GEOMETRIC DERIVED MOTION VECTOR FOR MOTION PREDICTION IN BLOCK-BASED VIDEO CODING

Yucheng Sun, Lu Yu

Institute of Information and Communication Engineering, Zhejiang University
Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking
Hangzhou, China
Email: {sunyucheng, yul}@zju.edu.cn

ABSTRACT

Motion compensation plays an important role in inter picture prediction but the representation of motion vector (MV) is highly redundant. The existing motion model only selects the MVs of neighbor blocks as candidates even if they are very different from the MV of current block, which results in additional RD cost. In this paper, a geometric motion model is proposed to derive a geometric derived motion vector (GDMV) when an object has a relative movement that is close to or away from the camera. The proposed GDMV is integrated to HM 16.0 to further improve the motion prediction in HEVC. In order to ensure the accuracy of the motion vector predictor, a refining process is adopted before integrating the model into HEVC motion vector prediction. Experimental results show that compared to HM 16.0, the proposed GDMV can bring 0.14%~1.13% BD bitrate saving in low delay P main test.

Index Terms**—**Motion prediction, High Efficiency Video Coding (HEVC), geometric, motion model

1. INTRODUCTION

In current video coding standards, the primary method for inter prediction is block-based motion compensation. Motion compensation exploits the fact that much of the information that represents one frame might be associated to the information used in other frames of the same sequence. The prediction operation then basically consists of copying the referenced block to the block in the current frame. When the quantization parameter (QP) is high and the compressed picture quality is low, motion vector (MV) coding becomes a bottleneck for the performance promotion of block-based video coding. It is because that the resulting motion vector representations are highly redundant.

This work was supported in part by the National Science Foundation of China (Grant No. 61431015), the National High Technology Research and Development Program of China (863 Program) (Grant No.2015AA015903) and the Fundamental Research Funds for the Central Universities 2017XZ ZX009-03.

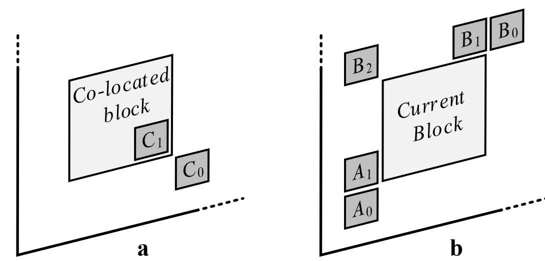


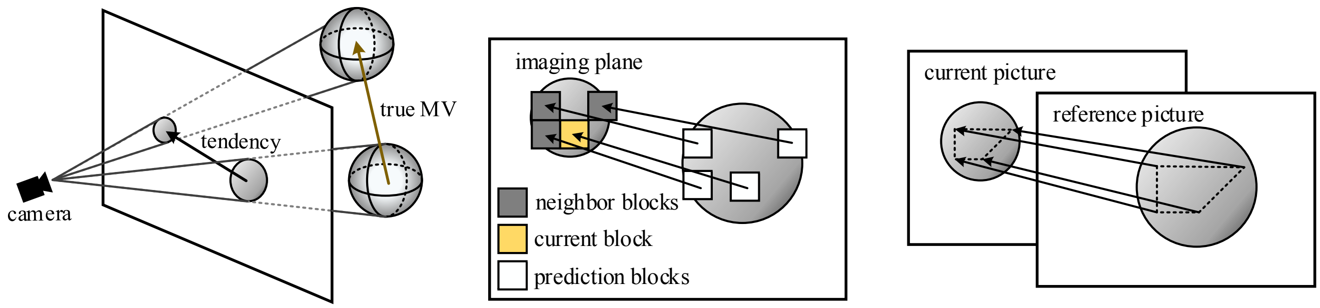
Fig.1: AMVP and merge candidates. (a) Temporal. (b) Spatial.

The High Efficiency Video Coding (HEVC) standard is the most recent joint video project of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG), working together in a partnership known as the Joint Collaborative Team on Video Coding (JCT-VC) [1]. In HEVC, advanced motion vector prediction (AMVP) and merge mode are adopted to reduce the cost of MV representation [2].

AMVP [3] derives several most probable MV candidates based on data from adjacent prediction blocks and reference picture. Merge mode [4] for MV coding allows the inheritance of MVs from temporally or spatially neighboring prediction blocks. Moreover, improved skipped and direct motion reference are also specified compared to H.264/AVC.

