An Improvement to Merge Mode in ECM With Template Matching

Ru-Ling Liao, Yan Ye, Jie Chen and Xinwei Li
Alibaba Group
DCC 2023

Outline

- Introduction
- Overview of merge modes
- Proposed improvement
- Experimental results
- Conclusion

Introduction

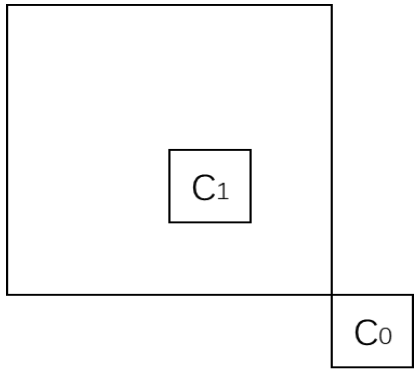
- Versatile Video Coding (VVC) was published in July 2020
  - Developed by the Joint Video Experts Team (JVET)
  - Based on hybrid video coding system with more coding tools supported
  - Significantly exceeded the compression efficiency of the HEVC standard

- Exploring coding technologies beyond VVC capability was started in JVET
  - Established Enhanced Compression Model (ECM) software platform
  - Included more decoder-side prediction technologies
  - Achieved 18.5% and 15.6% BD-rate reduction for RA and LB configurations as compared to VVC

- This paper proposes to improve merge modes with decoder-side derivation technologies

Temporal motion vector predictor (TMVP)

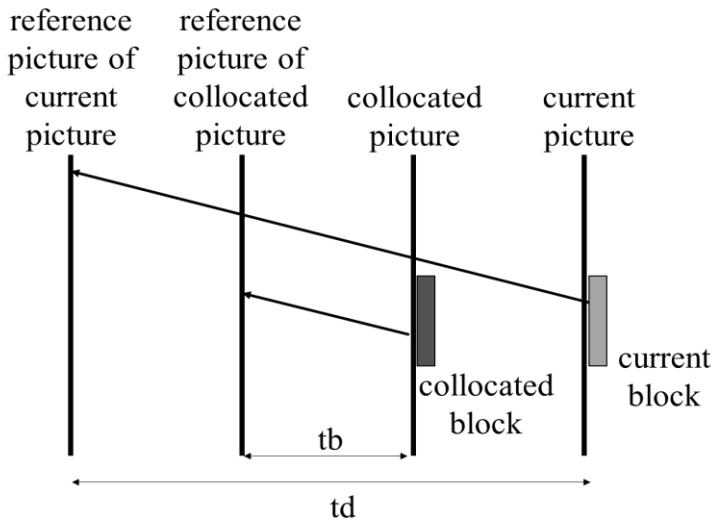
- Obtain TMVP from collocated block C_0 or C_1



- Perform motion vector scaling according to POC difference

$$TMVP = \frac{td}{tb} \times MV\ of\ the\ collocated\ block$$

- Result in TMVP representing bi-predicted motion in VVC



Bi-prediction with CU-level weight (BCW)

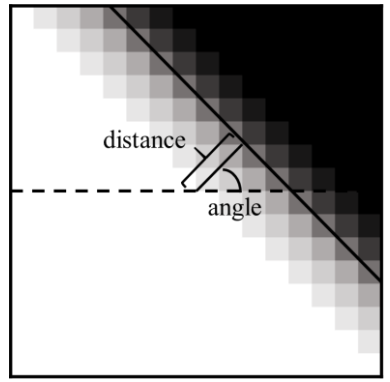
- Newly adopted tool in VVC
- Extend simple averaging to weighted averaging for bi-prediction blocks
- 5 weights {-2/8, 3/8, 4/8, 5/8, 10/8} for low delay pictures
  3 weights {3/8, 4/8, 5/8} for non-low delay pictures

$$P_{bi-pred} = (8 \times (1 - w) \times P_0 + 8 \times w \times P_1 + 4) \gg 3$$

- Signal BCW weights in the bitstream for non-merge coded blocks
- Inherit BCW weights from the corresponding neighboring block for merge coded blocks

