EFFICIENT MODE DECISION SCHEME FOR SPATIAL SCALABLE VIDEO CODING

[1]Madhuri M.Chandanbatwe, [2]Prof.P.M.Jadhav, [3]Prof.M.R.Ingale
[1]PG Student, [2]professor, Department E&TC
Maharashtra Institute of Technology,Pune, India
Email: [1]madhuric10@gmail.com , [2]parul.jadhav@mitpune.edu.in

Abstract:In the scalable video coding extension of H.264 standard, the concept of inter-layer prediction is adopted to reduce the redundancy between the base layers and the enhancement layers. Thus, additional coding modes should be calculated during the macroblock mode decision procedure at the enhancement layers. It increases the computational complexity and may obstruct its practical applications. A fast mode decision scheme is developed to speed up the encoding of the enhancement layers. It makes use of reducing the number of the candidate modes and early termination of the mode decision procedure adaptively.

Keywords :macroblock mode decision, spatial scalable video coding, enhancement layer

I. INTRODUCTION

H.264/SVC is developed by the joint video team (JVT) as an extension of H.264/advanced video coding (AVC).In fact, many features or tools are introduced to get a better coding efficiency which causes a major increase in terms of video encoder computational complexity.In comparison with previous video coding standards, the purpose of the scalable video coding is to encode the signal once, but enables its decoding from partial streams depending on the specific rate and resolution required by certain applications . Figure 1 shows the H.264/SVC encoder with three spatial layers. The base layer (BL) holds a lower resolution or a reduced quality version of each coded frame. For the enhancement layers (ELs), Layer 1 and Layer 2, the input is in a higher resolution and will be coded as an ordinary H.264/AVC along with inter-layer predictions that will provide additional coding choices, such as inter-layer motion vectors, intra-prediction, and residual information deduced from the BL . SVC presents three types of scalability : temporal, quality, and spatial scalability. Temporal scalability in the SVC is achieved using a structure of hierarchical B pictures inherited from H.264/AVC standard without any additionalspecific extension. Thus, a temporal SVC algorithm allowsthe extraction of a video of multiple frame-rates from asingle coded stream . However, quality scalabilitycan be considered as a special case of spatial scalabilitywith identical picture sizes for base and ELs. Yet,spatial scalability offers additional information using interlayer prediction mechanisms. This tool presents three types of inter-layer prediction: intra, motion, and residual.

The inter-layer intra prediction (ILIP) is defined to predict the EL macroblock (MB) from the previously coded BL intra MB. The ILIP may be used when theenhancement layer MB is coded with basemode flag equal to 1 and the co-located MB in its reference layer is intracoded.The purpose of adding the inter-layer motion prediction(ILMP) is to save bits by reusing the motion informationfrom the BL. This technique holds up the prediction at MBlevel. Thus, for a MB in EL, the associated referenceindexes and motion vectors are derived from the correspondinginformation of the co-located MB in the referencelayer .In order to operate the residual information coded forinter-coded MBs in a lower resolution, an inter-layer predictionis employed. Therefore, whatever the type of a MB,an additional flag is added to each MB syntax to notify theusage of inter-layer residual prediction (ILRP). If this flag is true, the residual signal of a reference layer is up-sampledwith the aim of being used as a prediction for theresidual signal of the enhancement layer MB . Consequently,the difference between the residual data of the EL and the BL reconstructed residual is coded.In this paper, a fast mode decision algorithm was proposed for inter- and intra-frame coding spatial scalability.The goal of our proposed algorithm was to reduce the number of candidates in mode decision**.**

II. CONCEPT OF RESIDUAL PREDICTION

Scalability is a highly desirable feature for video coding because a single scalable bit stream can be used to adapt to various network conditions of the transmission channel, or to different rendering capabilities of a terminal device where the video is decoded and displayed. Residual prediction is a mechanism that involves using the base layer prediction residual to predict the enhancement layer prediction residual. This mechanism is mainly designed for inter-coded macroblock so that a feature called single loop decoding can be enabled. Intra-coded macroblocks from the base layer are fully decoded and reconstructed so that they may be upsampled and used to directly predict the current enhancement layer being decoded. However inter-coded macroblock from the base layer are not fully decode and reconstructed. Instead, only prediction residual of each base layer inter-coded macroblock is decoded and may be used to prediction enhancement layer prediction residual, but no motion compensation is done on the base layer inter-coded macroblock itself.