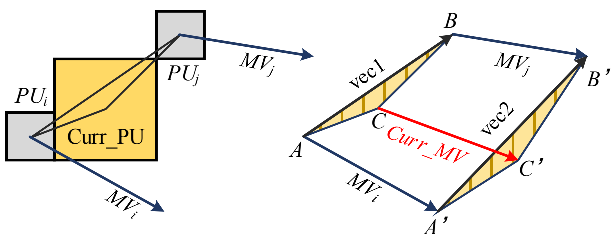
The locations of the spatial candidate blocks and temporal candidate blocks are the same for both AMVP and inter-prediction block merging. As shown in Fig.1 (a), the blocks to the bottom right of current block and at the center of the current block in the reference picture/collocated picture have been determined to be the most suitable ones to provide a good temporal motion vector prediction. The spatial candidates are derived from five spatially neighboring blocks, as shown in Fig.1 (b).

AMVP and merge candidates list pick the MV of reconstructed neighbor blocks. An index indicates which candidate in AMVP/merge candidate list is selected as prediction MV even though the candidate may be very different from the MV of current block. There are two main reasons. Firstly, encoder may choose the suboptimum MV



a b c

Fig.2: Object has a relative movement away from the camera. (a) Camera capture. (b)Imaging plane illustration. (c) Inter prediction.



a b

Fig.3: Derivation of the new motion vector.

with lower RD cost instead of the best one which may cause higher RD cost with higher mode representation. In [5], [6], the applicable motion parameters was refined at decoder side to further improve the coding performance. Secondly, objects may have a relative movement that is close to or away from camera. In order to derive a more accurate motion vector predictor (MVP), a geometric derived motion vector (GDMV) is proposed to improve video coding efficiency.

The rest of the paper is organized as follows. Section 2 presents the geometric motion model and how it is integrated into HEVC. Experiment results are given in section 3. Section 4 concludes the paper.

2. THE PROPOSED METHOD

2.1. Geometric motion model

In HEVC, the existing motion model only selects the MV of neighbor blocks as motion vector predictor (MVP). However, as shown in Fig.2-(a), when an object moves away from the camera, the shape of the object become smaller in imaging plane. It can be seen in Fig.2-(b) that none of the MVs of neighbor blocks are the same as the MV of the current block, which leads to worse coding efficiency. When the object has a relative movement that is close to or away from the camera, the MVs of neighbor block tend to spread or gather together. Considering this phenomenon, a geometric model is proposed based on Fig.3 to derive a more accurate MVP. Fig.3 shows the derivation of the geometric derived motion vector (GDMV). We assume the three blocks belong to one object. MV_i, MV_j represent the MV of the neighbor blocks PU_i, PU_j. MV_i is different from MV_j even if they belong to the same object. In Fig.3-(b) A, B is the center of PU_i, PU_j. C is the center of the current block $Curr\_PU$. $Curr\_MV$ is the MVP we want to derive. We assume the object in imaging plane just changing the size while remaining the topological structure, which can be written as

$$\Delta ABC \backsim \Delta A'B'C'. \tag{1}$$

According to (1), triangle ABC is a similar triangle of triangle $A'B'C'$. Based on this assumption, we can derive the unknown point C' to get the expectation MVP i.e. CC'. The calculation formula is given by

$$\begin{cases} \cos\theta_1 = \cos\theta_1' \\ \cos\theta_2 = \cos\theta_2' \end{cases}, \tag{2}$$

where θ_1 is the angle between the segment AB and the segment $A'B'$, θ_1' is the angle between the segment AC and the segment $A'C'$, θ_2 is the angle between the segment BA and the segment AC, θ_2' is the angle between the segment $B'A'$ and the segment $A'C'$.

