multimed Tools Appl*, Springer Science+Business Media New York ,Springer 2013.

[25] Do Quan and Y.S. Ho, "Categorization for Fast Intra Prediction Mode Decision in H.264/AVC", *IEEE Transactions on Consumer Electronics,* Vol. 56, No. 2, pp.1049-1056, May 2010.

[26] C. Kim and C.C. J. Kuo, "Feature-Based Intra/Inter Coding Mode Selection for H.264/AVC", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 17, No. 4, pp.441- 453,APRIL 2007.

[27] M. G. Sarwer, L.M. Po and Q.M. J. Wu, "Fast sum of absolute transformed difference based 4X4 intra-mode decision of H.264/AVC video coding stand-

ard", *Signal Processing: Image Communication* 23 , pp. 571– 580,Elsevier 2008.

[28] K. Lim, S. Kim, J. Lee, D. Pak and S. Lee, "Fast Block Size and Mode Decision Algorithm for Intra Prediction in H.264/AVC",*IEEE Transactions on Consumer Electronics*, Vol. 58, No. 2, pp. 654 – 660,May 2012.

[29] S. Moiron and M. Ghanbari, "Reduced Complexity Intra Mode Decision for Resolution Reduction on H.264/AVC Transcoders", *IEEE Transactions on Consumer Electronics,* Vol. 55, No. 2, pp.606-612,May 2009.

[30] H. Zeng, K. K. Ma and C. Cai, "Hierarchical Intra Mode Decision for H.264/AVC",*IEEE Transactions On Circuits and Systems For Video Technology,* Vol. 20, No. 6, pp.907-912, June 2010.

[31] W. Ping, H. Hua and T. Zheng, "Fast feature-based mode decision for 4×4 intra prediction in H.264/AVC",*SCIENCE CHINA (Information Sciences),* Vol.54,No. 11,pp 2386-2399,November 2011.

[32] D.Y. Kim and Y.L. Lee, "A fast intra prediction mode decision using DCT and quantization for H.264/AVC", *Signal Processing: Image Communication* 26 ,pp.455–465,Elsevier,2011.

[33] H. Pejman and F. Zargari, "An Efficient Fast Intra Mode Decision Method Based on Orthogonal Modes Elimination", *IEEE Transactions on Consumer Electronics,* Vol. 58, No. 4, pp. 1345 – 1352, November 2012.

[34] T. Wiegand, G. J. Sullivan, G. Bjøntegaard and A. Luthra, "Overview of the H.264/AVC Video Coding Standard", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 13, No. 7, pp. 560-576, July 2003.

[35] J. W. Chen, C.Y. Kao and Y. L. Lin, "Introduction to H.264 Advanced Video Coding", *IEEE Conference Publications*, 2006.