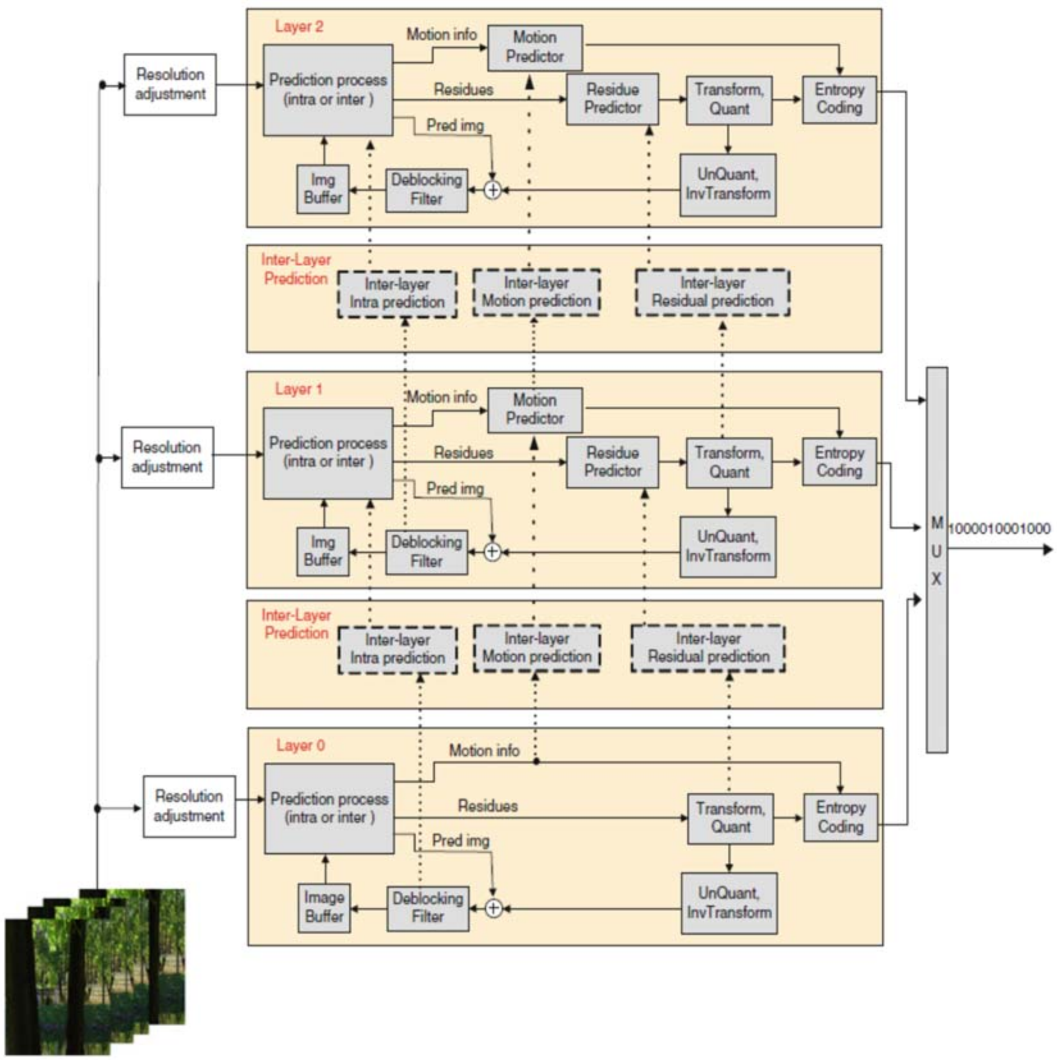

Fig 1.SVC Encoder structure with three spatial layers

The distinction is illustrated in figure 2. In this simple example, the enhancement layer macroblocks E, F, G, H are each covered by only one upsampled base layer macroblocks, and they are A, B, C, D respectively

