

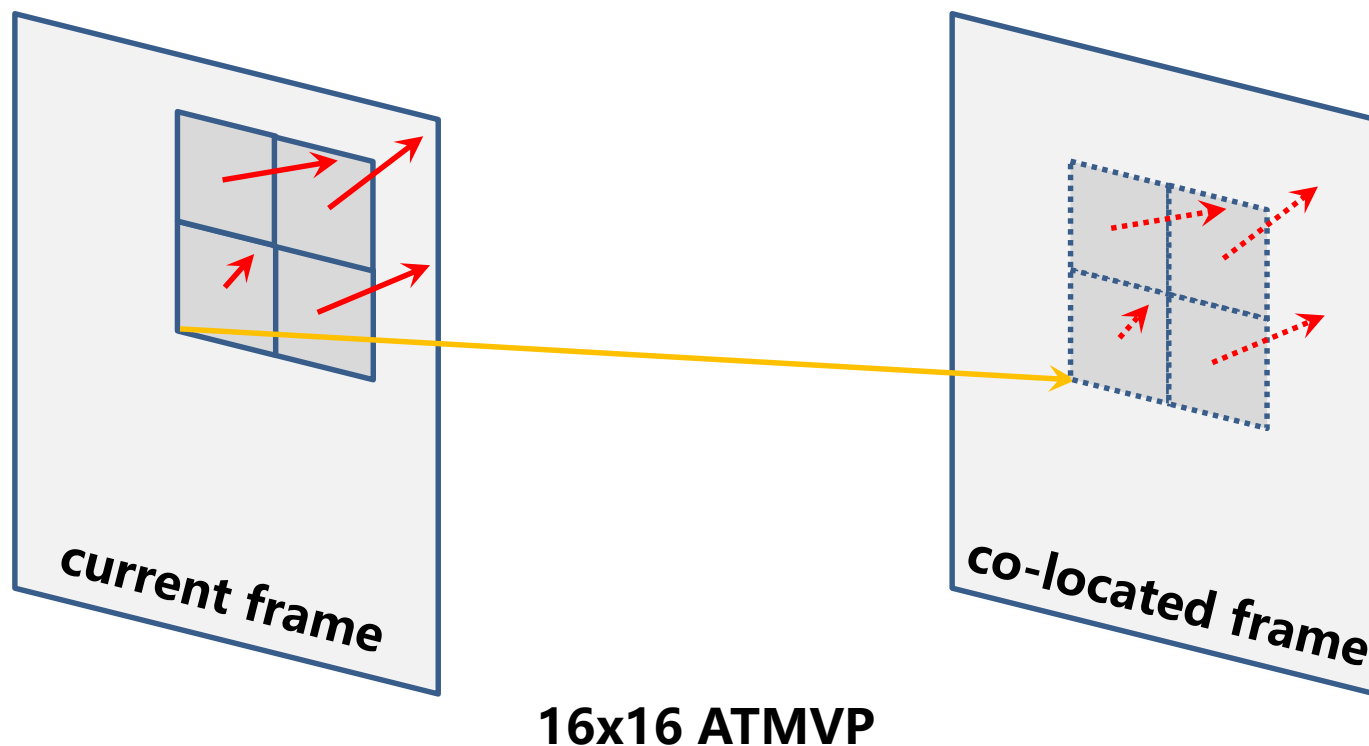


Non-CE4: Constraints on block size for ATMVP

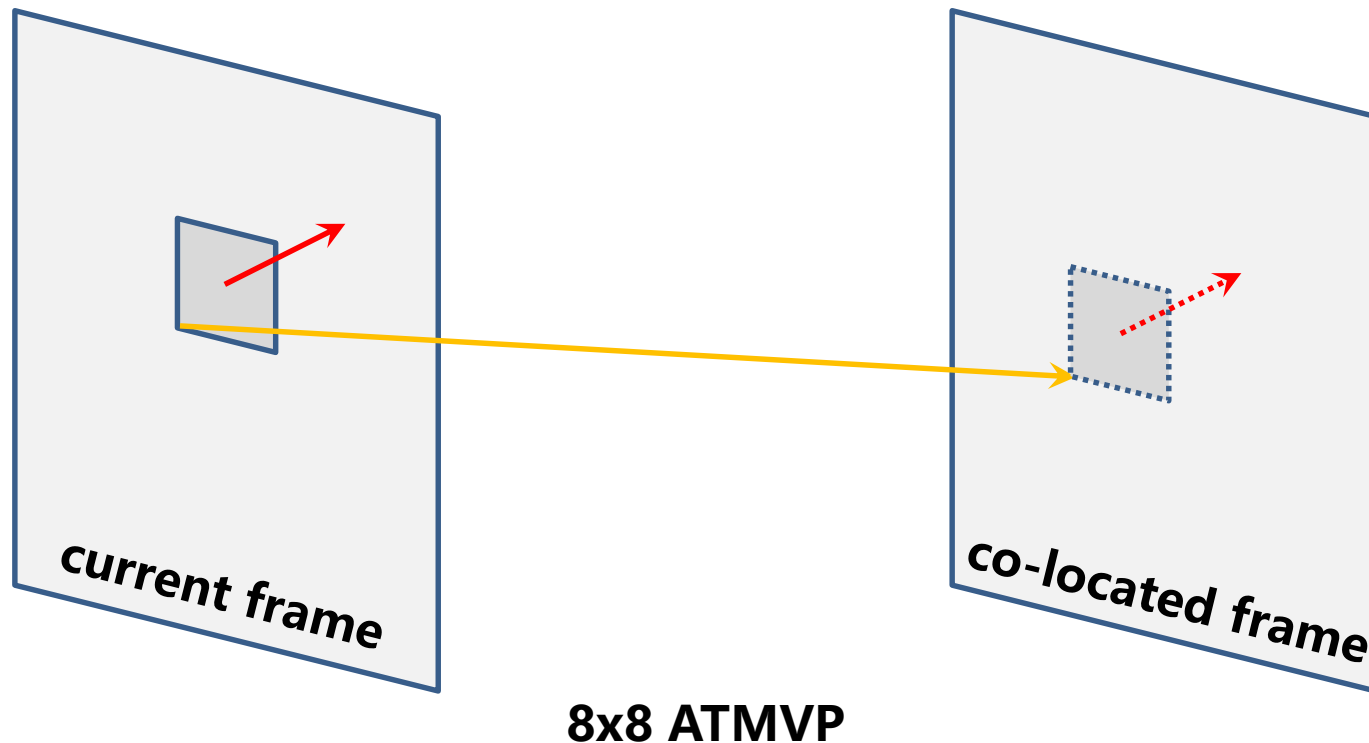
Yucheng Sun, Fangdong Chen, Li Wang
2020/1/12

- Motivation
- Proposed Method
- Experimental Results
- Conclusion

- Sub-blocks merge mode cleanup
 - When applying ATMVP on 8x8 block
 - ATMVP mode degenerates into a temporal merge mode



- Sub-blocks merge mode cleanup
 - When applying ATMVP on 8x8 block
 - ATMVP mode degenerates into a temporal merge mode



- Sub-blocks merge mode cleanup
 - Remove sub-blocks merge mode which contains only 1 sub-blocks
- Remove redundant spatial and temporal check for 8x8 blocks

• Spatial and temporal check

Otherwise, the following ordered steps apply:

1. The location (xCtb, yCtb) of the top-left sample of the luma coding tree block that contains the current coding block and the location (xCtr, yCtr) of the below-right center sample of the current luma coding block are derived as follows:

$$xCtb = (xCb \gg CtuLog2Size) \ll CtuLog2Size \quad (724)$$

$$yCtb = (yCb \gg CtuLog2Size) \ll CtuLog2Size \quad (725)$$

$$xCtr = xCb + (cbWidth / 2) \quad (726)$$

$$yCtr = yCb + (cbHeight / 2) \quad (727)$$

2. The luma location (xColCtb, yColCtb) is set equal to the top-left sample of the collocated luma coding block covering the location given by (xCtr, yCtr) inside ColPic relative to the top-left luma sample of the collocated picture specified by ColPic.
3. The derivation process for subblock-based temporal merging base motion data as specified in clause 8.5.5.4 is invoked with the location (xCtb, yCtb), the location (xColCtb, yColCtb), the availability flag availableFlagA₁, and the prediction list utilization flag predFlagLXA₁, and the reference index refIdxLXA₁, and the motion vector mvLXA₁, with X being 0 and 1 as inputs and the motion vectors ctrMvLX, and the prediction list utilization flags ctrPredFlagLX of the collocated block, with X being 0 and 1, and the temporal motion vector tempMv as outputs.
4. The variable availableFlagSbCol is derived as follows:
 - If both ctrPredFlagL0 and ctrPredFlagL1 are equal to 0, availableFlagSbCol is set equal to 0.
 - Otherwise, availableFlagSbCol is set equal to 1.

