

A BLOCK-MATCHING BASED INTRA FRAME PREDICTION FOR H.264/AVC

Jiheng Yang, Baocai Yin, Yanfeng Sun And Nan Zhang

Key Laboratory of Multimedia and Intelligent Software
Beijing University of Technology, Beijing
College of Computer Science and Technology
Beijing 100022, China
{jasonme@emails.bjut.edu.cn, yinbc@bjut.edu.cn}

ABSTRACT

In the latest video coding standard H.264, intra frame prediction is employed to reduce the blocks' spatial correlation. In this paper, a new algorithm is proposed to improve the performance of intra prediction of H.264/AVC. It changes mode 2 of H.264's standard prediction methods into a BMA (block-matching algorithm)-DC hybrid mode. Experiment results show that with the proposed algorithm, significant improvement of coding performance can be gained compared to the performance of H.264 algorithm.

1. INTRODUCTION

Intra prediction is employed to reduce the spatial correlation in video compression. The latest video coding standard H.264/AVC [1], which was finalized by Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, achieved great success in low bitrate coding of intra frame thanks to the precise prediction carried out by the new intra frame prediction algorithms. At present, H.264 CODECs mainly provide intra prediction for 4x4 blocks and 16x16 blocks. In this paper, however, we'll only focus on 4x4 block intra prediction.

For 16x16 luma blocks, there are four modes for intra prediction: DC mode, horizontal mode, vertical mode and plane mode in H.264.

For 4x4 luma blocks, there are nine modes for intra prediction, which are shown in Figure 1. The nine modes are labeled 0,1,3,4,5,6,7,8 while mode 2 is DC mode. As we can see from Figure 1, it uses 13 pixels, which have already been coded and are the nearest pixels to block X, for prediction.

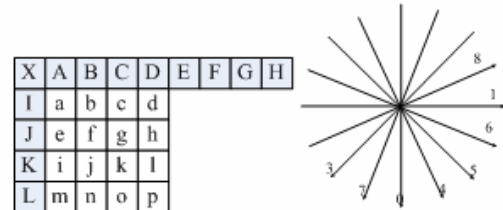


Figure 1 A 4x4 block with its boundary pixels and its 9 intra prediction modes

In recent years, tremendous work [2] [3] have been done to improve the accuracy of intra prediction. But most of them are pixel based intra prediction as shown in Figure 1. That is to say, the predicted block is generated using only the boundary pixels. Also a block-based intra prediction [4] is proposed. It uses the same technology as inter frame's motion compensation. But in this way, the position of the referenced block has to be coded into the bitstream too, and this results a lot more extra side information, which affects the performance significantly.

Besides all these achievements that have already been made, we can see that the performance of intra prediction still can be improved, which will lead to better compression efficiency for intra frames.

Block matching algorithm [5] is originally widely used in image restoration to recover missing blocks. The main concept behind this algorithm is that we can always find similar blocks for a block in one frame.

In this paper, we compiled the block-based intra prediction and block-matching algorithm into a novel BMA mode and introduced this concept into intra prediction. Then we generate a new hybrid mode with DC mode and BMA mode as a substitute to mode 2 of intra prediction in H.264/AVC and tested the idea with the reference software JM 9.0[6]. Results show that using the new mode, significant improvement can be gained compared to H.264 proposed intra prediction algorithm.

This paper is organized as follows. The proposed BMA-based intra prediction is explained in detail in Section 2. CODEC related issues are discussed in Section 3. Experimental results of the proposed method by comparing

to the standard method in H.264 are given in Section 4 and a short conclusion is given in Section 5.

2. BMA BASED INTRA PREDICTION

2.1 BMA Intra Prediction Mode

When a block is missing from a frame, the most common way to recover it is to use block-matching algorithm. We run a search in the current frame to find a matching block whose boundary is the closest to the missing block. Using the matching block's information, we can predict what the missing block could be and complete restoration process.

