2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM)

# Enhanced Motion Vector Prediction for Video Coding

Suhong Wang
*Institute of Digital Media*
*Peking University*
Beijing, China
suhong.wang@pku.edu.cn

Zhao Wang
*Institute of Digital Media*
*Peking University*
Beijing, China
zhaowang@pku.edu.cn

Falei Luo
*Institute of Computing Technology*
*Chinese Academy of Sciences*
Beijing, China
falei.luo@vipl.ict.ac.cn

Shanshe Wang
*Institute of Digital Media*
*Peking University*
Beijing, China
sswang@pku.edu.cn

Siwei Ma
*Institute of Digital Media*
*Peking University*
Beijing, China
swma@pku.edu.cn

Wen Gao
*Institute of Digital Media*
*Peking University*
Beijing, China
wgao@pku.edu.cn

***Abstract*—Motion vector prediction (MVP) plays a crucial role in reducing the rate cost for coding motion vector (MV). However, the relative limited construction of MVP list restricts the potential efficiency. In this paper, an enhanced MVP scheme is proposed. In particular, we first modify the insertion order of MVP candidates according to the statistics. Then, the scaling process during MVP construction is modified to extend the application range. Moreover, an expanded-area motion vector prediction (EMVP) approach is adopted to further utilize the spatial correlation of local motion field. Simulation results shows that the proposed scheme can achieve better prediction for coding motion information.**

***Index Terms*—MVP, High Efficiency Video Coding, expanded-area motion vector prediction, video coding**

## I. Introduction

Motion Vector Prediction (MVP) plays an important role in the previous video coding standards. The earliest standard H.261 [1] uses a simple predictor, where the predicted motion vector is given by the motion vector of the macroblock to the left of the current coding block. The H.263 motion compensation uses a more complex motion vector predictor [3]. The predicted motion vector is the median value of three MVP candidates, which are the motion vectors of the macroblock to the left, above and above-right relative to the current one. Later, in H.265/HEVC [2], the motion vector prediction was enhanced with advanced motion vector prediction (AMVP) based on motion vector competition [2] [4].

In the latest future video coding (FVC) development [5], this AMVP mechanism still plays a key role. From the description of AMVP, it is known that if the neighbor blocks of the current block are intra-mode coded, the blocks are excluded from MVP candidate list and thus no motion information from these kinds of blocks can be utilized. In this case, zero MVs will be added to the MVP list if the number of the candidates is less than two. It can be seen that this mechanism may be not very precise and some other useful motion information of neighbor inter-mode blocks is negligible. The surrounding MV information of inter-mode coded blocks is not fully utilized and the MVP accuracy of current block will be influenced. In addition, since numerous changes have been made during the development of FVC, the order of MVP lists and the mechanism of scaled MVP should also be taken into further consideration.

In this paper, we propose an optimized MVP scheme for FVC. To make full use of available MV information of surrounding blocks, some MVs of surrounding inter-mode blocks, which are named as Expanded-area Motion Vector Predictor (EMVP), are utilized instead of zero MVs or Temporal Motion Vector Predictor (TMVP) by expanding the area of MVP derivation. Besides, according to statistical tests, the order of MVP candidates lists and the applying conditions of scaling have been improved in this paper as well.

The rest of this paper is organized as follows. In Section II, the principle of AMVP will be introduced specifically and the problem statement is also given in this section. Section III presents the detailed description of the proposed optimized MVP scheme. The experimental results and performance comparison are given in Section IV to show the efficiency of our proposed method. Finally, the conclusion is presented in Section V.

## II. Overview of AMVP and Problem Statement

### A. Overview of AMVP

As shown in Fig. 1, the AMVP scheme of HEVC will choose two spatial MVPs and one temporal MVP as the candidate MVPs. For spatial motion vector candidate derivation, two motion vector candidates are derived based on motion vectors of each PU located in five different positions. The order of derivation for left side of the current PU is set as follow:{$A_0$, $A_1$, scaled $A_0$, scaled $A_1$} and the order of derivation for above side of the current PU is set as {$B_0$, $B_1$, $B_2$, scaled $B_0$, scaled $B_1$, scaled $B_2$}.

The derivation is processed in a two pass approach. In the first pass, it is checked whether any of the candidate blocks contain a reference index that is equal to the reference index of the current block. The first motion vector found will be taken as candidate A. When all reference indices from $A_0$ and $A_1$ are pointing to a different reference picture than the reference index of the current block, the MVs need to be scaled according to the temporal distances between the candidate reference picture and the current reference picture, which is shown in Fig. 2.