If (a_1, b_1) is the coordinates of the vector AB, (a_2, b_2) is the coordinates of the vector $A'B'$, (c_1, d_1) is the coordinates of the vector AC, (c_2, d_2) is the coordinates of the vector $A'C'$, (e_1, f_1) is the coordinates of the vector CB, α is the similarity ratio, β_1 is the ratio between the segment AC and the line segment AB, β_2 is the ratio between the segment BC and the line segment AB. Equation (2) can be written as

$$\begin{cases} \beta_1^2(a_1a_2 + b_1b_2) = c_1c_2 + d_1d_2 \\ \alpha^2(a_1c_1 + b_1d_1) = a_2c_2 + b_2d_2 \end{cases}, \tag{3}$$

where

$$\alpha = \frac{\sqrt{a_2^2 + b_2^2}}{\sqrt{a_1^2 + b_1^2}}, \beta_1 = \frac{\sqrt{c_1^2 + d_1^2}}{\sqrt{a_1^2 + b_1^2}}, \beta_2 = \frac{\sqrt{e_1^2 + f_1^2}}{\sqrt{a_1^2 + b_1^2}}. \tag{4}$$

If (x, y) is the coordinate of point C, (x', y') is the coordinate of point C', $(x' - x, y' - y)$ is the MV we want to derive, where x', y' are unknown. The definition of vector dot product

$$\begin{cases} g(m,n) = mx_1' + ny_1' \\ h(m,n) = mx_2' + ny_2' \end{cases}. \tag{5}$$

where (x_1', y_1') is the coordinate of point A', (x_2', y_2') is the coordinate of point B'.

Integrate (5) into (3), after calculating the binary linear equation, we get the final GDMV.

$$\begin{cases} mv_x = x' - x = (d_1(g(a_2,b_2) + \alpha^2(a_1c_1 + b_1d_1)) \\ -b_2(g(c_1,d_1) + \beta_1^2(a_1a_2 + b_1b_2)))/(a_2d_1 - b_2c_1) - x \\ mv_y = y' - y = (c_1(g(a_2,b_2) + \alpha^2(a_1c_1 + b_1d_1)) \\ -a_2(g(c_1,d_1) + \beta_1^2(a_1a_2 + b_1b_2)))/(b_2c_1 - a_2d_1) - y \end{cases}, \quad (6)$$

When $a_2d_1 = b_2c_1$, AC is parallel to $A'B'$, GMVD should be derived as

$$\begin{cases} mv_x = x' - x = (f_1(h(a_2,b_2) + \alpha^2(a_1e_1 + b_1f_1)) \\ -b_2(g(e_1,f_1) + \beta_2^2(a_1a_2 + b_1b_2)))/(a_2f_1 - b_2e_1) - x \\ mv_y = y' - y = (e_1(h(a_2,b_2) + \alpha^2(a_1e_1 + b_1f_1)) \\ -a_2(g(e_1,f_1) + \beta_2^2(a_1a_2 + b_1b_2)))/(b_2e_1 - a_2f_1) - y \end{cases}, \quad (7)$$

where $a_2f_1 \neq b_2e_1$ under the measurements in (10).

Besides, when two neighbor blocks have a specific positional relationship, the calculation process will be greatly simplified. When the two neighbor blocks come from the left region of the current block, a_1 is equal to zero. Then (3) can be simplified as

$$\begin{cases} \beta^2 b_1b_2 = c_1c_2 + d_1d_2 \\ \alpha^2 b_1d_1 = a_2c_2 + b_2d_2 \end{cases}. \quad (8)$$

2.2. Integration the proposed GDMV into HEVC

2.2.1. Refining motion vector of neighbor region

During video coding, the neighbor blocks may choose the suboptimum MV with lower RD cost instead of the optimum one with higher RD cost. However, the suboptimum MV is not accurate enough to be used for the prediction of current block. In order to ensure the accuracy of the MVs of neighbor blocks, a refining process is adopted before integrating the model into HEVC motion vector prediction.

In refining process, we utilize the reference frame to refine the MVs in AMVP and merge candidates list. In order to reduce the computational complexity, the MVs in AMVP and merge candidate list is utilized as the starting position of the motion search. Since template matching (TM) scheme [7], [8] does not work in our motion model because the shape of the object changes over frames, we choose small block size from the spatial blocks in Fig.1-(b) to operate searching process to assure search accuracy. When the neighbor PU size is larger than the specific block size, pixels closer to the current PU are chosen. If the neighbor PU size is smaller than the specific block size, the refining process will be early terminated. Furthermore, the sum of absolute differences (SAD) of the samples is utilized as cost measurement for motion search.