8.5.5.4 Derivation process for subblock-based temporal merging base motion data

Inputs to this process are:

- the location (xCtb, yCtb) of the top-left sample of the luma coding tree block that contains the current coding block,
- the location (xColCtb, yColCtb) of the top-left sample of the collocated luma coding block that covers the below-right center sample.
- the availability flag availableFlagA₁ of the neighbouring coding unit,
- the reference index refIdxLXA₁ of the neighbouring coding unit,
- the prediction list utilization flag predFlagLXA₁ of the neighbouring coding unit,
- the motion vector in 1/16 fractional-sample accuracy mvLXA₁ of the neighbouring coding unit.

Outputs of this process are:

- the motion vectors ctrMvL0 and ctrMvL1,
- the prediction list utilization flags ctrPredFlagL0 and ctrPredFlagL1,
- the temporal motion vector tempMv.

The variable tempMv is set as follows:

$$\text{tempMv}[0] = 0 \quad (740)$$

$$\text{tempMv}[1] = 0 \quad (741)$$

The variable currPic specifies the current picture.

When availableFlagA₁ is equal to TRUE, the following applies:

- If all of the following conditions are true, tempMv is set equal to mvL0A₁:
 - predFlagL0A₁ is equal to 1,
 - DiffPicOrderCnt(ColPic, RefPicList[0][refIdxL0A₁]) is equal to 0,
- Otherwise, if all of the following conditions are true, tempMv is set equal to mvL1A₁:
 - slice_type is equal to B,
 - predFlagL1A₁ is equal to 1,
 - DiffPicOrderCnt(ColPic, RefPicList[1][refIdxL1A₁]) is equal to 0.

The rounding process for motion vectors as specified in clause 8.5.2.14 is invoked with mvX set equal to tempMv, rightShift set equal to 4, and leftShift set equal to 0 as inputs and the rounded tempMv as output.

The location (xColCb, yColCb) of the collocated block inside ColPic is derived as follows.

- The following applies:

$$yColCb = \text{Clip3}(yCtb, \text{Min}(\text{pic_height_in_luma_samples} - 1, yCtb + (1 \ll CtbLog2SizeY) - 1), yColCtb + \text{tempMv}[1]) \quad (742)$$

- If subpic_treated_as_pic_flag[SubPicIdx] is equal to 1, the following applies:

$$xColCb = \text{Clip3}(xCtb, \text{Min}(\text{SubPicRightBoundaryPos}, xCtb + (1 \ll CtbLog2SizeY) + 3), xColCtb + \text{tempMv}[0]) \quad (743)$$

- Otherwise (subpic_treated_as_pic_flag[SubPicIdx] is equal to 0), the following applies:

$$xColCb = \text{Clip3}(xCtb, \text{Min}(\text{pic_width_in_luma_samples} - 1, xCtb + (1 \ll CtbLog2SizeY) + 3), xColCtb + \text{tempMv}[0]) \quad (744)$$

The array colPredMode is set equal to the prediction mode array CuPredMode[0] of the collocated picture specified by ColPic.

The motion vectors ctrMvL0 and ctrMvL1, and the prediction list utilization flags ctrPredFlagL0 and ctrPredFlagL1 are derived as follows:

- If colPredMode[xColCb][yColCb] is equal to MODE_INTER, the following applies:

- The variable currCb specifies the luma coding block covering (xCtb, yCtb) inside the current picture.
- The variable colCb specifies the luma coding block covering the modified location given by ((xCbCb >> 3) << 3, (yCbCb >> 3) << 3) inside the ColPic.
- The luma location (xColCb, yColCb) is set equal to the top-left sample of the collocated luma coding block specified by colCb relative to the top-left luma sample of the collocated picture specified by ColPic.
- The derivation process for collocated motion vectors specified in clause 8.5.2.12 is invoked with currCb, colCb, (xColCb, yColCb), refIdxL0 set equal to 0, and sbFlag set equal to 1 as inputs and the output being assigned to ctrMvL0 and ctrPredFlagL0.
- The derivation process for collocated motion vectors specified in clause 8.5.2.12 is invoked with currCb, colCb, (xColCb, yColCb), refIdxL1 set equal to 0, and sbFlag set equal to 1 as inputs and the output being assigned to ctrMvL1 and ctrPredFlagL1.

