MULTI-WEIGHTS INTRA PREDICTION WITH DOUBLE REFERENCE LINES

*Hailang Yang, Hongkui Wang, Yamei Chen, Junhui Liang, Li Yu**

School of Electron. Inf. & Commun., Huazhong Univ. of Sci. & Tech.

ABSTRACT

Conventional intra prediction applies the nearest reference line to predict the current block and improves the performance by adding more angular prediction modes. However, the performance of prediction near the bottom-right corner of the block suffers from non-negligible degradation due to the reference samples used for prediction locate at the above and left sides of the current block. Therefore, in this paper, to improve this situation, multi-weights intra prediction is proposed. Considering the coding complexity, the optimal two reference lines are picked out at first based on minimum distortion criterion. Following, the best weight combination is selected out from candidate weight combinations and is applied to predict the current block. Experimental results show that the proposed algorithm achieves 2.19% bitrate saving on average compared with the conventional intra prediction.

Index Terms— Video coding, multi-weights intra prediction, multiple reference lines

1. INTRODUCTION

The purpose of video coding is to remove the redundancy existing in the videos. Major redundancy in video comes from correlations among pixels, adjacent rows and adjacent frames in spatial and temporal domains. In order to reduce these correlations in video, the prevailing video coding standards [1–3] mainly adopt the hybrid coding framework based on predictive differential coding theory [1]. In this framework, intra prediction and inter prediction are separately designed for spatial and temporal redundancy. The motion-compensated prediction method is applied in inter prediction to find the best matching block to the current encoding block in reference images. As for intra prediction, multiple angular prediction modes are designed to handle blocks predictions with direction information.

Specifically, in intra prediction, the current block is predicted based on the single reference line by the limited prediction modes. The optimal mode is picked out based on the rate-distortion optimization (RDO). However, as for the complex block with multiple directions or with complicated texture, it is difficult to obtain high prediction performance due to the following reasons. Firstly, only the nearest reconstructed row or column is applied as the reference line. In some situation, the non-adjacent reconstructed pixels may obtain better prediction performance. Secondly, the number of intra prediction modes is limited. Finally, due to the reference samples for intra prediction locate at the above and left sides of the block, the correlation among samples generally decreases with the increase of the distance, which results in the performance of prediction near the bottom-right corner of the block is poorer.

To improve this situation, existing studies have explored various intra prediction methods [3–8]. By applying multiple reference lines, Li *et al.* [4] proposed the multi-line-based intra prediction including the full search scheme and the fast search scheme. The full search scheme effectively improves prediction performance while brings high coding complexity. The simplified algorithm of this method described in [5], named Multiple Reference Line (MRL), has been incorporated into the new generation coding standard under development (i.e., versatile video coding, VVC) [3]. Designing more intra prediction modes is another effective way to improve coding performance for videos containing multi-direction texture information. So, compared with H.264/AVC [1], high efficiency video coding (HEVC) [2] extends prediction modes from 9 to 35. Recently, VVC defined 67 modes in the intra prediction. But, these two methods have limited performance improvement for the bottom-right corner of the current block. Therefore, in order to improve the status, a series of new intra prediction methods have been proposed, such as DL-based (i.e., deep learning based) intra prediction [6,7], line-based intra prediction [8], etc. In DL-based intra prediction, the correlation between the current block and reference samples is analyzed by neural networks and is applied to predict the current block. In line-based intra prediction, the current block is divided into line-based sub-partitions and the preceding sub-partition is first coded and reconstructed as the reference samples for the later sub-partitions.

According to Li's experimental result [4], although the nearest reference line generally has the strongest statistical correlation with the current block, in some situation, the farther non-adjacent reference lines can still provide potential better prediction. Therefore, in this paper, multi-weights intra prediction with multiple reference lines is proposed. Considering the coding complexity, only two reference lines are picked out for multi-weights prediction. The first reference line is the nearest reference line and another one is selected

This work was supported by the NSFC (NO. 61871437 and 61702205)

