http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 9 (4): 643-651, 2010
ISSN 1812-5638

Performance Modeling and Optimization of Hierarchical B Pictures Based on Directed Tree

[1]Zheng Zhu, [1,2]Dong-Xiao Li and [1,2]Ming Zhang
[1]Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027, China
[2]Zhejiang Provincial Key Laboratory of Information Network Technology, Hangzhou 310027, China

Abstract: A general methodology for performance modeling and optimization of hierarchical B pictures used in video coding is proposed in this study. First, an approach of directed tree decomposition is used to denote hierarchical B pictures. In combination with a proposed linear model for compression efficiency, the performance of any hierarchical B prediction structure can be evaluated conveniently from the directed tree. Then with a dynamic programming method, the optimal tree for any group length of hierarchical B pictures can be set up elegantly from recursive subtrees. This method can be adapted to tradeoff between compression efficiency and random access ability. Besides, experimental results show that the optimal tree of compression efficiency achieves higher performance than the existing hierarchical B prediction structures.

Key words: Video coding, hierarchical B pictures, prediction structure, directed tree

INTRODUCTION

H.264/AVC offers much more flexibility on a picture/sequence level in comparison to prior video standards including coding using hierarchical B pictures (Leontaris and Cosman, 2007). Hierarchical B pictures play an important role in both traditional video coding and multi view video coding (MVC), because they have the benefit of significantly improved rate distortion efficiency besides providing temporal scalability (Winken *et al*., 2007). The HHI proposed the reference view-temporal prediction structure for the standardization of MVC, wherein hierarchical B pictures are used as basic structure for temporal prediction (Muller *et al*., 2006; Merkle *et al*., 2007). As efficient compression is essential for the applications of MVC, many other algorithms were proposed subsequently to achieve even better compression performance. For example, a method of minimum spanning tree was proposed by Li *et al*. (2007) and Kang *et al*. (2007), respectively. However, B picture cannot be involved in their prediction structures because it would yield loops. Park *et al*. (2008) proposed a temporal hierarchical B prediction structure, aiming to minimize the average reference distances. However, as the experimental results reveal, the sum of reference distances is not a proper index of coding efficiency of B pictures. Therefore, further studies are still necessary to establish a reasonable model for prediction structure including B picture. With respect to the quantitative evaluation of prediction structures, Chen *et al*. (2007) and Liu *et al*. (2007) presented a normalized exponential factor model for multi-hypothesis prediction. However, the non-linear model will bring high complexity in estimation procedure and difficulty in implementation, which is not suitable for practical applications.

In this study, a novel model is proposed to analyze the performance of hierarchical B pictures by mapping them to directed trees. The directed trees have some interesting traits that are helpful to the research on the following two problems: (1) the quantitative evaluation of compression efficiency and random access ability and (2) search for the optimal prediction structure. For the evaluation of compression efficiency, a linear model is presented. This model dovetails well with the prediction performance of B pictures by using the Least Squares Estimation (LSE). During the process of searching for the optimal picture structure, a Dynamic Programming (DP) approach is implemented to set up the optimal tree from the recursive subtrees. The optimal tree can be set adapted to the tradeoff between compression efficiency and random access ability. Experimental results show that the optimal tree of compression efficiency achieves higher performance than the existing hierarchical B prediction structure supported by H.264/AVC Joint Model (JM). It can be seen that the proposed approach is applicable to both traditional and multi-view video coding.

MAPPING HIERARCHICAL B TO DIRECTED TREE METHOD

Hierarchical B pictures represent a coding structure that uses bidirectional predictive pictures (B pictures) as

Corresponding Author: Dong-Xiao Li, Room 501, Department of Information Science and Electronic Engineering, Zheda Road 38, Zhejiang University, Hangzhou 310027, China Tel/Fax: +86-571-87952017

references for other B pictures within One Group of Pictures (GOP). A typical hierarchical B prediction structure is depicted in Fig. 1 for a GOP of eight pictures with four temporal levels (Schwarz *et al*., 2006). A picture is called key picture (black in Fig. 1), when all previously coded pictures precede the picture in display order. Pictures between the current and previous key pictures as well as the current key picture itself are considered to build a GOP. The key pictures are coded in regular intervals and the remaining pictures of a GOP are hierarchically predicted by using the two nearest pictures of the next higher temporal level as references.

In this study, a novel model is used to evaluate the performance of hierarchical B pictures by mapping them to directed trees. Figure 2a-d demonstrate the mapping process for a typical case in Fig. 1. Root of the tree is used to denote the GOP length L (L = 8 in this case). Then a GOP is divided into two parts by the first layer of B picture (B_3 in Fig. 2), each part with a length of 4. We use two nodes to denote the two parts, with the value of their lengths, as shown in Fig. 2a. These two nodes are drawn as child nodes of the root of tree and the sum of them equals to the value of their father node exactly. Next, each part is again divided into two parts by the second layer of B pictures B_1 and B_5, respectively, as shown in Fig. 2b and c. In the same way, we get the nodes of this layer that indicate the lengths of these new smaller parts. This process continues until all the B pictures are expressed in the tree (Fig. 2d). In the dyadic cases, one part is divided into two parts by B pictures each time and the corresponding results are binary trees. Figure 3 shows another dyadic case for a GOP length of eight.