- Otherwise, the following applies:

$$\text{ctrPredFlagL0} = 0 \quad (745)$$

$$\text{ctrPredFlagL1} = 0 \quad (746)$$

• Spatial and temporal check

8.5.2.12 Derivation process for collocated motion vectors

Inputs to this process are:

- a variable currCb specifying the current coding block,
- a variable colCb specifying the collocated coding block inside the collocated picture specified by ColPic,
- a luma location (xColCb, yColCb) specifying the top-left sample of the collocated luma coding block specified by colCb relative to the top-left luma sample of the collocated picture specified by ColPic,
- a reference index refIdxLX, with X being 0 or 1,
- a flag indicating a subblock temporal merging candidate sbFlag.

Outputs of this process are:

- the motion vector prediction mvLXCol in 1/16 fractional-sample accuracy,
- the availability flag availableFlagLXCol.

The variable currPic specifies the current picture.

The arrays predFlagL0Col[x][y], mvL0Col[x][y] and refIdxL0Col[x][y] are set equal to PredFlagL0[x][y], MvDmvrL0[x][y] and RefIdxL0[x][y], respectively, of the collocated picture specified by ColPic, and the arrays predFlagL1Col[x][y], mvL1Col[x][y] and refIdxL1Col[x][y] are set equal to PredFlagL1[x][y], MvDmvrL1[x][y] and RefIdxL1[x][y], respectively, of the collocated picture specified by ColPic.

[Ed. (BB): Define ColPic NoBackwardPredFlag.]

The variables mvLXCol and availableFlagLXCol are derived as follows:

- If colCb is coded in an intra, IBC, or palette prediction mode, both components of mvLXCol are set equal to 0 and availableFlagLXCol is set equal to 0.
- Otherwise, the motion vector mvCol, the reference index refIdxCol and the reference list identifier listCol are derived as follows:
 - If sbFlag is equal to 0, availableFlagLXCol is set to 1 and the following applies:
 - If predFlagL0Col[xColCb][yColCb] is equal to 0, mvCol, refIdxCol and listCol are set equal to mvL1Col[xColCb][yColCb], refIdxL1Col[xColCb][yColCb] and L1, respectively.
 - Otherwise, if predFlagL0Col[xColCb][yColCb] is equal to 1 and predFlagL1Col[xColCb][yColCb] is equal to 0, mvCol, refIdxCol and listCol are set equal to mvL0Col[xColCb][yColCb], refIdxL0Col[xColCb][yColCb] and L0, respectively.
 - Otherwise (predFlagL0Col[xColCb][yColCb] is equal to 1 and predFlagL1Col[xColCb][yColCb] is equal to 1), the following assignments are made:
 - If NoBackwardPredFlag is equal to 1, mvCol, refIdxCol and listCol are set equal to mvLXCol[xColCb][yColCb], refIdxLXCol[xColCb][yColCb] and LX, respectively.

- Otherwise, mvCol, refIdxCol and listCol are set equal to mvLNCol[xColCb][yColCb], refIdxLNCol[xColCb][yColCb] and LN, respectively, with N being the value of collocated_from_l0_flag.
- Otherwise (sbFlag is equal to 1), the following applies:
 - If PredFlagLXCol[xColCb][yColCb] is equal to 1, mvCol, refIdxCol, and listCol are set equal to mvLXCol[xColCb][yColCb], refIdxLXCol[xColCb][yColCb], and LX, respectively, availableFlagLXCol is set to 1.
 - Otherwise (PredFlagLXCol[xColCb][yColCb] is equal to 0), the following applies:
 - If NoBackwardPredFlag is equal to 1 and PredFlagLYCol[xColCb][yColCb] is equal to 1, mvCol, refIdxCol, and listCol are set to mvLYCol[xColCb][yColCb], refIdxLYCol[xColCb][yColCb] and LY, respectively, with Y being equal to !X where X being the value of X this process is invoked for. availableFlagLXCol is set equal to 1.
 - Otherwise, both components of mvLXCol are set equal to 0 and availableFlagLXCol is set equal to 0.
- When availableFlagLXCol is equal to TRUE, mvLXCol and availableFlagLXCol are derived as follows:
 - If LongTermRefPic(currPic, currCb, refIdxLX, LX) is not equal to LongTermRefPic(ColPic, colCb, refIdxCol, listCol), both components of mvLXCol are set equal to 0 and availableFlagLXCol is set equal to 0.
 - Otherwise, the variable availableFlagLXCol is set equal to 1, refPicList[listCol][refIdxCol] is set to be the picture with reference index refIdxCol in the reference picture list listCol of the slice containing coding block colCb in the collocated picture specified by ColPic, and the following applies:

$$\text{colPocDiff} = \text{DiffPicOrderCnt}(\text{ColPic}, \text{refPicList}[\text{listCol}][\text{refIdxCol}]) \quad (609)$$

$$\text{currPocDiff} = \text{DiffPicOrderCnt}(\text{currPic}, \text{RefPicList}[X][\text{refIdxLX}]) \quad (610)$$