Intra prediction is a lot like recovering a missing block of a picture. We'll call a block which is currently being coded the target block for the rest of this paper. The target block is just like a block that is missing in the current frame. So if we can restore it with a matching block, we can achieve better prediction result, therefore better intra frame compression result can be gained. Under this consideration, we introduce block-matching algorithm (BMA) into intra prediction.

The problem is that we don't have all the boundary information of the target block because of the coding order defined by the CODEC, so we can only make use of part of the boundary to perform block matching. More specifically, for 4x4 block prediction in H.264, we have only nine pixels available, four at the upper side, four at the left side, and one at the left up side, the distribution of the boundary pixels is shown in Figure 2.

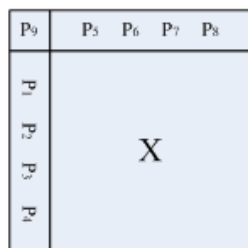


Figure 2: Target block X and its boundary pixels used for block matching

We use F to denote current frame, and use C to denote the area that have already been coded, for Figure 2, the prediction process is to find a block M for the target block X , and let:

$$M = BMA(X) \quad \text{where } M \subset C$$

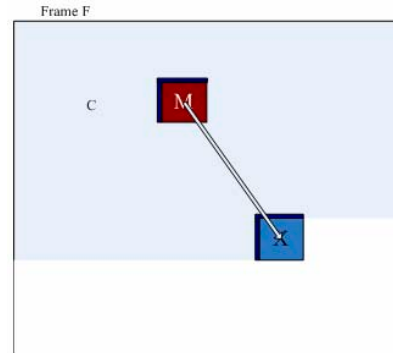


Figure 3 Block M, which belongs to coded area C, is the matching block of target block X, in frame F

2.2 Block Matching Criterion

We use a classical way to check if one block is matching to another. For Figure 2, we compute the Mean Square Error (MSE) between the boundary pixels of the target block X (whose gray level values are denoted with P_1 to P_9) and the boundary pixels of a candidate block M' (whose gray level values are denoted with P'_1 to P'_9). Mathematically depicted with:

$$MSE(X, M') = \frac{1}{N} \sum_{i=1}^N (P_i - P'_i)^2$$

Here, N equals to 9, which stands for the number of pixels we're using for block matching.

The candidate block M that has the lowest MSE is considered the matching block to the target block X , that is: $M = \arg \min_{M'} \{MSE(X, M')\}$ where $M' \subset C$

After we find the matching block, we use its reconstructed information for prediction directly.

2.3 Complexity Issues

Because we're using a block searching scheme to perform intra prediction instead of simply performing directional prediction, so it is obvious that this algorithm is more time consuming. But statistically, blocks which have shorter distance from one another in a frame tend to have stronger spatial correlation than the ones that have longer distance. So for a block, it's unnecessary to run a block search throughout C to locate a matching block. Therefore we choose a threshold T for the algorithm used as a search-range parameter for block matching. Empirically we choose $T = 24$ in our algorithm as the search range, which is a result of the consideration for both complexity and also quality. In this way, we have a lot less candidate blocks and reduce the complexity of the algorithm to a fairly low level.

We can also apply fast search algorithm to further speed up the process. A lot of fast algorithms (such as [7]) have been proposed to accelerate the speed of block

matching process. By applying these technologies into BMA based intra prediction, huge amount of computation will be saved with neglectable coding performance loss.

Because the block matching process for different candidate blocks are independent from one another, so another option is to apply parallel algorithms for block matching, which is one of the major trends of video coding.

3. CODEC RELATED ISSUES

3.1 Encoder side

We integrated the block-matching based intra frame prediction into H.264/AVC as one of the mode for intra prediction. For easy implementation and also saving extra bits to code the mode, we substitute mode 2 (which is DC mode in intra prediction for 4x4 blocks) with a hybrid mode, which is depicted in Figure 4. The encoder uses this hybrid mode along with the other 8 modes to perform prediction for 4x4 blocks, and choose the mode which produces lowest rate-distortion cost as the optimal mode for prediction. Since there is no extra mode used, so the proposed algorithm is compliant with the original standard of H.264/AVC.