This model is not only applicable to dyadic cases of hierarchical B prediction structures. Figure 4a and b show a triadic case of hierarchical B structure with a GOP length of eight. The GOP is divided into three parts by the

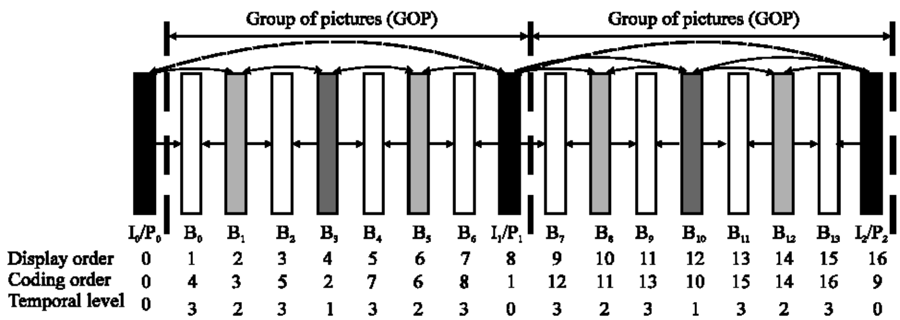

Fig. 1: Typical hierarchical B coding structure with multiple levels for a GOP length of eight

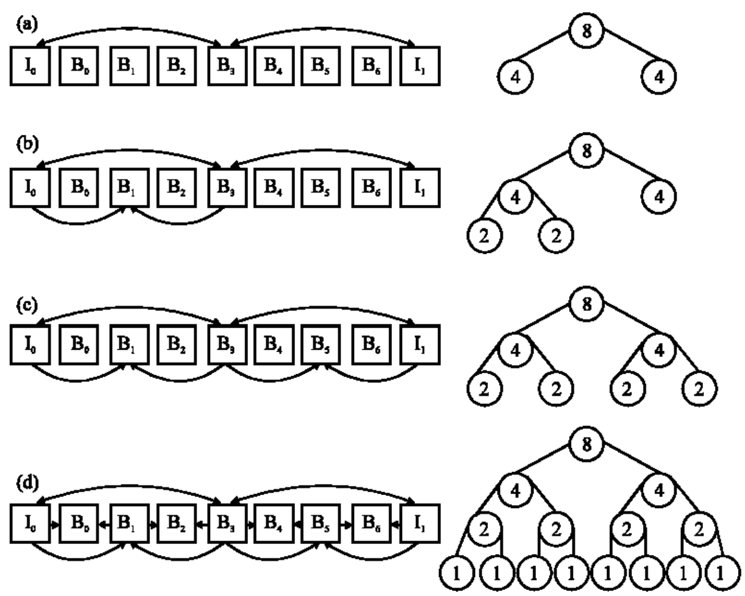

Fig. 2: (a-d) Mapping process for a typical hierarchical B structure for a GOP with length of eight to a directed tree

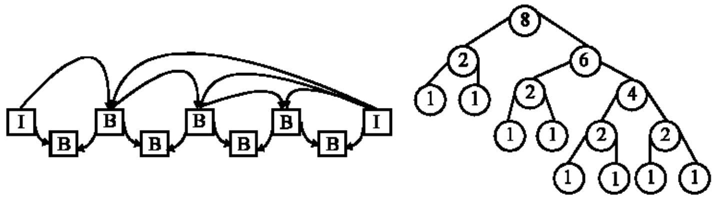

Fig. 3: Directed tree for another dyadic case of a hierarchical B structure with a GOP length of eight

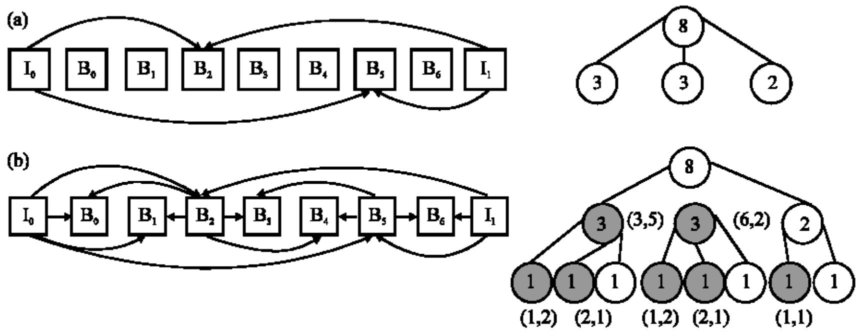

Fig. 4: Mapping process for a triadic case with a GOP length of eight

first layer of B pictures B_2 and B_5, with the lengths of 3, 3 and 2. And then the left two parts are again divided into three parts and the right part is divided into two parts. Accordingly, the mapping result is no more a binary tree.