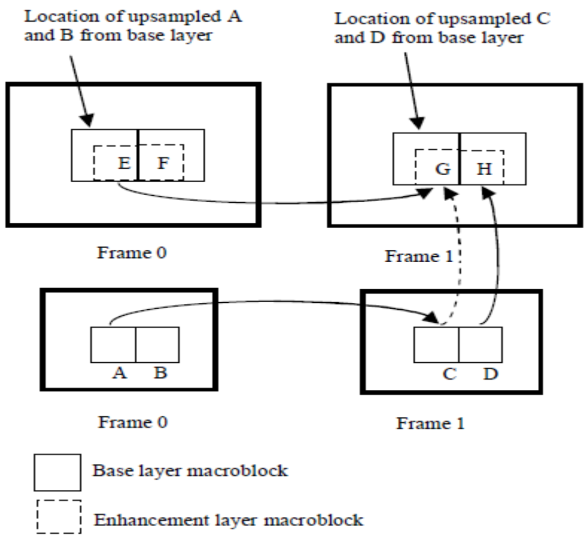

Fig 2.Distinction between regular upsampling and residual prediction

According to this example, macroblock D is inter-coded, and the enhancement layer macroblock H is coded using inter-layer prediction. This means that the prediction for H is formed from the fully reconstructed and upsampled version of D, and the difference(or residual) between the original pixels of H and the prediction is coded into the enhancement layer bit stream. Using 'O' to indicate the original pixels, 'U' the upsampling function and 'R' the decoding and reconstruction function, we can write the prediction residual of H as in(1).

E(H)=O(H)-U(R(D))…………..(1)

Continuing the example, base layer macroblock C is inter-coded, with prediction from base layer macroblock A. This prediction can be written as PAC, and thus

E©=O©-PAC ……………….(2)

Similarly, the enhancement layer macroblock G is inter-coded. In H.264/AVC , two types of inter-coding are possible in the enhancement layer: with and without residual prediction. Without residual prediction, enhancement layer macroblock G is predicted from enhancement layer macroblock E, written as PEG, in the standard fashion, which can be written along the lines of equation (2) as

E(G)=O(G)-PEG …………………..(3)

However, if the residual prediction technique is used for macroblock G, the residual of base layer macroblock C, i.e. O©-PAC, is also used together with the full reconstruction of enhancement layer macroblock E to form a prediction for G. in this case, prediction residual of macroblock G can be expressed as in (4).

E(G)=O(G)-P'EG-U(O©)-PAC …………………(4)

Here, U(O©-PAC) is simply the upsampled prediction residual from macroblock C that is decoded from the bit stream P'Egis the prediction formed from macroblock E, under residual prediction mode. Depending on the motion vectors of G when coded with regular and residual prediction respectively, P'EG, may or may not be same as PEG.

Whether regular or residual prediction is used in the coding of inter-coded enhancement layer macroblocks involves a ratedistortion analysis, and is performed and the encoder's mode decision module.

It is important to note that equation (4) cannot be simplified to a direct inter-layer prediction, i.e. O(G)-U(R©), along the lines of equation (1). This is because the fully reconstructed value C, i.e. R©, is not available when decoding the enhancement layer due to the principle of single-loop decoding. As stated earlier, this principle is borne from the desire to performing motion compensation at only one layer, regardless of how many layers are to be decoded.

III. MODE DECISION SCHEME

Scalable video coding (SVC) has been standardized to extend the capabilities of the H.264 advanced video coding (AVC). The SVC can compress several video sequences of various resolutions as a single bit-stream. In the SVC enhancement layer, for Joint Scalable Video Model (JSVM) software implementation, an exhaustive mode decision process based on the base layer mode predictions is performed to obtain the best mode for each macroblock (MB). This technique may achieve a higher coding efficiency; however, it induces a significant computational complexity in the encoding engine. In order to speedup the SVC encoder, a fast mode decision algorithm was proposed . In other words, our aim was to decrease the number of candidate modes to reduce the computational complexity and maintain the same level of coding efficiency, this approach used the spatial and temporal correlation between MB situated at the enhancement layer and its co-located MB at the base layer.For intra- and inter-prediction, SVC presents several modes.

IV. INTRA-MODE CODING

Each MB can be coded in one of several coding types depending on the slice-coding type. For all slice-coding types, SVC offers three choices which are denoted as Intra 16* 16, Intra 8 *8, and Intra 4 *4. Figure illustrates candidate modes of intra coding.

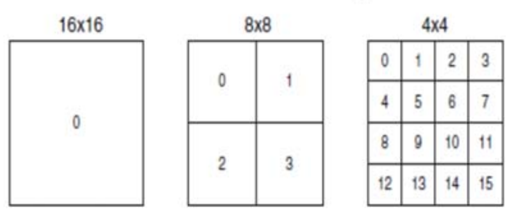

Fig. 3 Candidate modes of intra coding H.264/SVC

Intra 4 *4 mode is based on predicting each 4 *4 block independently. This mode is well suited for coding parts of a picture with significant details. Whether Intra 16 *16 or Intra 8 *8 modes, they both are usually used to code smooth areas of a picture. Furthermore, for intra coding area, SVC provides an additional intra-prediction mode in which the prediction signal is formed by the reconstructed and up-sampled BL signal.