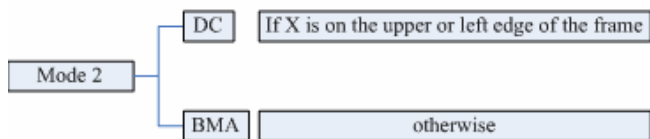


Figure 4 Definition of new hybrid mode 2. X stands for the block being predicted

The reason we're changing DC mode is for following 2 reasons.

- 1) DC mode performs well for smooth areas of the picture; BMA mode also performs well for smooth areas.
- 2) For blocks located at the texture area of the picture, or at the monotone background, etc. BMA have better chances to outperform DC mode.

But we keep using DC mode to predict the blocks located at the upper or left edge of the frame, because they only have 4 boundary pixels, which is insufficient for accurate block matching. Keep using DC mode in these areas can maintain the coding efficiency of H.264/AVC.

3.2 Decoder Side

At the decoder, if we know that a 4x4 block is coded using mode 2, then we check the whether the block is located at the upper or the left side of the frame. If it is, then decode it using DC mode; otherwise, decode it using BMA mode.

By decoding a block in BMA mode, the decoder also run a block search for a search range of T in the current frame, which is the same value as the encoder has set earlier, and use the matching block's information for prediction.

This way, we can assure the block can be decoded correctly. The same fast scheme applied at the encoder should also be performed at decoder not only to speed up the process, but also to synchronize with the encoder.

4. EXPERIMENT RESULTS

The block-matching based intra prediction algorithm proposed in this paper is implemented with the reference software of H.264/AVC JM 9.0. We set the coding environment as follows:

Source Sequence Format:	QCIF (176x144), 4:2:0
Frame Rate:	30Hz
Frame Skip:	NO
Coded Frame Number:	100
Coded Sequence Format:	IIII.....I
Use Hadamard:	Yes
Entropy Coding Algorithm:	CABAC
Values of QP	16, 20, 24, 28

Table 1 shows the result of the numeral comparison between the proposed method and standard intra prediction algorithms of H.264/AVC for sequence Foreman, Carphone, Bus and Hall Monitor. We use 'Gain' to denote average PSNR gain. We denote 'A' to the performance of JM 9.0 and 'B' to the performance of our proposed algorithm. We measure PSNR by db and Bitrate by kbps (kilo bits per second).

Figure 5 and Figure 6 shows the R-D curve comparison of the proposed method and the standard H.264/AVC intra prediction algorithms for sequence Foreman and Carphone.

From the results we can see that our algorithm can outperform the performance of H.264/AVC's algorithm by over 0.1-0.4db, which proves the effectiveness of the proposed method.

Table 1: PSNR and Bitrate performance of proposed intra prediction method

		QP	16	20	24	28	Gain
Foreman	A	PSNR	46.39	42.74	39.57	36.71	0.420
		Bitrate	2219.98	1579.98	1083.14	734.13	
	B	PSNR	46.41	42.77	39.6	36.73	
		Bitrate	2144.63	1513.54	1030.58	694.95	
Carphone	A	PSNR	47.15	44.16	41.18	38.3	0.204
		Bitrate	1638.53	1175.07	836.56	584.54	
	B	PSNR	47.15	44.13	41.16	38.28	
		Bitrate	1606.86	1146.62	812.75	566.18	
Bus	A	PSNR	46.4	42.48	38.74	35.26	0.104
		Bitrate	2980.93	2295.14	1705.46	1228.78	
	B	PSNR	46.41	42.49	38.76	35.27	
		Bitrate	2962.8	2279.97	1692.6	1218.55	
Hall_Mon itor	A	PSNR	46.46	43.5	40.82	38.14	0.100
		Bitrate	1841.99	1293.96	926.85	672.6	
	B	PSNR	46.47	43.51	40.84	38.17	
		Bitrate	1826.88	1280.9	917.03	665.32	