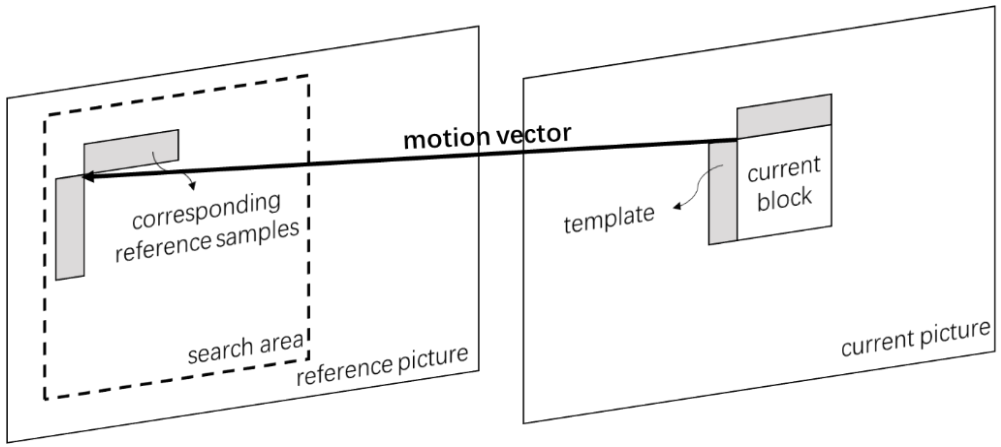
Geometric partition mode (GPM)

- Partition a block into two segments
- Support 64 partition modes in VVC
- Predict each segment with its own motion information
- Derive motion information from neighboring blocks
- Apply blending to samples along partition edge



Template matching (TM)

- Decoder-side motion vector derivation technology adopted in ECM
- Refine motion vector by finding the minimum TM cost of a template and its references
- Reorder candidates according to TM cost
- Determine best prediction mode based on TM cost



Proposed improvement

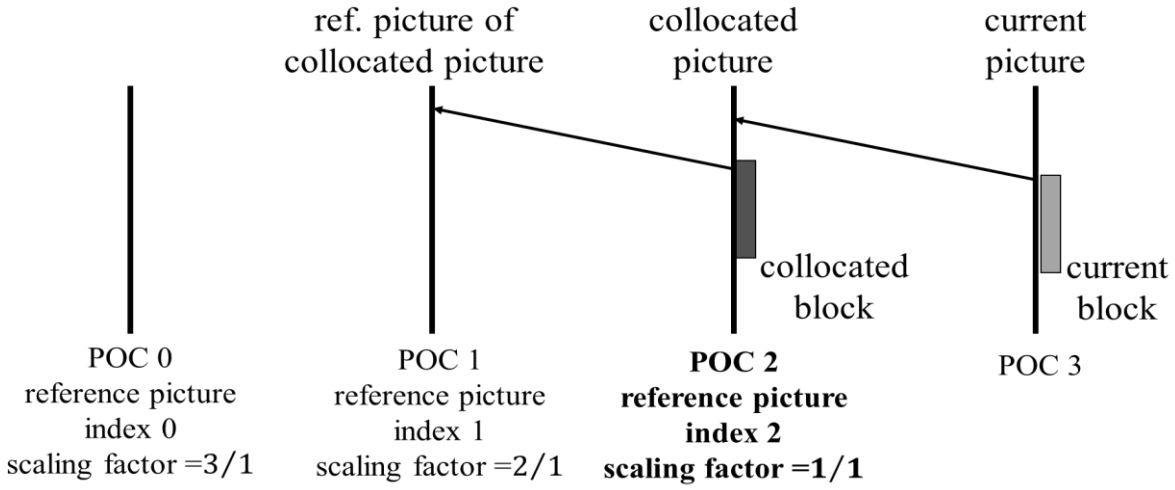
- Utilize the benefits of decoder-side motion information derivation technology

- Propose to apply template matching to TMVP, BCW and GPM
  - Determine inter prediction direction of TMVP
  - Decide bi-prediction weights of merge blocks
  - Refine motion information of two individual geometric partitions

Improvement to TMVP

- Decide the prediction direction for TMVP based on TM cost
  - Calculate TM costs for L0-predicted, L1-predicted and bi-predicted TMVP
  - Select the one with minimum TM cost as TMVP for merge mode

- Determine the reference picture index according to scaling factor instead of fixed to 0
  - Select the one with scaling factor closest to 1



Improvement to BCW

- Adjust bi-predicted weights for merge blocks according to TM cost
  - Select the one with minimum TM cost