We can get the forward and backward reference distances of B pictures from a directed tree easily. A group of brother nodes with a number of N is corresponding to N-1 B pictures of this layer. The left N-1 brother nodes are used to denote the corresponding N-1 B pictures, as grayed in Fig. 4b. We call them B-nodes. Let B_i be the value of the ith B-node in a group of brother nodes and F be the value of father node. The forward reference distance of the nth B picture

$$D_{n1} = \sum_{i=1}^{n} B_i$$

and the backward reference distance $D_{n2} = F\text{-}D_{n1}$, as shown in the brackets in Fig. 4b. For the dyadic cases, the left one of a paired brother nodes is always the B-node. It is straightforward to get the reference distances of B pictures from a binary tree because the paired brother nodes exactly indicate that, the proposed directed tree offers an intuitionistic way to exploit the features of prediction structures.

EVALUATION MODEL FOR HIERARCHICAL B PICTURES

Evaluation for compression efficiency: Motioncompensated coding schemes achieve compression by exploiting the similarities between successive frames of a video signal (Girod, 1987, 2000). In this sense, the prediction efficiency determines the compression performance of the prediction structure. Let PE denote the prediction efficiency of a B picture, we define PE as:

$$PE = OUTBITS_B / OUTBITS_I \qquad (1)$$

where, $OUTBITS_B$ denotes the output bits number of a frame coded as B picture and $OUTBITS_I$ denotes the output bit number of the same frame when it is intra-coded.

Assumed that PE is only relative to the reference distances (Chen *et al*., 2007; Liu *et al*., 2007), a general monotonically increasing function PE = f (D_1, D_2) can be used, where, D_1 and D_2 denote the reference intervals. A common approach to get an approximation of PE is to set up a linear model with LSE (Steven, 1993). In the linear model, PE is expressed using matrix notation as $PE = H\theta$, where, H is an observation matrix related to D_1 and D_2 and θ is a vector of parameters to be estimated. The LSE is found by minimizing the cost:

$$J(\theta) = (x-H\theta)^T(x-H\theta) \qquad (2)$$

where, x is a vector of actual data of PE that can be observed. Setting the gradient equal to zero yields the LSE (Steven, 1993):

$$\theta = (H^TH)^{-1}H^Tx \qquad (3)$$

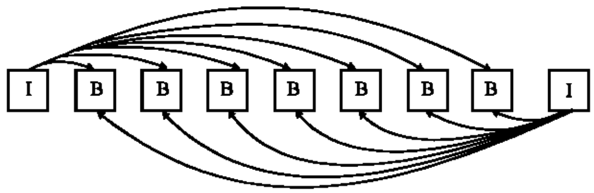

Fig. 5: Coding structure to get the prediction efficiency of B pictures

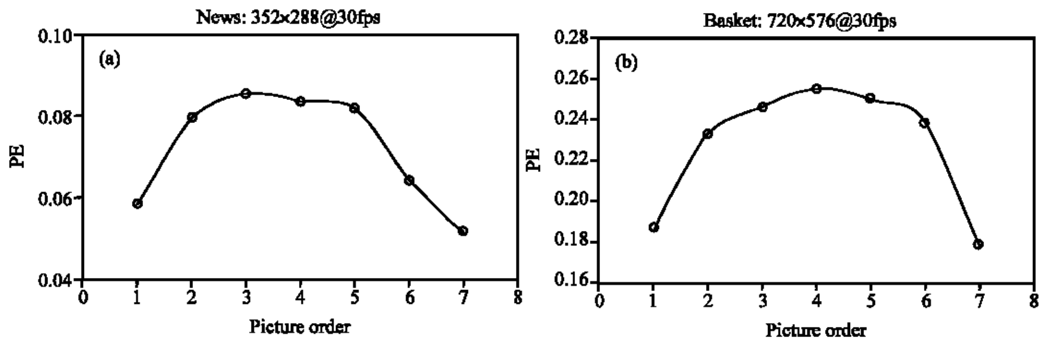

Fig. 6: PE for sequence news_cif and basket_ccir when the reference intervals satisfy $D_1+D_2 = 8$

Experiments are implemented on H.264/JM11.0 to see the behavior between PE and reference intervals D_1 and D_2. As shown in Fig. 5, I pictures are coded in regular intervals L (e.g., L = 8 in Fig. 5) with B pictures located between them. As none of the B pictures is marked as reference, thus, we can get the actual values of PE for the case when $D_1+D_2 = L$. Figure 6a and b show the averaged results for sequences news_cif and basket_ccir with L = 8. The x-axis denotes the display order of B pictures in a GOP. Meanwhile, it also expresses the combination of reference intervals D_1 and D_2 like this: $(D_1, D_2) = (1,7), (2,6), (3,5), \ldots (6,2), (7,1)$. It is noted that the curves show the logarithmic property when L is fixed.