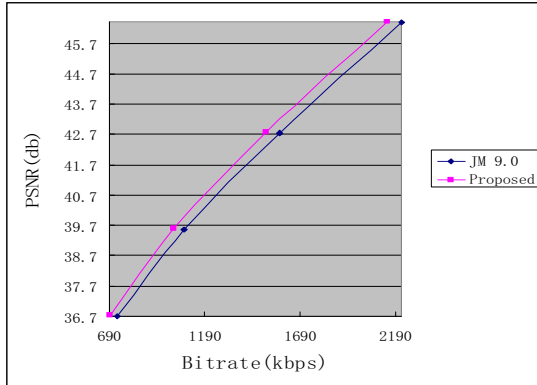


Figure 5: The R-D curve of Foreman.yuv

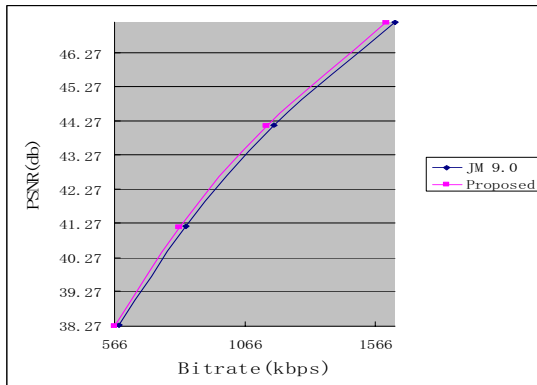


Figure 6: The R-D curve of Carphone.yuv



Figure 7: Original Y component for frame 89 of sequence Foreman



Figure 8: Y component of predicted image for frame 89 using H.264 standard algorithm (Left) and proposed algorithm (Right) of sequence Foreman (QP=20)

In Figure 7 and Figure 8, we show original and predicted picture of frame 89 for sequence Foreman as a sample to check the subjective and objective performance of our proposed algorithm and H.264 standard algorithm.

We can see that, the predicted frame using our proposed algorithm is subjectively closer to the original frame. Especially the edges of the background are smoother than using H.264's prediction algorithm. Objectively, the PSNR of the predicted picture of this frame using the H.264 standard algorithm is 27.058db, and the PSNR using the proposed algorithm is 27.745db, which is 0.687db higher than H.264. This can also prove that the proposed algorithm predicts more accurate than H.264 standard algorithm.

5. CONCLUSION

In this paper, we proposed a BMA-based intra prediction algorithm to improve the performance of intra prediction in H.264/AVC. It used a novel hybrid mode which compiles DC mode and BMA mode and tested its performance with JM 9.0. The results of the experiments show that this method could improve the overall performance of the compressed I frame by 0.1-0.4dB compared to H.264/AVC standard.

6. ACKNOWLEDGEMENT

This project is supported by National Natural Science Foundation of China (60572104、60533030) ; Beijing Natural Science Foundation (40410031)

7. REFERENCES

- [1] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 ISO/IEC 14496-10 AVC)", JVT-G050, March 2003.
- [2] Zhang Nan, Yin Baocai, Kong Dehui and Yue Wenying, "Spatial Prediction Based Intra-Coding", ICME2004, June 2004.
- [3] Zhu Gang, Guoping Li and Yun He, "The Intra Prediction Based on Sub-Block", ICSP2004, Aug, 2004.
- [4] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG "New Intra Prediction using Intra-Macroblock Motion Compensation", JVT-C151, 2002.
- [5] Zhou Wang, Yinglin Yu and David Zhang, "Best Neighborhood Matching: An Information Loss Restoration Technique for Block-Based Image Coding System", IEEE Transaction on Image Processing, July 1998.
- [6] H.264 JVT reference software, <http://bs.hhi.de/~suehring/tml/download>.
- [7] Wen Li, David Zhang, Zhiyong Liu, Xiangzhen Qiao, "Fast Block-Based Image Restoration Employing the Improved Best Neighborhood Matching Approach", IEEE Transaction On Systems, Man and Cybernetics-Part A: Systems and Humans, July 2005