



(51) International Patent Classification:
H04N 19/52 (2014.01)

(21) International Application Number:
PCT/CN2019/130725

(22) International Filing Date:
31 December 2019 (31.12.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PCT/CN2018/126066
31 December 2018 (31.12.2018) CN
PCT/CN2019/070636
07 January 2019 (07.01.2019) CN
PCT/CN2019/071159
10 January 2019 (10.01.2019) CN

(71) Applicants: **BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.** [CN/CN]; Room B-0035, 2/F, No. 3 Building, No. 30, Shixing Road, Shijingshan District, Beijing 100041 (CN). **BYTEDANCE INC.** [US/US];

12655 West Jefferson Boulevard, Sixth Floor, Suite No. 137, Los Angeles, California 90066 (US).

(72) Inventors: **ZHANG, Kai**; 12655 West Jefferson Boulevard, Sixth Floor, Suite No. 137, Los Angeles, California 90066 (US). **ZHANG, Li**; 12655 West Jefferson Boulevard, Sixth Floor, Suite No. 137, Los Angeles, California 90066 (US). **LIU, Hongbin**; Jinritoutiao Post Office, China Satellite Communications Tower, No. 63, Zhichun Road, Haidian District, Beijing 100080 (CN). **XU, Jizheng**; 12655 West Jefferson Boulevard, Sixth Floor, Suite No. 137, Los Angeles, California 90066 (US). **WANG, Yue**; Jinritoutiao Post Office, China Satellite Communications Tower, No. 63, Zhichun Road, Haidian District, Beijing 100080 (CN).

(74) Agent: **LIU, SHEN & ASSOCIATES**; 10th Floor, Building 1, 10 Caihefang Road, Haidian District, Beijing 100080 (CN).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,

(54) Title: MAPPING BETWEEN DISTANCE INDEX AND DISTANCE IN MERGE WITH MVD

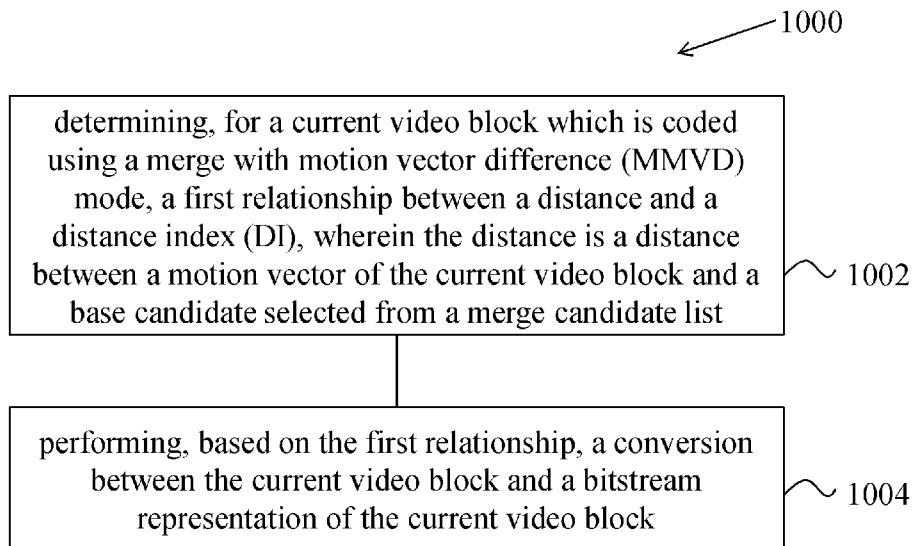


FIG. 10

(57) Abstract: Disclosed and systems, methods and devices for video processing. An exemplary method for video processing includes: determining, for a current video block which is coded using a merge with motion vector difference (MMVD) mode, a first relationship between a distance and a distance index (DI), the distance being a distance between a motion vector of the current video block and a base candidate selected from a merge candidate list; and performing, based on the first relationship, a conversion between the current video block and a bitstream representation of the current video block.



DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

MAPPING BETWEEN DISTANCE INDEX AND DISTANCE IN MERGE WITH MVD

TECHNICAL FIELD

[001] Under the applicable patent law and/or rules pursuant to the Paris Convention, this application is made to timely claim the priority to and benefits of International Patent Application No. PCT/CN2018/126066, filed on December 31, 2018, International Patent Application No. PCT/CN2019/070636, filed on January 7, 2019, and International Patent Application No. PCT/CN2019/071159, filed on January 10, 2019. The entire disclosures thereof are incorporated by reference as part of the disclosure of this application.

[002] The present document relates to video and image coding and decoding.

BACKGROUND

[003] Digital video accounts for the largest bandwidth use on the internet and other digital communication networks. As the number of connected user devices capable of receiving and displaying video increases, it is expected that the bandwidth demand for digital video usage will continue to grow.

SUMMARY

[004] The present document discloses video coding tools that, in one example aspect, improve the signaling of motion vectors for video and image coding.