The logarithmic property between PE and (D_1, D_2) implies that the function $PE=f(D_1, D_2)$ should also has a logarithmic feature. Hence, in the proposed model, a logarithmic function is used for approximation of PE:

$$PE = f(D_1, D_2) = \theta_1+\theta_2\log(D_1.D_2) \quad (4)$$

where, $\theta = [\theta_1, \theta_2]$ denotes the parameters to be estimated. Experimental platform is built on H.264/JM11.0 for further verification. By switching NumberBFrames from 2 to 7, the actual values of PE for B pictures with different reference intervals are obtained. Then LSE is applied and the results of sequences news_cif and basket_ccir are shown in Fig. 7a and b. The x-axis denotes the combination of D_1 and D_2 and is arranged like this: first L = 2, $(D_1, D_2) = (1,1)$; then L = 3, $(D_1, D_2) = (1,2), (2,1)$; then L = 4, $(D_1, D_2) = (1,3), (2,2), (3,1)$; ... last of all, L = 8, $(D_1, D_2) = (1,7), (2,6), (3,5), \ldots (6,2), (7,1)$. As the results reveal, the proposed model dovetails well with the actual data in most cases.

From the model proposed, the prediction efficiency of a picture structure can be evaluated by the sum of the efficiency of B pictures:

$$PE_{GOP} = \sum_{i=1}^{L-1} f(D_{i1}, D_{i2}) = \theta_1(L-1) + \theta_2 \log(\prod_{i=1}^{L-1} D_{i1} \cdot D_{i2}) \quad (5)$$

where, L denotes the length of a GOP. To evaluate the relative efficiency of prediction structures, the actual values are not concerned, so the expression can be simplified to:

$$PE_{GOP} = \log\prod_{i=1}^{L-1} D_{i1} \cdot D_{i2} \quad (6)$$

The average efficiency of B pictures is $Pe_{aver} = PE_{GOP}/(L-1)$. Thus, the compression performance of a prediction structure can be evaluated straightforward by the products of reference distances of B pictures. For the dyadic cases, we can simply use the value of the paired nodes of the corresponding binary tree. For example, in case of GOP length of eight, PE_{GOP} of prediction structure in Fig. 2 can be obtained immediately to be log (256) and PE_{GOP} of prediction structure in Fig. 3 is log (384). For the non dyadic case of Fig. 4, the reference distances must be figured out for each B-node first and then we can get the PE_{GOP} of log (2880). Thus, the prediction structure in