V. INTER-MODE CODING

H.264/SVC presents a large range of inter-modes, such as{Skip, Inter 16 9 16, Inter 16 9 8, Inter 8 *16, Inter 8 *8, Inter 8 *4, Inter 4*8, Inter 4* 4}. A MB may be split into MB partitions {Inter 16 *16, Inter 16 *8, Inter 8 *16, Inter 8 * 8}, and Inter 8 *8 may be portioned into sub-MB partitions {Inter 8 *4, Inter 4 *8, Inter 4 *4}.Figureilustrates all the candidate modes of H.264/SVC.

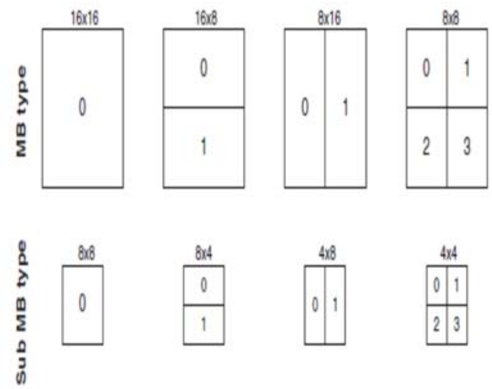

Fig. 4 Candidate modes for inter coding H.264/SVC

In addition to those modes, other candidates are derived from the up-sampling inter information inherited from BL.

Mode decision algorithm used by Joint Scalable Video Model (JSVM) reference software is called full mode decision (FMD). The idea of FMD is to compute rate distortion cost (rdcost) of candidate modes for a MB to decide the less costly mode. This model leads to enormous computational complexity. The major idea of a fast mode decision algorithm proposed in this paper was to decrease the number of candidate modes according to their occurrence probabilities

VI. PROPOSED ALGORITHM

For I frame, we limit the candidates in the mode decision process to Intra 4 * 4 and the BL_mode that was selected at BL. According to previous analyses, each type of frame has its own characteristics.

For P frame, the proposed method is divided into two parts according to the mode coding. For intra coding type, I frame type. In fact, only BL_mode and Intra 4 *4 are the candidates in mode decision process. For inter coding type, we proceed as follows: if the BL mode is Inter 16*16, check the BL_mode, the skip mode and Inter 16 *16 mode at EL; if BL mode is Inter 16* 8, check the BL_mode, the skip mode and Inter 16 *8 at EL; if BL mode is Inter 8 *16, check the BL_mode, the skip mode and Inter 8*16 at EL; finally, when the mode at BL is Inter 8 *8 we limit the mode decision to the BL_mode, the skip mode and Inter 8 9 8. In this process, we eliminate the sub-block candidates. Otherwise, we check the skip, Inter 16 *16, Inter 16 *8,Inter 8 *16, and Inter 8 *8 only.

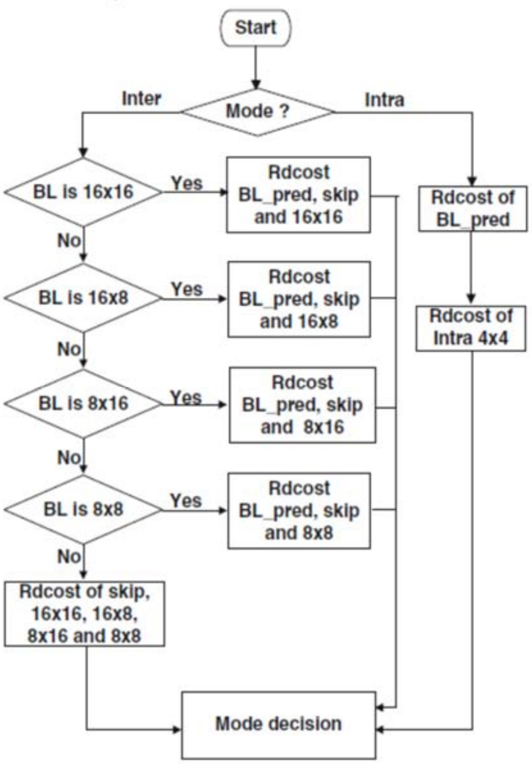

Fig.5 Algorithm for P frame

VII. SIMULATION RESULTS:

Simulation is done on Microsoft Visual Studio using the Joint Scalable Video Model(JSVM).The main configuration file consists of layer configuration files of base layer input file with spatial resolution as that of QCIF format and Enhancement layer input file of CIF format.To evaluate computational complexity reduction, the total time saving (TS) is computed using the following formula:

$$TS\,(\%) = \frac{T_{JSVM} - T_{alg}}{T_{JSVM}} \times 100$$

where TJSVM =CPU time needed to encode the original EL

Talg =CPU time necessary to encode the applied algorithm.

Change in PSNR:

ΔPSNR=PSNRalgorithm – PSNR0

Layer macroblock coded:

11*9=99 for QCIF(176*144)i.e base layer L0

Layer macroblock coded:

22*18=396 for CIF(352*288)i.e Enhancement layer L1

frame resoln		bitrate	Min-btr	Y-PSNR	U-PSNR	V-PSNR
176x144	3.75	39.36	39.36	34.678	37.998	39.5252
176x144	7.5	40.83	40.83	34.635	37.998	39.538
176x144	15	55.8	55.8	34.63	37.998	39.5376
176x144	30	71.61	71.616	34.641	38.007	39.5531
176x144	60	82.032	82.032	34.664	38.001	39.5581
352x288	3.75	105.84	105.84	36.297	39.778	41.6176
352x288	7.5	110.85	110.85	36.15	39.782	41.6114
352x288	15	155.32	155.32	36.178	39.803	41.6669
352x288	30	211.44	211.44	36.216	39.888	41.6906
352x288	60	277.344	277.34	36.262	39.913	41.7566

SUMMARY

VIII. CONCLUSION

Scalable video coding is a good solution for video transmission over heterogeneous networks. To satisfy a higher coding efficiency, a good mode decision process is required. In this paper, we presented a fast mode decision algorithm for intra- and inter-frame coding in SVC The number of candidate modes in MB that takes part in rate distortion computation has been reduced significantly.

ACKNOWLEDGMENT

I am indeed thankful to my guide Prof. P.M.Jadhav and Prof.M.R.Ingale for their able guidance to complete this paper. I extend my special thanks to Head of Department of Electronics and Telecommunications Prof. (Dr.) G. N. Mulay who extended the preparatory steps of this paper-work

I am also thankful to the Principle Dr. L. K. Kshirsagar, Maharashtra Institute of Technology for his valued support and faith on me.

IX. REFERENCES

i. H. Schwarz, D. Marpe, and T. Wiegand, **“Overview of the scalable video coding extension of the H.264/AVC standard**,” *IEEE Trans.Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

ii. T. Wiegand, G. J. Sullivan, G. Bjøntegaard, 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

iii. C. Andrew Segall, *Member, IEEE*, and Gary J. Sullivan, *Fellow, IEEE* **“Spatial Scalability Within H.264/AVC Video Coding Extension** “IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 17, NO. 9, SEPTEMBER 2007 .

iv. Xianglin Wang, Justin Ridge, Nokia Research Center, Palo Alto California, USA“**Improved Video Coding with Residual Prediction for Extended Spatial Scalabilty,”** ISCCSP 2008,Malta,12-14 March 2008.

v. Richardson, L.E.G.: H.264 and MPEG-4 video compression: video coding for next-generation multimedia. Wiley, New York.

vi. Li, X., Amon, P., Hutter, A., Kaup, A.: “**Performance analysis of inter-layer prediction in scalable video coding extension of H.264/AVC**.” In: Proceedings of the IEEE Transactions on Broadcasting, vol. 57, no. 1 (2011)

vii. Lin, H.C., Peng, W.H., Hang, H.M.: **“Fast context-adaptive mode decision algorithm for scalable video coding with combined coarse-grain quality scalability (CGS) and temporal scalability”.**In: Proceedings of the IEEE Transactions Circuits and Systems for Video Technology, vol. 20, no. 5, pp. 732–748 (2010).

viii. Lee, B.S., Kim, M.C., Hahm, S.J., Park, C.S., Park, K.S.: **“A fast mode selection scheme in inter-layer prediction of H.264 scalable extension coding.”** In: Proceedings of the IEEE International Symposium on Broadband Multimedia Systems Broadcasting ,pp. 1–5 (2008).