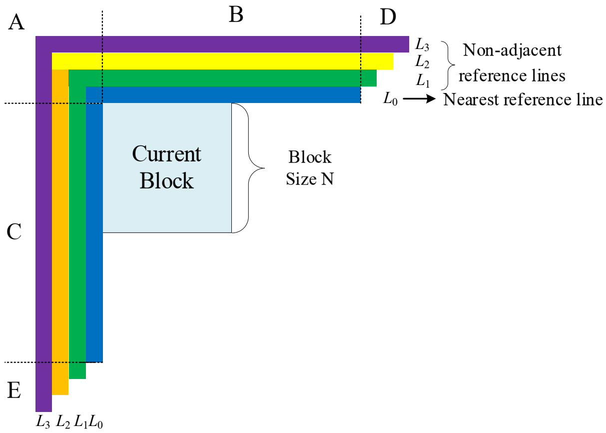

Fig. 1. The structure of multiple reference lines

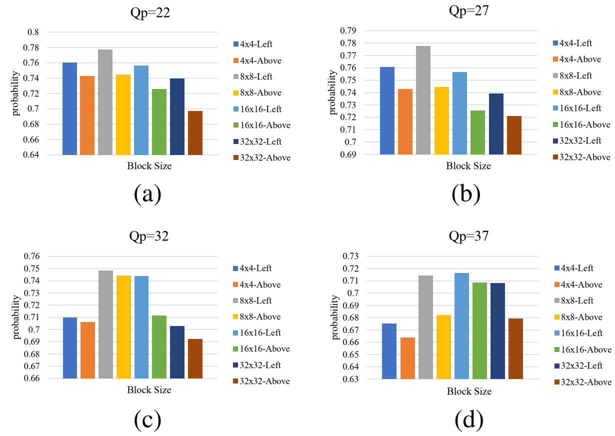

Fig. 2. The positive judgment probability of the optimal reference lines selection

from non-adjacent reference lines based on minimum distortion criterion. In order to obtain the better performance for the bottom-right corner, multiple candidate weights are picked out for both reference lines. Then, based on the RD cost minimum criterion of the total block, the optimal weight is selected and applied for final intra prediction. Therefore, the performance of prediction can be improved effectively.

2. OPTIMAL REFERENCE LINES SELECTION

Li's statistical result [4] indicates that the nearest reference line (L_0) has about 60% probability of being selected as the best reference line in the case of selecting four reference lines. Each of non-adjacent reference lines ($L_1 - L_3$) has about 10%-20% probability of being selected as the best reference line. Based on above statistical result, four reference lines at the above and the left sides of the current block are chosen as our candidate reference lines. The structure of the multiple reference lines is shown in Fig.1. As for the $N \times N$ block, the total reference samples from 4 reference lines are divided into 5 parts. The reference samples in the A-C parts are the reconstructed samples and the reference samples in the D and E parts are the extended samples from B and C.

Intra prediction performance can be effectively improved by adding more reference lines. However, the resulting coding complexity will also increase dramatically. According to the experiment results, adding three reference lines which increases encoding time by about 363% [4]. In this paper, we will choose only 2 reference lines from 4 candidate reference lines. Because of the strong correlation of the nearest reference line, this reference line is chosen for each block. Another reference line will be chosen from the remaining three reference lines. Usually, the optimal reference line can be picked out according to the RD cost, formulated by

$$L_{RDopt} = \underset{L_i \in \{L_1, L_2, L_3\}}{\arg\min} (D_{L_i} + \lambda \cdot R_{L_i}), \qquad (1)$$

here, D_{L_i} is the distortion between the reconstructed block using the reference lines (L_0, L_i) and the original block. R_{L_i} is the total bits. And λ is the Lagrangian multiplier used in the mode decision process. However, it is obvious that the best reference line selection process by RD optimization is much complex in the encoder. In this paper, by simulating the intra prediction process, the nearest reference line is predicted by non-adjacent reference lines. Another optimal reference line is picked out based on minimum distortion criterion. The specific process is as follows: a) Pick out the best M candidate intra modes as described in HEVC [2], in HM-16.20, M is set according to the block size; b) In the corresponding mode, the nearest reference line (L_0) is predicted by L_1, L_2, L_3, respectively; c) Another optimal reference line is picked out from L_1, L_2, L_3 based on minimum distortion criterion by

$$L_{MSEopt} = \underset{L_i \in \{L_1, L_2, L_3\}}{\arg\min} (\sum_{j=1}^{2N} (O_{L_0}(j) - P_{L_i}(j))^2). \qquad (2)$$

Here, $O_{L_0}(j)$ is the j-th sample in L_0 and $P_{L_i}(j)$ is the predictive sample of $O_{L_0}(j)$ by L_i reference line. To verify the performance of the method for obtaining the best reference line, the probability in the case of $L_{RDopt} = L_{MSEopt}$ for diffident block size (N) and different QPs (i.e., 22-37) is calculated. The statistical result is shown in Fig.2. The Left and Above in the label of Fig.2 represent the positive judgment probability of left and above reference lines of the current block. It can be found that the probability falls in the range of 65% to 80%. Due to samples in the non-adjacent reference lines are much similar, which results in the MSEs obtained by formula (2) are much similar and the probability is not high enough. Under these circumstances, the selected