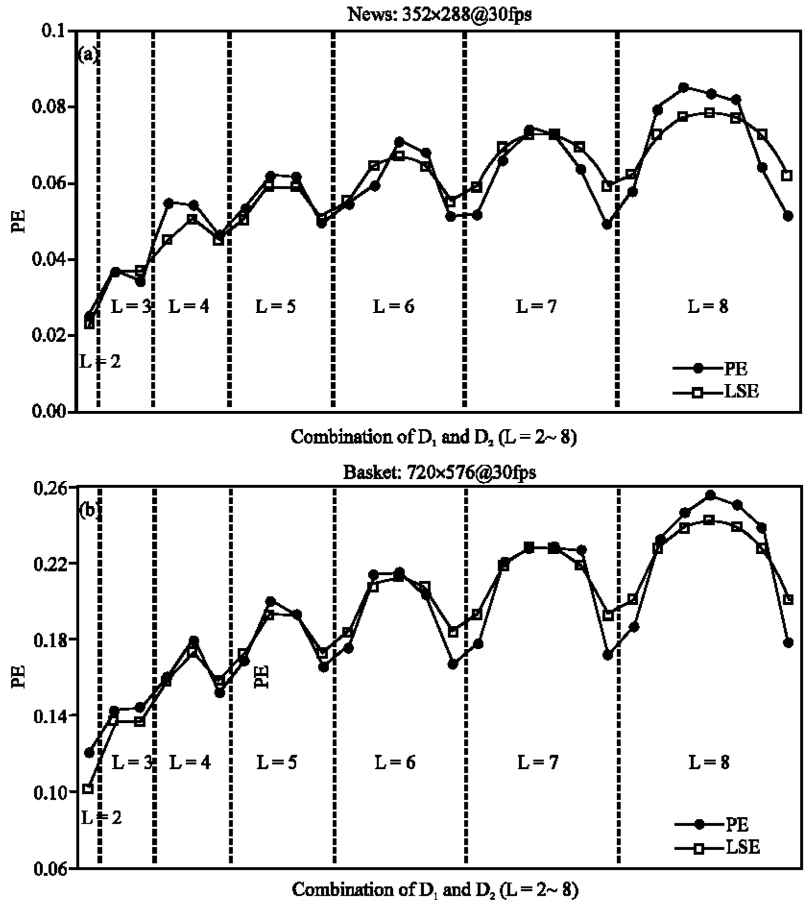

Fig. 7: The estimation results for PE with the logarithmic model in temporal dimension. PE denotes the actual values of prediction efficiency and LSE denotes the results of the proposed linear model by implementing least squares estimation. L denotes the coding interval of I pictures and x-axis expresses the combination of reference distances D_1 and D_2 and is arranged like this: first L = 2, (D_1, D_2) = (1,1); then L = 3, (D_1, D_2) = (1,2), (2,1); then L = 4, (D_1, D_2) = (1,3), (2,2), (3,1); … last of all, L = 8, (D_1, D_2) = (1,7), (2,6), (3,5), …(6,2), (7,1)

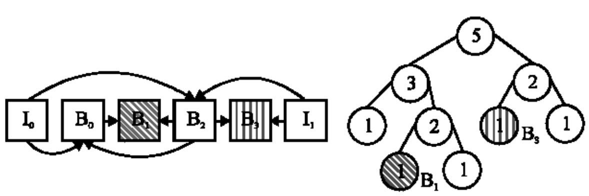

Fig. 8: Random access ability indicated by the height of nodes

Fig. 2 is regarded to be more efficient than that in Fig. 3 and both of them are regarded to be much more efficient than that in Fig. 4.

Evaluation for random access ability: Though the compression efficiency is usually the most important requirement, random access ability (RA) is also a desirable feature in many video communications. Besides, some other important functionality such as coding delay and structure complexity are closely related to random access ability. A straightforward and logical way to evaluate the random access ability is to investigate the number of pictures that must be pre-decoded before each B picture. It is observed that in our proposed directed tree, the distance from root to a B-node exactly reflects that. As illustrated in Fig. 8, the number of pictures must be pre-decoded for B_1 is three in the GOP (I_1, B_2, B_0) and for B_3 the number is two (I_1, B_2). In the corresponding directed tree, it is clear to see that the distance from root to the B-node indicating B_1 is three (twilled in Fig. 8) and for B_3 the distance is exactly two (vertical lined in Fig. 8).

In graph theory, the nodes at distance k from root are defined to have height k and form the kth level of tree (Reinhard, 2005). Let H_i be the height of the ith B-node, the average random access ability of prediction structure can be simply defined by:

$$RA_{aver} = \sum_{i=1}^{L-1} H_i /(L-1) \tag{7}$$

Again, we compare the prediction structures for a GOP length of eight in Fig. 2-4 for example. For the prediction structure in Fig. 2, RA_{aver} equals 17/7 and in Fig. 3 it is 19/7 and in Fig. 4 it is 12/7. Thus, the prediction structure in Fig. 4 has the best random access ability and the prediction structure in Fig. 2 is considered to be better than that in Fig. 3.

OPTIMAL TREE

An interesting feature of the directed tree is that the optimal tree can be constructed elegantly from recursive subtrees with a Dynamic Programming (DP) approach. This solution can be used for the tradeoff optimization between compression efficiency and random access ability. We define the cost function of prediction structure as:

$$C = PE_{aver} + \lambda . RA_{aver} \tag{8}$$

where, λ is the penalty factor for random access ability. Minimizing the cost function yields the optimal prediction structure. The optimization is a recursive process. For any GOP length L, we do not need to investigate all the possible structure of trees but only have to be concerned about the son nodes of root. For an optimal tree, each son node of root and its posterity nodes should have to form an optimal subtree. Thus, if the optimal tree for GOP lengths from 2 to L-1 have been obtained, it is easy to get the optimal tree for GOP length of L.

When λ is set to 0, the cost function is expressed only by PE_{aver} and we can get the optimal trees of compression efficiency. Note that a non-binary tree can be modified to a binary tree to get smaller reference distances for some B pictures, so an optimal tree of compression efficiency must be a binary tree. The searching process for optimal tree of compression efficiency can be simplified to be implemented among binary trees. Based on the proposed model for prediction efficiency, we simply use multiplication as the cost function. Let C_i be the minimum cost for GOP length of i. (1) Starting from GOP length of 2, there is only one prediction structure and the optimal tree is certainly that one, C_2= 1.1; (2) If C_i (I = 2…L-1) have been obtained, the minimum cost of C_L can be found out by compare the following values: $1.(L-1).C_{L-1}$, $2.(L-2).C_2.C_{L-2}$,…, $(L/2).(L/2).C_{L/2}.C_{L/2}$ (when L is even) or $((L-1)/2).((L+1)/2).C_{(L-1)/2}.C_{(L+1)/2}$ (when L is odd). For example, for L=8, if C_i (i=2~7) are obtained, we can simply compare $1.7.C_7$, $2.6.C_2.C_6$, $3.5.C_3.C_5$ and $4.4.C_4.C_4$ to find out the minimum cost C_8, as shown in Fig. 9.