[005] In one aspect, there is disclosed a method for video processing, comprising: determining, for a current video block which is coded using a merge with motion vector difference (MMVD) mode, a first relationship between a distance and a distance index (DI), wherein the distance is a distance between a motion vector of the current video block and a base candidate selected from a merge candidate list; and performing, based on the first relationship, a conversion between the current video block and a bitstream representation of the current video block.

[006] In one aspect, there is disclosed a method for video processing, comprising: performing a conversion between a current video block and a bitstream representation of the current video block, wherein the current video block is coded using a merge with motion vector difference (MMVD) mode; wherein the conversion comprises parsing MMVD side information from or writing the MMVD side information into the bitstream representation, wherein the MMVD side

information comprises at least one of an MMVD flag indicating whether MMVD syntaxes are parsed, a first syntax element indicating a distance of MMVD between a motion vector of the current video block and a base candidate selected from a merge candidate list, a second syntax element indicating a direction of MMVD representing a direction of motion vector difference (MVD) relative to the base candidate.

[007] In one aspect, there is disclosed a method for video processing, comprising: determining at least one distance for motion vector difference(MVD) associated with a current video block, which is coded in a merge with motion vector difference (MMVD) mode, from a first distance with a rough granularity and one or more distances with fine granularities; and performing a conversion between a current video block and a bitstream representation of the current video block based on the distance for MVD.

[008] In an aspect, there is disclosed an apparatus in a video system, the apparatus comprising a processor and a non-transitory memory with instructions thereon, wherein the instructions upon execution by the processor, cause the processor to implement the method in any one of examples described as above

[009] In an aspect, there is disclosed a computer program product stored on a non-transitory computer readable media, the computer program product including program code for carrying out the method in any one of examples described as above

[0010] In yet another example aspect, the above-described method may be implemented by a video encoder apparatus or a video decoder apparatus that comprises a processor.

[0011] In yet another example aspect, these methods may be embodied in the form of processor-executable instructions and stored on a computer-readable program medium.

[0012] These, and other, aspects are further described in the present document.

BRIEF DESCRIPTION OF DRAWINGS

[0013] FIG. 1 shows an example of simplified affine motion model.

[0014] FIG. 2 shows an example of affine motion vector field (MVF) per sub-block.

[0015] FIG. 3A-3B show 4 and 6 parameter affine models, respectively.

[0016] FIG. 4 shows an example of motion vector predictor (MVP) for AF_INTER.

[0017] FIG. 5A-5B show examples of candidates for AF_MERGE.

[0018] FIG. 6 shows an example of candidate positions for affine merge mode.

[0019] FIG. 7 shows an example of distance index and distance offset mapping.

- [0020] FIG. 8 shows an example of ultimate motion vector expression (UMVE) search process.
- [0021] FIG. 9 shows an example of UMVE search point.
- [0022] FIG. 10 is a flowchart for an example method for video processing.
- [0023] FIG. 11 is a flowchart for another example method for video processing.
- [0024] FIG. 12 is a flowchart for yet another example method for video processing.
- [0025] FIG. 13 shows an example of a hardware platform for implementing a technique described in the present document.

DETAILED DESCRIPTION

[0026] The present document provides various techniques that can be used by a decoder of video bitstreams to improve the quality of decompressed or decoded digital video. Furthermore, a video encoder may also implement these techniques during the process of encoding in order to reconstruct decoded frames used for further encoding.

[0027] Section headings are used in the present document for ease of understanding and do not limit the embodiments and techniques to the corresponding sections. As such, embodiments from one section can be combined with embodiments from other sections.

[0028] 1. Summary

[0029] This patent document is related to video coding technologies. Specifically, it is related to motion compensation in video coding. It may be applied to the existing video coding standard like HEVC, or the standard (Versatile Video Coding) to be finalized. It may be also applicable to future video coding standards or video codec.

[0030] 2. Introductory comments

[0031] Video coding standards have evolved primarily through the development of the well-known ITU-T and ISO/IEC standards. The ITU-T produced H.261 and H.263, ISO/IEC produced MPEG-1 and MPEG-4 Visual, and the two organizations jointly produced the H.262/MPEG-2 Video and H.264/MPEG-4 Advanced Video Coding (AVC) and H.265/HEVC standards. Since H.262, the video coding standards are based on the hybrid video coding structure wherein temporal prediction plus transform coding are utilized. To explore the future video coding technologies beyond HEVC, Joint Video Exploration Team (JVET) was founded by VCEG and MPEG jointly in 2015. Since then, many new methods have been adopted by JVET and put into the reference software named Joint Exploration Model (JEM). In April 2018, the Joint Video

Expert Team (JVET) between VCEG (Q6/16) and ISO/IEC JTC1 SC29/WG11 (MPEG) was created to work on the VVC standard targeting at 50% bitrate reduction compared to HEVC.

[0032] 2.1 Affine motion compensation prediction

[0033] In HEVC, only translation motion model is applied for motion compensation prediction (MCP). While in the real world, there are many kinds of motion, e.g. zoom in/out, rotation, perspective motions and the other irregular motions. In the JEM, a simplified affine transform motion compensation prediction is applied. As shown FIG. 1, the affine motion field of the block is described by two control point motion vectors.