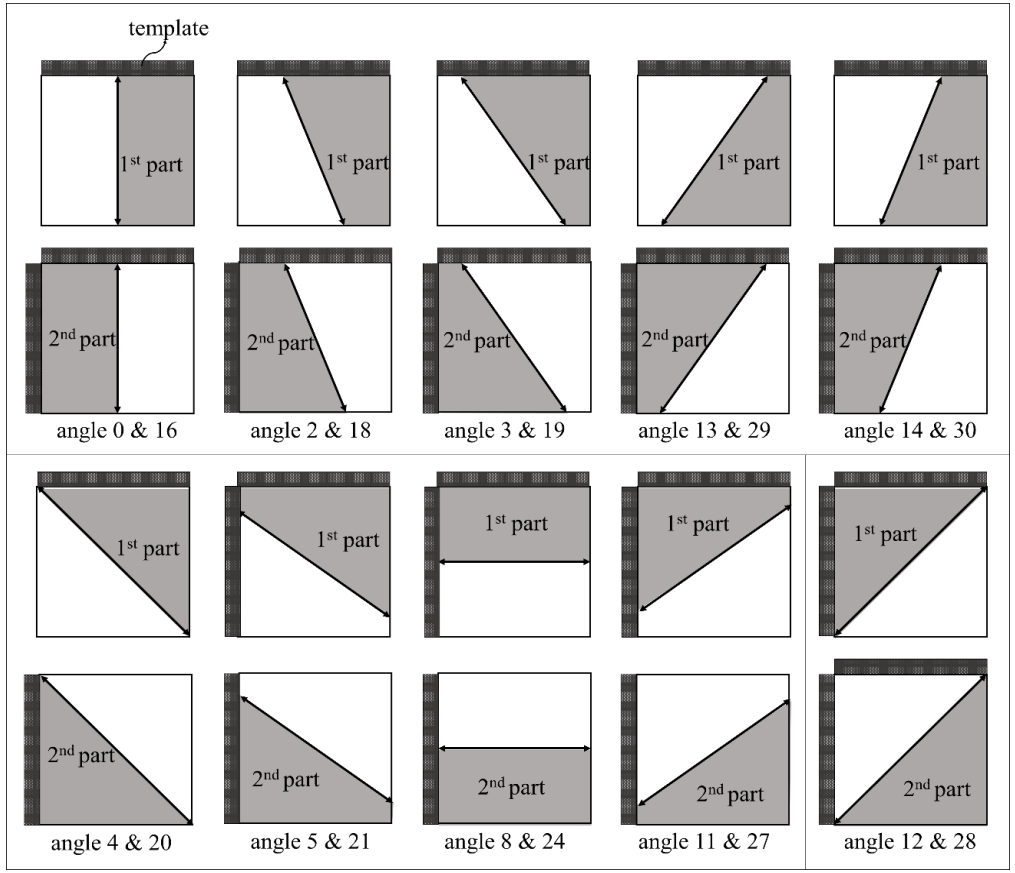
- Extend bi-predicted weights to $\{\frac{-4}{8}, \frac{-3}{8}, \frac{-2}{8}, \frac{-1}{8}, \frac{1}{8}, \frac{2}{8}, \frac{3}{8}, \frac{4}{8}, \frac{5}{8}, \frac{6}{8}, \frac{7}{8}, \frac{8}{8}, \frac{9}{8}, \frac{10}{8}, \frac{11}{8}, \frac{12}{8}\}$

- Apply three rules when calculating TM cost
  - Only the two neighboring weights (i.e. ± 1) and the inherited weight are considered
  - The TM cost of the inherited weight is reduced by $\frac{3}{32}$
  - The TM cost of the equal weight (i.e. $\frac{4}{8}$) is reduced by $\frac{3}{32}$

Improvement to GPM

- Refine two GPM motion using TM individually

- Construct template according to partition angle
  - Only left neighboring samples
  - Only above neighboring samples
  - Left and above neighboring samples



Experimental results

Class	Resolution	Random Access Main10			Low delay B Main10		
		Y	U	V	Y	U	V
Class A1	3840×2160	-0.26%	-0.49%	-0.54%			
Class A2	3840×2160	-0.14%	-0.19%	-0.29%			
Class B	1920×1080	-0.19%	-0.40%	-0.38%	-0.35%	-0.58%	-0.58%
Class C	832×480	-0.20%	-0.29%	-0.19%	-0.55%	-0.46%	-0.49%
Class E	1280×720				-0.01%	-0.71%	-0.11%
Overall Summary	Average	-0.20%	-0.35%	-0.34%	-0.33%	-0.57%	-0.43%
	EncT	101%			103%		
	DecT	99%			100%		

- Tested on top of ECM-6.0 and under JVET common test condition
- 0.20% BD-rate saving for RA
- 0.33% BD-rate saving for LB
- Negligible encoding and decoding runtime

Results of individual improvement

- Largely additive with minor overlap in coding gain of proposed methods
- Higher gain in LB since TMVP has higher probability of being selected

Class C results

Methods	Random Access Main10			Low delay B Main10		
	Y	U	V	Y	U	V
TMVP	0.00%	0.01%	0.02%	-0.37%	-0.18%	-0.29%
BCW	-0.04%	-0.03%	0.06%	-0.05%	0.20%	-0.04%
GPM	-0.16%	-0.15%	-0.02%	-0.18%	0.31%	-0.33%
Overall	-0.20%	-0.29%	-0.19%	-0.55%	-0.46%	-0.49%

Conclusion

- Improved TMVP, BCW and GPM using decoder-side derivation technology
- 0.20% and 0.33% BD-rate saving for RA and LB configurations, respectively
- No impact on encoding and decoding runtime
- Adopted in ECM software

Thank you!