With this method, we set up the optimal trees of compression efficiency with GOP length from two to twenty and the results are illustrated in Table 1. Note that the Pe_{aver} reflect the relative prediction efficiency of B pictures, which cannot be used as an index of the actual coding bit rates of different GOP lengths because the I pictures are not counted in. In fact, for a smaller GOP length, the average prediction efficiency of B pictures is better because of the shorter reference intervals, just as Table 1 shows, whereas the entire bit rate may be higher for I pictures are adopted more frequently.

By looking up the Table 1, the optimal prediction structure of compression efficiency for any given GOP length can be set up immediately. As shown in Fig. 10,

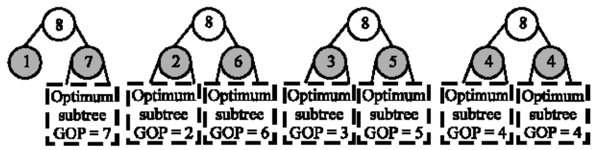

Fig. 9: The recursive structure of optimal tree for a GOP length of eight

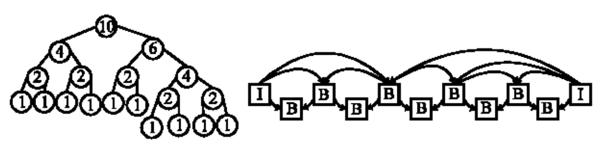

Fig. 10: Optimal tree for a GOP length of ten

Table 1: Optimal trees for GOP lengths from 2 to 20

GOP length	Left subtree	Right subtree	PE_{aver}	RA_{aver}
2	1	1	0.00	1.0000
3	1	2	0.3466	1.5000
4	2	2	0.4621	1.6667
5	2	3	0.6212	2.0000
6	2	4	0.6931	2.2000
7	3	4	0.7607	2.3333
8	4	4	0.7922	2.4286
9	4	5	0.8584	2.6250
10	4	6	0.8922	2.7778
11	4	7	0.9283	2.9000
12	4	8	0.9452	3.0000
13	5	8	0.9766	3.0833
14	6	8	0.9909	3.1538
15	7	8	1.0096	3.2143
16	8	8	1.0166	3.2667
17	8	9	1.0430	3.3750
18	8	10	1.0563	3.4706
19	8	11	1.0725	3.5556
20	8	12	1.0793	3.6316

for GOP length of ten, first in item 10 of the GOP length column we get the root of left and right subtree are four and six; then in the item 4 and 6 of GOP length column, subtrees of the next level can be obtained; repeating the process until the binary tree is finished; finally the prediction structure can be drawn up according, to the tree.

It is observed that the prediction structures based on optimal trees of compression efficiency are rather regular. Given any GOP length L, one of the subtrees is always the power of two, $D_1 = 2^m$. The power m is a positive integer and can be decided as follows: let $n = \lfloor \log_2 L \rfloor - 1$, if $L = 2^n \times 3$, m = n+1, otherwise, m = n. The root of another subtree satisfies $D_2 = L - D_1$. For some GOP lengths the optimal tree turns to be identical with the existing typical hierarchical B prediction structure supported by H.264/JM (for example, when L equals the power of two).

For tradeoff optimization, the searching process is a bit more complicated than compression efficiency optimization. The penalty factor for random access ability λ can be set arbitrarily as requirement and non-binary trees should have to be taken into account during the searching process.

RESULTS FOR OPTIMAL TREE OF COMPRESSION EFFICIENCY

The experimental platform is built on H.264/JM11.0 codec to evaluate the performance of hierarchical B prediction structure based on optimal tree of compression efficiency. Two existing hierarchical B prediction structures supported by H.264/JM are used as reference by setting HierarchicalCoding to 1 and 2. Then switching HierarchicalCoding to 3 and setting ExplicitHierarchyFormat yields the hierarchical B prediction structure of optimal trees. Eight typical test sequences of different resolutions (QCIF, CIF, CCIR and HD) are used and three GOP lengths (seven, eleven and fifteen) are selected for test. Both the reference and proposed hierarchical B prediction structures based on optimal tree are set on quantization parameter cascading (Schwarz *et al*., 2006).

Figure 11 shows the coding results for sequence mobile_cif, where HC1 and HC2 means the reference hierarchical B prediction structures by setting HierarchicalCoding to 1 and 2 and OT denotes the test hierarchical B prediction structure based on optimal tree. To calculate the average differences the Rate-Distortion (RD) curves in detail, the Bjontegaard measure is used (Bjontegaard, 2001, 2008) and experimental results are listed in Table 2. When L = 7, the optimal tree prediction structure is close to the reference structure of HC1, but about 0.1 dB PSNR gains have been achieved over HC2. As GOP length increases, coding gains over HC1 increase too. When L = 15, 0.1~0.4 dB gains can be achieved. The results show that the proposed prediction structure based on optimal tree performs better than the reference prediction structures at any given GOP lengths.