2.2.2. Selecting the MV pair and deriving GDMV

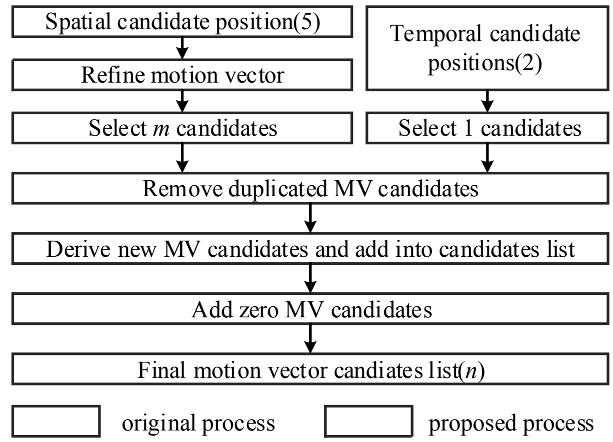


Fig.4: Illustration of candidate list construction, in AMVP, $m = 2$, $n = 2$. In merge mode, $m = 4$, $n = 5$.

In our proposed motion model the three blocks in Fig.3-(a) must belong to the same object. So we need to select the suitable neighboring MV pair to derive the MVP for current block. Here two criteria are followed when choosing the neighboring MV pair. First, considering the fact that the size changing degree is not arbitrary, similarity ratio α in (4) should be taken into consideration. The other one should be considered is the degree of rotation of the directed segment AB and $A'B'$, which is defined as

$$\tau = \frac{a_1a_2 + b_1b_2}{\sqrt{a_1^2 + b_1^2}\sqrt{a_2^2 + b_2^2}}. \quad (9)$$

Finally (10) determine the possibility this two blocks belongs to the same object, which empirically leads the best coding efficiency. Besides, the MV pair must come from the same reference frame.

$$\begin{cases} 0.7 \le \alpha \le 1.3 \\ 0.94 \le \tau \end{cases}. \quad (10)$$

Then MV pair is integrated into the proposed motion model to derive a new motion vector.

2.2.3. Adding the GDMV into AMVP and merge list

In AMVP, each PU is associated with a single scaled MV, while in merge mode, each PU is associated with a single MV without scaling and one corresponding reference frame from this list. As shown in Fig.4, the similarities and differences between AMVP and merge candidates list construction are listed as follows.

1) AMVP has up to two spatial candidates while merge mode has four spatial merge candidates which can be derived from five spatial neighboring blocks shown in Fig.1-(a).

2) AMVP and merge candidates lists have one temporal candidates derived from two temporal, co-located blocks.

Table. 1: BD bit rate performance of the proposed algorithm, compared to HM 16.0 low delay P configuration

Class	Sequence	BD-rate		
		Y	U	V
Class A	Traffic	**-1.0%**	-1.1%	-1.1%
	PeopleOnStreet	-0.4%	-0.4%	-0.6%
	Nebuta	-0.1%	-0.6%	-0.4%
	SteamLocomotive	-0.5%	0.2%	-2.0%
Class B	Kimono	-0.5%	-0.4%	-0.4%
	ParkScene	**-0.9%**	-0.7%	-0.9%
	Cactus	-0.7%	-0.6%	-0.5%
	BasketballDrive	-0.3%	-0.3%	-0.2%
	BQTerrace	**-1.1%**	-0.8%	-0.3%
Class C	BasketballDrill	-0.6%	-1.3%	-0.6%
	BQMall	-0.4%	-0.6%	-0.5%
	PartyScene	-0.4%	-0.3%	0.0%
	RaceHorses	-0.3%	-0.5%	-0.2%
Class D	BasketballPass	-0.2%	-0.3%	-0.4%
	BQSquare	-0.4%	-0.1%	-0.8%
	BlowingBubbles	-0.4%	-0.6%	-0.6%
	RaceHorses	-0.3%	-0.7%	-0.2%
Class E	FourPeople	-0.5%	-0.2%	-0.3%
	Johnny	**-1.1%**	-1.2%	-1.5%
	KristenAndSara	-0.5%	-1.2%	-0.6%
Class F	BasketballDrillText	**-0.8%**	-1.1%	-0.6%
	ChinaSpeed	-0.6%	-0.5%	-0.6%
	SlideEditing	-0.1%	0.1%	0.1%
	SlideShow	-0.3%	-0.5%	-1.8%
Average		-0.5%	-0.6%	-0.6%

3) AMVP and merge mode will add zero motion vectors when the candidates list is insufficient.