Then the temporal motion vector candidates search one motion vector from two different co-located candidates, i.e. ($C_0$, $C_1$). The first position is $C_0$, which is at the bottom-right position of the current PU. When position $C_0$ is unavailable or outside of the current coding tree unit (CTU), position $C_1$ is used. When the number of searched candidates is greater than two, only the first two candidates will be preserved. Otherwise, the zero motion vector is added to the candidates list, if the number of searched motion vector is less than two.

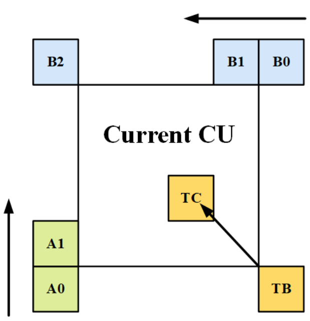



Fig. 1. AMVP mechanism in HEVC.

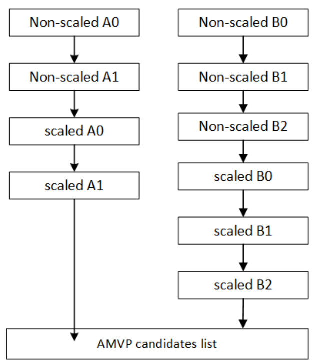



Fig. 2. The flowchart of AMVP mechanism.

### *B. Problem statement*

Motion vector difference (MVD) is the vector subtracted by MVP from MV, which can be eliminate applied to the correlation between motion vectors of neighboring blocks. The calculation of both MVD components is shown in (1) and (2). From these equations, we can know that, the more accurate MVP is performed, the fewer MVD will be coded and the coding efficiency will be improved.

$$MVD_x = \Delta x - MVP_x \tag{1}$$

$$MVD_y = \Delta y - MVP_y \tag{2}$$

To learn about the performance of MVP mechanism more specifically, we calculate the proportion of MVD in video bitstream with configuration of Lowdelay-P in one second. From Table.1, we found that MVD still takes up a large proportion in bitstream, average of which is up to 11%. In this condition, the performance of MVP mechanism still have much room for improvement according to the statistical data.

When the number of MVP candidates in the list is smaller than two, zero motion vector (0, 0) will be added into the list till the list is full. As mentioned in [6], the percentage of intra-mode macroblocks is up to 12% and 9% in sequence "stefan" and "foreman", separately. According to this, D. Liu proposes an improved MVP scheme for H.264 by exploiting more neighbor motion information of the first layer blocks around the current block in [5]. While in JEM 7.0, the percentage of intra-mode blocks is still up to 12% on average and MVP accuracy of nearly 48% blocks will be influenced, as shown in Table. 1. But this "add-zero" scheme stays unchange in JEM 7.0, which means that great amounts of coding blocks MVPs are not accurate enough.

TABLE I
PROPORTION OF MVD IN BIT STREAM(QP=47) AND THE PERCENTAGE OF INTRA-MODE BLOCKS(QP=47)

| Sequences | Resolution | MVD Proportion | I-mode blocks |
|---|---|---|---|
| BasketballPass | 416x240 | 14% | 9% |
| RaceHorses | 416x240 | 15% | 23% |
| BQMall | 832x480 | 9% | 13% |
| PartyScene | 832x480 | 9% | 15% |
| BQTerrace | 1920x1080 | 11% | 4% |
| BasketballDrive | 1920x1080 | 9% | 10% |
| Cactus | 1920x1080 | 7% | 17% |
| ParkScene | 1920x1080 | 10% | 6% |
| **Overall** | | **11%** | **12%** |