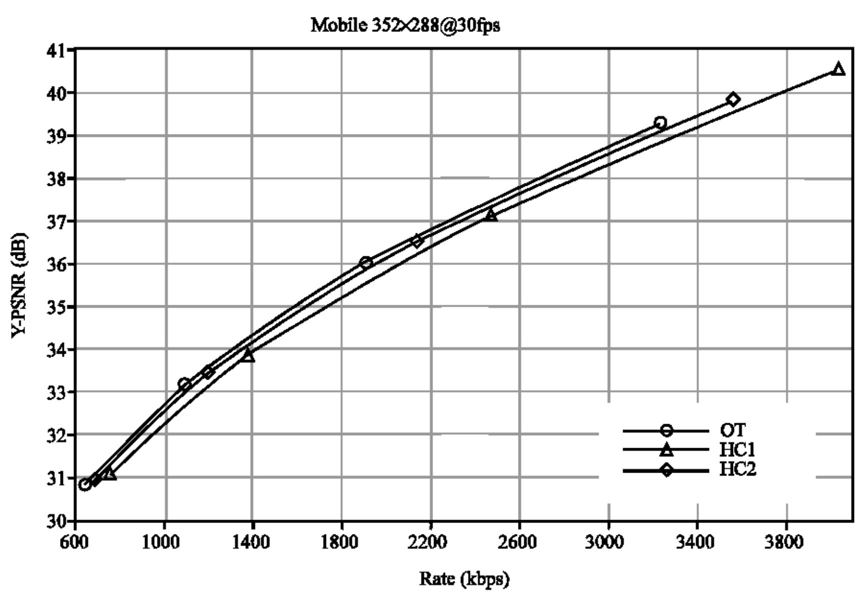

Fig. 11: Coding results for sequence mobile_cif for a GOP length of fifteen. HC1 and HC2 denote the reference hierarchical B prediction structures supported by H.264/AVC JM by setting HierarchicalCoding to 1 and 2 separately and OT denotes the test hierarchical B prediction structure based on optimal tree

Table 2: Coding results for the proposed and reference hierarchical B prediction structures

		Gain over HC1		Gain over HC2	
Test sequences	GOP length	PSNR(dB)	Bit saving (%)	PSNR(dB)	Bit saving (%)
Paris: 176×144@30fps	7	-0.014	-0.159	0.063	0.708
	11	0.028	0.344	0.050	0.613
	15	0.226	2.856	0.073	0.948
Mother_dauter: 176×144@30fps	7	0.002	0.054	0.043	0.675
	11	0.056	0.939	0.021	0.338
	15	0.235	3.937	0.050	0.844
Mobile: 352×288@30fps	7	0.096	1.430	0.121	1.815
	11	0.153	2.577	0.054	0.917
	15	0.369	6.566	0.098	1.805
News: 352×288@30fps	7	-0.022	-0.345	0.036	0.578
	11	0.010	0.174	0.034	0.562
	15	0.115	1.939	0.033	0.564
Flowergarden: 720×576@30fps	7	0.037	0.548	0.128	1.883
	11	0.088	1.376	0.084	1.343
	15	0.417	6.601	0.098	1.600
Basket: 720×576@30fps	7	0.007	0.117	0.109	1.659
	11	0.072	1.145	0.069	1.102
	15	0.211	3.333	0.134	2.163
Crew: 1280×720@30fps	7	-0.026	-1.006	0.013	0.630
	11	0.065	2.923	-0.039	-1.757
	15	0.132	5.889	0.082	3.654
City: 1280×720@30fps	7	0.006	0.235	0.042	1.079
	11	0.014	0.794	0.011	0.408
	15	0.287	9.272	0.065	2.133

HC1 and HC2 denote the reference hierarchical B prediction structures supported by H.264/AVC JM by setting HierarchicalCoding to 1 and 2 separately

DISCUSSION

The prediction structure of hierarchical B pictures has been widely used in video coding for its excellent compression performance. A detailed analysis on hierarchical B pictures was presented by Schwarz *et al.* (2006) regarding their coding delay, memory requirements and compression efficiency. As for the quantitative evaluation of prediction structures, Chen *et al.* (2007) and Liu *et al.* (2007) presented a normalized exponential factor model for multi-hypothesis prediction. In contrast, the linear model proposed in this study is much more convenient to implement. Experimental results showed that the linear model dovetails well with the prediction efficiency of B pictures in most cases. This model can be further simplified to the product of the reference distances of B pictures free from parameters. This feature makes it very easy to compare the compression performance among hierarchical B prediction structures with the same GOP length.