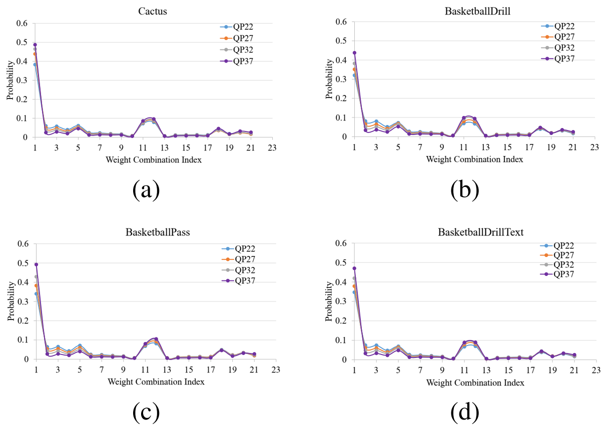

Fig. 3. Distribution of the optimal weight combination

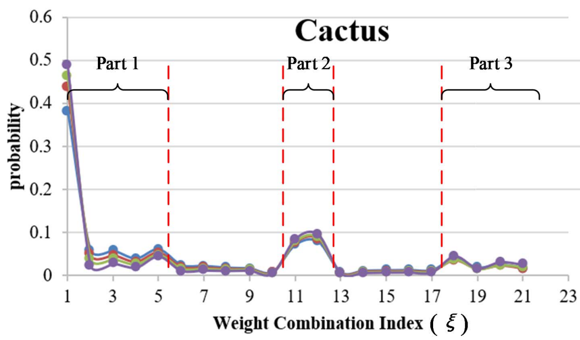

Fig. 4. Candidate weight combinations classification

reference line may not be optimal. Thus, multiple weight combinations for intra prediction are designed, which ensures the improvement of prediction performance.

3. MULTI-WEIGHTS INTRA PREDICTION

After obtaining two reference lines (i.e., L_0, L_{MSEopt}), multiple weight combinations are designed for the two reference lines to further improve the intra prediction performance of the whole block. The multiple weight combinations (Ω) is defined as follows:

$$\Omega = (\omega_1, \omega_2) = \{\overbrace{(1,0),(1,0.1),\cdots,(1,1),\cdots,(0,1)}^{21\ Weight\ Combinations}\}. \tag{3}$$

There are 21 weight combinations totally and the corresponding reference line aggregate can be expressed by

$$L(\omega_1, \omega_2) = (L_0 \times \omega_1 + L_{MSEopt} \times \omega_2) \times N(\omega_1, \omega_2), \tag{4}$$

here, N is the normalization aggregate and defined as

$$N(\omega_1, \omega_2) = \big(N(1), \cdots, N(21)\big), \tag{5}$$

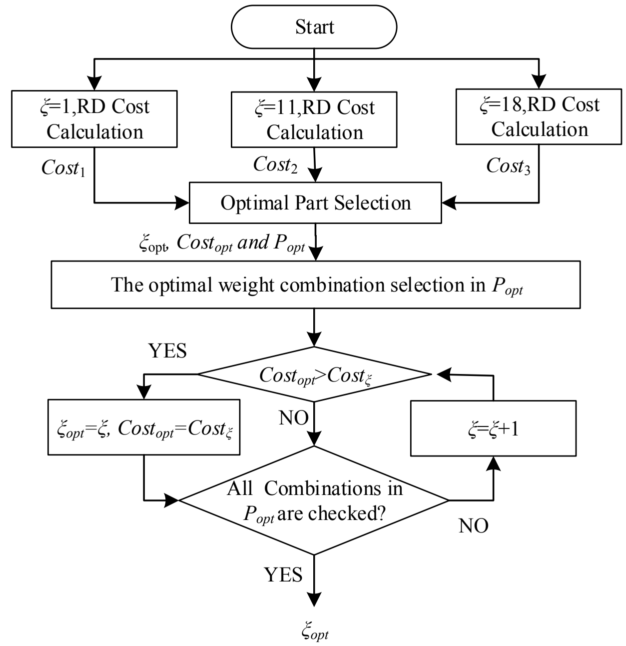

Fig. 5. Flowchart of optimal weight selection process

and $N(i) = 1/(\omega_1(i) + \omega_2(i))$. Yet, to choose the optimal weight combination from Ω will consume too much coding time. We calculate the probability that each weight combination is selected as the best combination for different sequences and 11 combinations with the highest probability are selected as candidate weight combinations. Statistical results is shown in Fig.3. It can be found that the distribution of optimal weight combinations for different sequences is similar.