In AMVP, scaling process causing the two spatial candidates in candidate list always have the same reference frame. In order to reduce computational complexity, we pick one spatial candidate from the left block and another from the above block to derive a GDMV when they meet the measurements in (10). In merge mode, there are relatively fewer MV pairs with the same reference frame. Thus, we check all combinations of available spatial candidates to derive a new candidates and select the one with smallest τ value.

After deriving the GDMV, one of the two original MVs in AMVP with smaller MV difference compared to GDMV is eliminated and the GDMV is added after the remaining one. While in merge mode, if candidate list is insufficient before adding zero motion vector, the derived MV will be added into the list. Otherwise the one of the two original MVs with smaller MV difference compared to GDMV is eliminated and the GDMV is added into the end of the candidates list.

3. EXPERIMENTAL RESULTS

We integrate the proposed GDMV into HEVC reference platform HM16.0 [9] and conduct extensive experiments to verify its performance.

3.1. Test conditions

The test sequences are from class A to class F specified in the HEVC common test conditions (CTC) [10]. The details of the test conditions are listed as follows:

- The anchor is the original HM 16.0 software.
- Experimental results are evaluated in terms of BD-rate, which is calculated by four rate and PSNR points generated with four QP values {22, 27, 32, 37}[11].
- Experiments are conducted under low delay P.

3.2. Experimental results

The coding performance of proposed geometric derived motion prediction is presented in Table. 1. Results show that compared to HM 16.0, the proposed method can bring 0.14%~1.13% BD bitrate saving in low delay P main test, which achieves 0.5%, 0.6% and 0.6% BD rate reduction for one luminance component and two chrominance components on average, respectively.

4. CONCLUSION

In this paper, a geometric motion model is presented to derive a more suitable MVP called GDMV. By refining the MVs of neighbor blocks of the current PU, and choosing a neighbor MV pair for the proposed model, GDMV can be derived. Then the coding efficiency of AMVP, merge mode and skip mode can be improved with GDMV. The experimental results show the BD rate saving is up to 1.13% compared to HM 16.0. The proposed motion model is only integrated into P frame. In the proposed motion model, object has only a slight rotation with large τ value. In future work, the combination of affine motion model will be explored when the parameter τ is small. And an extension of the proposed motion model to Bi - Prediction will be investigated.

5. REFERENCES

[1] B. Bross, W.-J. Han, G. J. Sullivan, J.-R. Ohm, and T. Wiegand, "High Efficiency Video Coding (HEVC) text specification draft 9", *Document JCT-VC*, *K1003*, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), 2012.

[2] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard", *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, 2012.

[3] J.-L. Lin, Y.-P. Tsai, Y.-W. Huang, and S, Lei, "Improved advanced motion vector prediction", *Document JCT-VC, D125*, Jan, 2011.

[4] P. Helle, K. Ugar, "Block merging for quandtree-based partitioning in HEVC", *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 22, no. 12, pp. 1765-1777, 2012.

[5] S. Kamp, M. Wien, "Decoder-Side Motion Vector Derivation for Block-Based Video Coding", *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 22, no. 12, pp. 1732-1745, 2012.

[6] S. Kamp, M. Wien, "Decoder-side motion vector derivation for hybrid video inter coding", *2010 IEEE International Conference on Multimedia and Expo (ICME)*, vol. 26, no. 2, pp. 1277-1280, 2010.

[7] K. Sugimoto, M. Kobayashi, Y. Suzuki, S. Kato, and C. S. Boon, "Inter frame coding with template matching prediction", *IEEE Int. Conf. Image Process.*, pp. 465-468, 2004.

[8] Y. Suzuki, C. S. Boon, and S. Kato, "Block-based reduced resolution inter frame coding with template matching prediction", *IEEE Int. Conf. Image Process.*, pp. 1701-1704, 2006.

[9] HM Reference Software for HEVC. Available at:https://hevc.hhi.fraunhofer.de/

[10] F. Bossen, H. Common, "Test condition and software reference configurations", *Document JCT-VC*, L1100, Jan, 2013.

[11] G. Bjontegaard, "Calculation of average psnr differences between rdcurves", *Document VCEG*, *M33*. Austin, Texas, USA, 2001.