The optimization design of prediction structures is an important issue for video coding, especially for multi-view video coding. The recursive searching approach for optimal tree proposed in this study can be used to construct optimal hierarchical B prediction structures of any GOP length. This method can be adapted to tradeoff between compression efficiency and random access ability by adjusting a penalty factor for random access ability. We set the penalty factor to zero to obtain the optimal prediction structures of compression efficiency. The configuration of these optimal prediction structures can be summarized to a simple formula. In comparison to the existing hierarchical B prediction structures supported by H.264/JM, the configuration complexity of the proposed optimal prediction structures is about the equal. However, as experimental results revealed, higher coding performance was achieved by the proposed optimal prediction structures for any GOP length.

CONCLUSIONS

The directed tree decomposition offers a new perspective to analyze and arrange hierarchical B prediction structures. In combination with the proposed linear model for compression efficiency, it is straightforward to evaluate the performance of any hierarchical B prediction structure from the directed tree. With a dynamic programming method, the optimal tree is set up elegantly, which can be used for tradeoff optimization of compression efficiency and random access ability. Experimental results show that the optimal tree of compression efficiency achieves higher performance than the existing hierarchical B prediction structure. The method is applicable to both traditional and multi-view video coding.

ACKNOWLEDGMENTS

This study was supported in part by the National Science Foundation of China under Grant No. 60802013 and the Zhejiang Provincial Natural Science Foundation of China under Grant No. Y106574.

REFERENCES

Bjontegaard, G., 2001. Calculation of average PSNR differences between RD curves. ITU-T SC16/Q6, 13th VCEG Meeting, Austin, Texas, USA., April 2001, Document VCEG-M33.

Bjontegaard, G., 2008. Improvements of the BD-PSNR model. ITU-T SG16/Q6, Berlin, Germany, Document VCEG-AI11.

Chen, C., Y.B. Liu, Q.H. Daiand and X.D. Liu, 2007. Performance modeling and evaluation of prediction structures in multi-view video coding. Proceedings of the International Conference on Multimedia and Expo, July 2-5, Beijing, pp: 1335-1338.

Girod, B., 1987. The efficiency of motion-compensating prediction for hybrid coding of video sequence. IEEE Trans. Selected Areas Commun., 5: 1140-1154.

Girod, B., 2000. Efficiency analysis of multihypothesis motion-compensated prediction for video coding. IEEE Trans. Image Process., 9: 173-183.

Kang, J.W., S.H. Cho, N.H. Hur, C.S. Kim and S.U. Lee, 2007. Graph theoretical optimization of prediction structure in multiview video coding. Proceedings of the International Conference on Image Processing, Sept. 16-19, San Antonio, Texas, USA., pp: 429-432.

Leontaris, A. and P.C. Cosman, 2007. Compression efficiency and delay tradeoffs for hierarchical B-pictures and pulsed-quality frames. IEEE Trans. Image Process., 16: 1726-1740.

Li, D.X., W. Zheng, X.H. Xie and M. Zhang, 2007. Optimizing inter-view prediction structure for multiview video coding with minimum spanning tree. Elect. Lett., 43: 1269-1271.

Liu, Y.B., Q.H. Dai and Z.X. You, 2007. Rate-prediction structure complexity analysis for multi-view video coding using hybrid genetic algorithms. Proc. SPIE, 6508: 650804-650804.

Merkle, P., A. Smolic, K. Muller and T. Wiegand, 2007. Efficient prediction structures for multiview video coding. IEEE Trans. Circuits Syst. Video Technol., 17: 1461-1473.

Muller, K., P. Merkle, H. Schwarz, T. Hinz, A. Smolic, T. Wiegand and T. Oelbaum, 2006. Multi-view video coding based on H.264/MPEG4-AVC using hierarchical B pictures. Proceedings of the Picture Coding Symposium, April 24-26, Beijing, pp: 385-390.

Park, P.K., K.J. Oh and Y.S. Ho, 2008. Efficient view-temporal prediction structures for multi-view video coding. IEEE Trans. Elect. Lett., 44: 102-103.

Reinhard, D., 2005. Graph Theory. Springer-Verlag, Heidelberg, New York, pp: 13-16.

Schwarz., H., D. Marpe and T. Wiegand, 2006. Analysis of hierarchical B pictures and MCTF. Proceedings of the IEEE International Conference on Multimedia and Expo, July 9-12, Toronto, Ont., pp: 1929-1932.

Steven, M.K., 1993. Fundamentals of Statistical Signal Processing: Estimation Theory. Prentice Hall PTR, Upper Saddle River, pp: 219-226.

Winken, M., H. Schwarz, D. Marpe and T. Wiegand, 2007. Joint optimization of transform coefficients for hierarchical b picture coding in H.264/AVC. Proceedings of the International Conference on Image Processing, Sept. 16-Oct. 19, San Antonio, TX, pp: 89-92.