## III. OPTIMIZED MVP SCHEME

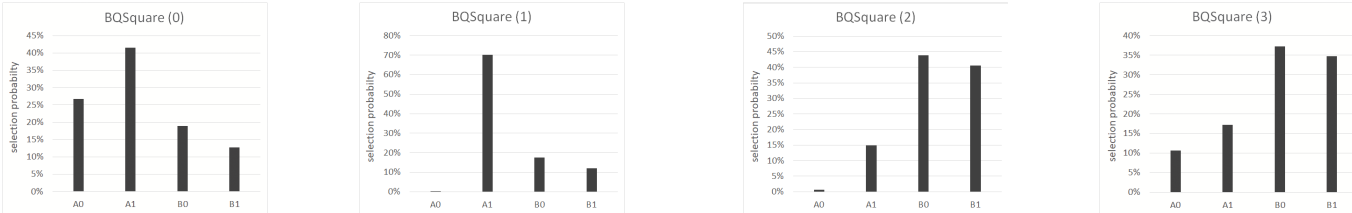
In this section the proposed optimized MVP scheme will be described. With the proposed MVP scheme, we firstly modified the scaling mechanism in AMVP to make the MVP candidates list more manifold. After constructing a better MVP candidates list, we add an expanded area motion vector prediction (EMVP) to enhance the accuracy of the MVP candidates. The detailed descriptions of this optimized MVP scheme are addressed in the following subsections.

### *A. Enhanced AMVP scaling mechanism*

In this section, the relationship between MVP list order and probability of selection for different candidates is analyzed first. We change the order of original AMVP mechanism:{$A_0$, $A_1$, scaled $A_0$,scaled $A_1$},{$B_0$, $B_1$, $B_2$, scaled $B_0$, scaled $B_1$, scaled $B_2$}into the following different cases:

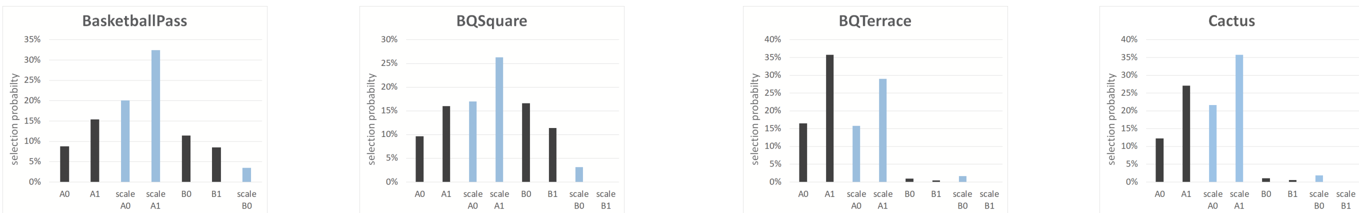
- case 1:{$A_1$, $A_0$, scaled $A_1$, scaled $A_0$},{$B_0$, $B_1$, $B_2$, scaled $B_0$, scaled $B_1$, scaled $B_2$}
- case 2:{$B_0$, $B_1$, $B_2$, scaled $B_0$, scaled $B_1$, scaled $B_2$},{$A_1$, $A_0$, scaled $A_1$, scaled $A_0$}
- case 3:{$B_0$, $B_1$, $B_2$, scaled $B_0$, scaled $B_1$, scaled $B_2$},{$A_0$, $A_1$, scaled $A_0$, scaled $A_1$}

As shown in Fig. 3, with the histogram of probability calculation, we can find that the MVP list order has a major impact on list construction. The selection probability for A0 is nearly zero in case 1 and case 2, which will desert the useful




(a) (b) (c) (d)

Fig. 3. Selection probability for each candidate for “BQSquare” video sequence in: (a) case 0; (b) case 1; (c)case 2; (d)case 3.




(a) (b) (c) (d)

Fig. 4. Selection probability for each candidate (including scaled MVP) for: (a) “BasketballPass”; (b) “BQSquare”; (c)“BQTerrace”; (d)“Cactus” video sequences.

motion information of block located at A0 and therefore the accuracy of motion vector prediction will be reduced.

We calculate the selection probability in AMVP mechanism more specifically and the result is shown in Fig. 4. From the histogram, we can figure out several facts: (1) The percentage of selection for scaling A0 and A1 is much higher than B0 and B1, which means the useful motion information of blocks above the current block will not be used in some cases; (2) The percentage of selection for scaling B0 and B1 is extremely low. With these facts, we try to modify the construction process of AMVP list as: (1) Change the candidates order to $\{A_0, A_1\}$,$\{B_0, B_1, B_2\}$,{(scaled $A_0$, scaled $A_1$),(scaled $B_0$, scaled $B_1$, scaled $B_2$)}; (2) Change the applying conditions for scaling MVs of blocks above the current block as shown in Fig. 5.