As shown in Fig.4, the candidate weight combinations can be divided into three parts. The first part has 5 weight combinations (i.e., weight combination index (ξ) is from 1 to 5). The second part has 2 weight combinations (i.e., weight combination index is 11 and 12) and the last part has 4 weight combinations (i.e., weight combinations index is from 18 to 21). Therefore, the optimal weight combination selection process can be divided into 2 steps (i.e., optimal part selection and optimal weight combination selection). a) Best part selection: three weight combinations (i.e., ξ=1, 11 and 18) are picked out to represent three parts. The optimal part (P_{opt}) is selected out with $P_{opt} = \arg\min(D_{P_i} + \lambda \cdot R_{P_i})$ in the case of $P_i \in \{P_1, P_2, P_3\}$. b) Best weight combination selection: the optimal weight combination is selected out in the optimal part based on the RD optimization. The flowchart of optimal weight combination selection process is shown in Fig.5.

Finally, the current block is predicted with two optimal reference lines and the optimal weight combination. It should be noted that additional four bits are required to code the optimal combination index (11 combination indexes are redefined as 0-10) in the encoder. In the decoder, the index can

Table 1. The RD-rate result of the proposed method

	Sequence	Y(%)	U(%)	V(%)
4K	Tango	-2.42	1.62	1.12
	Drums	-1.87	-0.71	-0.86
	Campre	-2.65	-0.11	0.15
	ToddlerFuntain	-0.95	1.05	0.18
	CatRobot	-3.08	-0.79	-0.77
	TrafcFlow	-3.54	-0.11	-0.17
	DaylightRoad	-2.61	0.86	0.17
	RollerCoaster	-2.17	-0.15	-0.54
Average		-2.41	0.21	-0.09
Class A	Trafc	-1.93	-0.62	-0.60
	PeopleOnStreet	-2.78	-1.12	-0.84
	Nebuta	-0.80	-0.75	-0.54
	SteamLocomotive	-0.66	0.74	-1.75
Average		-1.54	-0.44	-0.93
Class B	Kimono	-0.83	-0.16	-0.23
	ParkScene	-1.50	-0.02	-0.11
	Cactus	-2.46	-0.55	-0.39
	BasketballDrive	-2.32	0.14	-0.38
	BQterrace	-2.31	-0.86	-0.78
Average		-1.88	-0.28	-0.37
Class C	BasketballDrill	-3.51	-1.19	-0.85
	BQMall	-2.75	-1.37	-1.26
	PartyScene	-2.45	-1.36	-1.38
	RaceHorsedC	-1.80	-0.60	-0.49
Average		-2.62	-1.13	-0.99
Class D	BasketballPass	-2.41	-1.23	-1.13
	BQSquare	-2.84	-1.47	-1.33
	BlowingBubbles	-2.47	-1.21	-1.12
	RaceHorses	-1.96	-1.26	-0.98
Average		-2.42	-1.30	-1.14
Class E	Johnny	-2.93	-0.66	-0.12
	FourPeople	-2.76	-0.33	-0.62
	KristenAndSara	-2.83	-0.22	-0.88
Average		-2.84	-0.40	-0.54
Class F	BasketballDrillText	-3.34	-1.19	-1.04
	ChinaSpeed	-0.42	0.37	0.33
	SlideEditing	-0.38	0.53	0.16
	SlideShow	-2.37	-0.68	-0.92
Average		-1.62	-0.24	-0.37
All Average		-2.19	-0.42	-0.56

be decoded normally and the optimal reference lines can be obtained by the same way described in Section 2.