- The temporal motion buffer compression process for collocated motion vectors as specified in clause 8.5.2.15 is invoked with mvCol as input, and the modified mvCol as output.
- If RefPicList[X][refIdxLX] is a long-term reference picture, or colPocDiff is equal to currPocDiff, mvLXCol is derived as follows:

$$\text{mvLXCol} = \text{mvCol} \quad (611)$$

- Otherwise, mvLXCol is derived as a scaled version of the motion vector mvCol as follows:

$$\text{tx} = (16384 + (\text{Abs}(\text{td}) >> 1)) / \text{td} \quad (612)$$

$$\text{distScaleFactor} = \text{Clip3}(-4096, 4095, (\text{tb} * \text{tx} + 32) >> 6) \quad (613)$$

$$\text{mvLXCol} = \text{Clip3}(-131072, 131071, (\text{distScaleFactor} * \text{mvCol} + 128 - (\text{distScaleFactor} * \text{mvCol} >= 0)) >> 8) \quad (614)$$

where td and tb are derived as follows:

$$\text{td} = \text{Clip3}(-128, 127, \text{colPocDiff}) \quad (615)$$

$$\text{tb} = \text{Clip3}(-128, 127, \text{currPocDiff}) \quad (616)$$

- Skip ATMVP candidate when construct sub-block merge list for 8x8 blocks
- SPEC modification
 - Add one line
 - If one or more of the following conditions is true, availableFlagSbCol is set equal to 0.
 - slice_temporal_mvp_enabled_flag is equal to 0.
 - sps_sbtmvp_enabled_flag is equal to 0.
 - cbWidth is less than 8.
 - cbHeight is less than 8.
 - cbWidth is equal to 8 and cbHeight is equal to 8.

RA over VTM7.0					
	Y	U	V	EncT	DecT
Class A1	0.00%	0.05%	0.03%	100%	99%
Class A2	-0.01%	-0.01%	0.03%	100%	102%
Class B	0.01%	-0.02%	0.02%	100%	98%
Class C	0.02%	0.03%	0.06%	100%	100%
Class E					
Overall	0.01%	0.01%	0.03%	100%	100%
Class D	0.01%	-0.10%	0.06%	100%	100%
Class F	0.01%	0.00%	0.00%	100%	100%

LDB over VTM7.0					
	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B	0.00%	-0.10%	0.14%	100%	101%
Class C	0.05%	-0.19%	-0.05%	100%	97%
Class E	-0.09%	0.37%	-0.68%	99%	99%
Overall	0.00%	-0.01%	-0.13%	100%	99%
Class D	-0.01%	0.25%	0.25%	100%	100%
Class F	0.02%	-0.42%	-0.11%	100%	100%

Thank Huawei for the cross-checking

- JVET-Q0313 test block size constraint on ATMVP
 - Add one line in SPEC
 - Avoid redundant spatial and temporal check for 8x8 blocks of sub-block merge list construction
- Experimental results compared to VTM7.0
 - RA: 0.01% (Y), 0.01% (U), 0.03% (V)
 - LD: 0.00% (Y), -0.01% (U), -0.13% (V)
- Based on the above results, it is suggested to adopt the modification

THANKS !

公司总部：杭州市滨江区东流路700号

股票代码：002415

客服电话：400-700-5998

www.hikvision.com

© 杭州海康威视数字技术股份有限公司

分公司：北京、长沙、成都、重庆、长春、福州、广州、贵阳、杭州、合肥、哈尔滨、呼和浩特、济南、昆明、兰州、南昌、
南京、南宁、宁波、青岛、上海、沈阳、深圳、石家庄、太原、天津、武汉、厦门、乌鲁木齐、西安、郑州.....

海外分公司：洛杉矶、香港、阿姆斯特丹、孟买、圣彼得堡、迪拜、新加坡.....