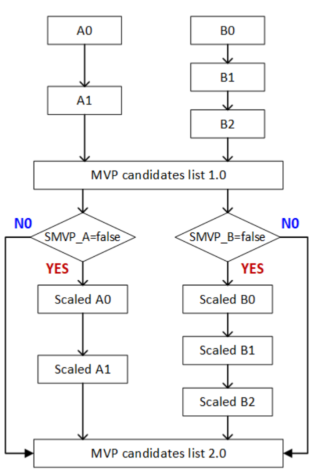



Fig. 5. The flowchart of Proposed modified AMVP mechanism

### *B. Expanded-area Motion Vector Prediction*

To better make use of available MV information of surrounding blocks, we proposed an expanded area EMVP to enhance the accuracy of AMVP. From 2.2, we find that the percentage of intra-mode blocks in JEM 7.0 still constitutes a high proportion. Thus, we try to perform EMVP after Pattern matched motion vector derivation (PMMVD) mode [7], which is at the end of the construction process.

The proposed EMVP scheme is illustrated in Fig. 6. The expanded area of the current block is the left block of A0, A1 and the upper block of B0, B1, which are already compressed and their motion information can be utilized to derive MVP for the current block. The applying conditions of EMVP are set as follows:

- Case 1: when PMMVD is performed and a $MV_{PMMVD}$ will be added as the first candidate in the MVP lists. If no candidate has been added in the first two process (SMVP and TMVP), zero MV will be added into the list. In this condition, we apply EMVP to replace zero MV when the following condition is satisfied:

$$D(MV_{PMMVD}, MV_{EMVP}) < D(MV_{PMMVD}, MV_{ZERO}) \quad (3)$$

where $MV_{PMMVD}$ is the MVP derived by PMMVD, $MV_{EMVP}$ is the MVP derived by EMVP, and $MV_{ZERO}$ means the zero MV. D(x, y) is the Manhattan distance between point x and y and i depends on the dimensions of the vectors:

$$D(x, y) = \Sigma_i |x_i - y_i| \quad (4)$$

- Case 2:when there are two candidates in the list after PMMVD, one each from SMVP process and TMVP

TABLE II
THE PERFORMANCE OF PROPOSED OPTIMIZED MVP SCHEME COMPARED WITH JEM 7.0

| Sequences | | Lowdelay-P | | | Lowdelay-B | | | Random Access | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Y | U | V | Y | U | V | Y | U | V |
| Class A | Campfire | 0.11% | -0.06% | -0.35% | 0.02% | 0.04% | -0.23% | 0.06% | -0.05% | -0.04% |
| | CatRobot1 | -0.09% | -0.38% | -0.53% | -0.17% | -0.51% | -0.25% | -0.04% | -0.13% | -0.24% |
| | DaylightRoad2 | **-0.25%** | -0.44% | 0.02% | **-0.20%** | -0.13% | 0.01% | -0.12% | 0.06% | 0.03% |
| | FoodMarket4 | 0.03% | -0.14% | -0.07% | -0.15% | 1.05% | 0.20% | 0.05% | 0.39% | 0.32% |
| | ParkRunning3 | -0.10% | -0.17% | 0.06% | -0.02% | -0.13% | 0.01% | -0.08% | 0.08% | -0.15% |
| **Overall** | | **-0.06%** | **-0.24%** | **-0.17%** | **-0.11%** | **0.06%** | **-0.05%** | **-0.03%** | **0.07%** | **-0.02%** |
| Class B | BasketballDrive | **-0.23%** | 0.33% | -0.06% | 0.11% | -0.04% | 0.45% | -0.05% | 0.10% | -0.42% |
| | BQTerrace | **-0.29%** | -1.51% | -0.70% | -0.07% | 0.38% | -0.05% | 0.04% | -0.32% | -0.08% |
| | Cactus | -0.08% | 0.65% | 0.35% | -0.07% | -0.33% | -0.10% | -0.07% | 0.18% | -0.17% |
| | MarketPlace | 0.11% | 0.28% | -1.97% | -0.17% | 0.35% | -0.53% | 0.02% | 0.95% | 0.79% |
| | RitualDance | **-0.44%** | -0.32% | -0.29% | -0.08% | -0.18% | -0.41% | 0.07% | -1.18% | -0.55% |
| **Overall** | | **-0.19%** | **-0.12%** | **-0.53%** | **-0.05%** | **0.04%** | **-0.12%** | **-0.00%** | **-0.05%** | **-0.09%** |

process separately. We try to apply EMVP to replace TMVP when the following condition is satisfied:

$$D(MV_{SMVP}, MV_{EMVP}) =< 8 \tag{5}$$

where $MV_{SMVP}$ is the MVP derived by spatial motion vector prediction. And the threshold is determined by experimental results.