4. EXPERIMENTAL RESULTS

In order to verify the performance of the proposed intra prediction algorithm, we implement it into the HEVC latest reference software (HM-16.20). The all-intra (AI) main configuration with QPs (i.e., 22 to 37) are tested. The test sequences for RD performance evaluation include the whole range of HEVC standard test sequences in common test conditions (CTC) [9] and another eight 4K sequences [10]. All sequences are coded 50 frames. BD-Rate [11] is used to measure the RD performance. The concrete experimental result of our method is shown in Table.1. As it can be seen, the average BD-rate saving of the simulations is 2.19% on luma component. Specifically, the average BD-rate saving of different class is in the range of 1.54% - 2.84%, which indicates that our algorithm is applicable for all sequences. The maximum bitrate saving is 3.54% for 'TrafficFlow' in 4K. As shown in Fig.6, The RD curves of 'FourPeople' is taken as example for perference comparison. It can be found that the proposed algorithm

Table 2. Comparison with the state-of -the-art proposal

Performance	Lis algorithm [4]		Proposed algorithm
	Full Search	Fast Search	
BD-rate(Y)-4K	-2.7%	-2.3%	-2.41%
BD-rate(Y)-CTC	-2.2%	-1.9%	-2.16%
BD-rate(Y)-All	-2.3%	-2.0%	-2.19%
Encoding Time	463%	212%	155%
Decoding Time	113%	110%	108%

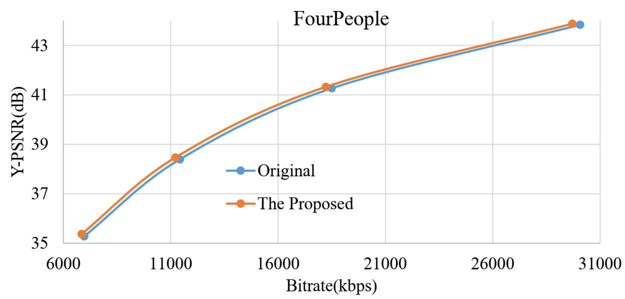

Fig. 6. RD curves of the FourPerple sequence

is generally better than the traditional prediction algorithm.

To further verify the performance and complexity of the proposed algorithm, we compare it with multiple-line-based intra prediction algorithm [4]. In [4], the full search scheme and the fast search scheme were proposed. The comparison result is shown in Table.2. As shown in this table, from a performance point of view, our algorithm is better than Li's fast search scheme, and is very close to Li's full search scheme in the Y component. Besides, in the view of the complexity, the encoding time of our algorithm is much smaller than Li's fast scheme and is only 1/3 of his full search scheme. This is for several reasons. First, our algorithm is only applied in Y component. Second, in optimal weights selection, it is not necessary to traverse all weight combinations. Third, we only use two reference lines and the optimal reference lines are chose based on minimum distortion criterion. In the decoder, compared with HM-16.20, our decoder adds optimal reference lines selection and weighted prediction, which increases decoding time by 8% totally. In summary, the proposed effectively improves the performance of intra prediction with lower coding complexity.

5. CONCLUSIONS

To further improve the performance of intra prediction, multi-weights intra prediction is proposed in this paper. Two optimal reference lines are picked out based on minimum distortion criterion. Multiple candidate weight combinations are designed for these reference lines and applied to predict the current block. Experimental results show that the proposed algorithm achieves 2.19% bitrate saving on average with lower complexity increase.

6. REFERENCES

[1] T. Wiegand, G. J. Sullivan, G. Bjontegaard, 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, 2003.

[2] G. J. Sullivan, J. Ohm, W. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, 2013.

[3] B. Bross, J. Chen, and S. Liu, "Versatile Video Coding (Draft 2)," *JVET-K1001, Ljubljana, SI*, Jul. 2018.

[4] J. Li, B. Li, J. Xu, and R. Xiong, "Efficient multiple line-based intra prediction for HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 4, pp. 947–957, 2018.

[5] B. Bross, H. Schwarz, D. Marp, and T. Wiegand, "CE3:Multiple reference line intra prediction," *JVET-K0051, Ljubljana, SI*, Jul. 2018.

[6] J. Li, B. Li, J. Xu, and R. Xiong, "Intra prediction using fully connected network for video coding," in *IEEE International Conference on Image Processing*, 2018.

[7] W. Cui, T. Zhang, S. Zhang, F. Jiang, W. Zuo, Z. Wan, and D. Zhao, "Convolutional neural networks based intra prediction for HEVC," in *Data Compression Conference*, 2017.

[8] S. De-Luxán-Hernández, H. Schwarz, D. Marpe, and T. Wiegand, "Line-based intra prediction for next-generation video coding," in *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018, pp. 221–225.

[9] F. Bossen, "Common test conditions and software reference cofigurations," *JCTVC-L1100, Geneva, Switzerland*, Jan. 2013.

[10] K. Suehring and X. Li, "Jvet common test conditions and software reference cofigurations," *JVET-B1010, San Deigo, USA*, Feb. 2016.

[11] G. Bjontegaard, "Improvements of the (BD-PSNR model," *VCEG-AI11*, Jul. 2008.