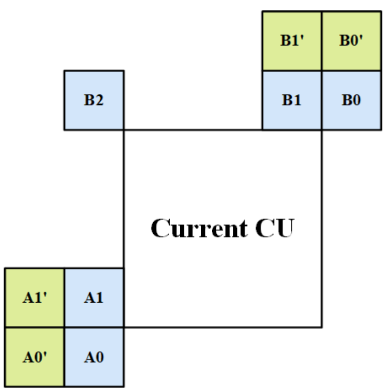



Fig. 6. Proposed EMVP scheme

## IV. EXPERIMENTAL RESULTS

The simulations of the proposed optimized MVP scheme was performed with JEM 7.0, which is the reference software for FVC. To verify the performance of the proposed scheme, we conduct simulations with three inter-related configurations used in FVC: Lowdelay-P(LDP), Random Access(RA) and Lowdelay-B(LDB). Ten video sequences [8] with various video characteristics are utilized as test sequences in our experiments. For each sequence, Quantization Parameter(QP) value of 32, 37, 42 and 47 are specified in our experiments. The coding performance is measured by Bjontegaard's method [9] in terms of BD-rate(Y component). The experimental results are shown in Table 2. From Table 3, it can be seen that for Class B, the average BD-rate gains are 0.19% for luma component at LDP configuration. Specially, for sequences "RitualDance" and "BQTerrace", the BD-rate gains can achieve up to 0.44% and 0.29% respectively.

This result shows that the proposed optimized MVP scheme succeeds to improve the coding performance.

## V. CONCLUSION

In this paper, we propose an optimized MVP mechanism, which includes two schemes: modifying the scaling process during the construction of AMVP mechanism and expanded-area motion vector prediction scheme. Having these two schemes work together can enhance the accuracy of MVP and get 0.13% BD-rate gains for luma component at Lowdelay-P configuration on average compared with JEM 7.0.

## ACKNOWLEDGMENT


This work was supported in part by National Natural Science Foundation of China (61632001), General Financial Grant from the China Postdoctoral Science Foundation(2016M600866), Top-Notch Young Talents Program of China, High-performance Computing Platform of Peking University, which are gratefully acknowledged.


## REFERENCES


[1] H.261: CCITT, "Draft revision of recommendation H.261: video codec for audiovisual services at p64 kbits/s," Signal Processing: Image Communication, Vol. 2, No. 2, pp. 221-240, Aug. 1990

[2] J. -R. Ohm, G. J. Sullivan, W. J. Han and T. Wiegand, "Overview of the High Efficiency Video Coding(HEVC) Standard," IEEE Trans. Circuits and Systems for Video Tech, Vol. 22, No. 12, pp. 1649-1668, Dec. 2012.

[3] ITU-T/SG15 Draft Recommendation H.263 Video Coding for Low Bitrate Communication, July 1995.

[4] J. L. Lin, Y. W. Chen, Y. W. Huang and S. M. Lei, "Motion Vector Coding in the HEVC standard," IEEE Journal of Selected Topics in Signal Processing, vol. 7, no. 6, pp. 957-968, 2013.

[5] J. L. Chen, E. Alshina, G. J. Sullivan, J. -R. Ohm, J. Boyce. "Algorithm Description of Joint Exploration Test Model 7 (JEM 7)", JVET of ITU-T SG 16 WP 3 and ISO/IEC JTC1/SC29/WG11, Doc. JVET-G1001-v1, 7th Meeting, Torine, Italy, 2017

[6] D. Liu, D. B. Zhao, Q. Wang, W. Gao, “An improved motion vector prediction scheme for video coding,” Pacific-Rim Conference on Multimedia, pp. 598-605, Hangzhou, China, 2006.
[7] S. Kamp, M. Evertz, M. Wien, “Decoder side motion vector derivation for inter frame video coding”, IEEE International Conference on Image Processing, pp. 1120-1123, Oct. 2008.
[8] A. Segall, V. Baroncini, J. Chen and T. Suzuki, “Joint call for proposals on video compression with capability beyond HEVC”, JVET of ITU-T SG 16 WP 3 and ISO/IEC JTC1/SC29/WG11, Doc. JVET-H1102, 8th Meeting, Macao, CN, 2017.
[9] G. Bjontegaard, “Calculation of average PSNR differences between RD-curves”, VCEG-M33, 13th VCEG Meeting, Austin, TX, USA, 2001.