[0034] The motion vector field (MVF) of a block is described by the following equation:

$$\begin{cases} v_x = \frac{(v_{1x} - v_{0x})}{w} x - \frac{(v_{1y} - v_{0y})}{w} y + v_{0x} \\ v_y = \frac{(v_{1y} - v_{0y})}{w} x + \frac{(v_{1x} - v_{0x})}{w} y + v_{0y} \end{cases} \quad (1)$$

[0035] Where (v_{0x}, v_{0y}) is motion vector of the top-left corner control point, and (v_{1x}, v_{1y}) is motion vector of the top-right corner control point.

[0036] In order to further simplify the motion compensation prediction, sub-block based affine transform prediction is applied. The sub-block size $M \times N$ is derived as in Equation 2, where $MvPre$ is the motion vector fraction accuracy (1/16 in JEM), (v_{2x}, v_{2y}) is motion vector of the bottom-left control point, calculated according to Equation 1.

$$\begin{cases} M = clip3(4, w, \frac{w \times MvPre}{\max(abs(v_{1x} - v_{0x}), abs(v_{1y} - v_{0y}))}) \\ N = clip3(4, h, \frac{h \times MvPre}{\max(abs(v_{2x} - v_{0x}), abs(v_{2y} - v_{0y}))}) \end{cases} \quad (2)$$

[0037] After derived by Equation 2, M and N should be adjusted downward if necessary to make it a divisor of w and h, respectively.

[0038] To derive motion vector of each $M \times N$ sub-block, the motion vector of the center sample of each sub-block, as shown in FIG. 2, is calculated according to Equation 1, and rounded to 1/16 fraction accuracy.

[0039] After MCP, the high accuracy motion vector of each sub-block is rounded and saved as the same accuracy as the normal motion vector.

[0040] 2.1.1 AF_INTER mode

[0041] In the JEM, there are two affine motion modes: AF_INTER mode and AF_MERGE mode. For CUs with both width and height larger than 8, AF_INTER mode can be applied. An affine flag in CU level is signaled in the bitstream to indicate whether AF_INTER mode is used.

In this mode, a candidate list with motion vector pair $\{(v_0, v_1) | v_0 = \{v_A, v_B, v_C\}, v_1 = \{v_D, v_E\}\}$ is constructed using the neighbour blocks. As shown in FIG. 4, v_0 is selected from the motion vectors of the block A, B or C. The motion vector from the neighbour block is scaled according to the reference list and the relationship among the POC of the reference for the neighbour block, the POC of the reference for the current CU and the POC of the current CU. And the approach to select v_1 from the neighbour block D and E is similar. If the number of candidate list is smaller than 2, the list is padded by the motion vector pair composed by duplicating each of the AMVP candidates. When the candidate list is larger than 2, the candidates are firstly sorted according to the consistency of the neighbouring motion vectors (similarity of the two motion vectors in a pair candidate) and only the first two candidates are kept. An RD cost check is used to determine which motion vector pair candidate is selected as the control point motion vector prediction (CPMVP) of the current CU. And an index indicating the position of the CPMVP in the candidate list is signaled in the bitstream. After the CPMVP of the current affine CU is determined, affine motion estimation is applied and the control point motion vector (CPMV) is found. Then the difference of the CPMV and the CPMVP is signaled in the bitstream.

[0042] FIG. 3A shows an example of a 4-parameter affine model. FIG. 3B shows an example of a 6-parameter affine model.

[0043] In AF_INTER mode, when 4/6 parameter affine mode is used, 2/3 control points are required, and therefore 2/3 MVD needs to be coded for these control points, as shown in FIG. 3A. In an example, it is proposed to derive the MV as follows, e.g., mvd_1 and mvd_2 are predicted from mvd_0 .

$$\begin{aligned} mv_0 &= \overline{mv}_0 + mvd_0 \\ mv_1 &= \overline{mv}_1 + mvd_1 + mvd_0 \\ mv_2 &= \overline{mv}_2 + mvd_2 + mvd_0 \end{aligned}$$

[0044] Wherein \overline{mv}_i , mvd_i and mv_i are the predicted motion vector, motion vector difference and motion vector of the top-left pixel ($i = 0$), top-right pixel ($i = 1$) or left-bottom pixel ($i = 2$) respectively, as shown in FIG. 3B. Please note that the addition of two motion vectors (e.g., $mvA(xA, yA)$ and $mvB(xB, yB)$) is equal to summation of two components separately, that is, $newMV = mvA + mvB$ and the two components of newMV is set to $(xA + xB)$ and $(yA + yB)$, respectively.

[0045] 2.1.2 Fast affine ME algorithm in AF_INTER mode

[0046] In affine mode, MV of 2 or 3 control points needs to be determined jointly. Directly searching the multiple MVs jointly is computationally complex. A fast affine ME algorithm is proposed and is adopted into VTM/BMS.

[0047] The fast affine ME algorithm is described for the 4-parameter affine model, and the idea can be extended to 6-parameter affine model.

$$\begin{cases} x' = ax + by + c \\ y' = -bx + ay + d \end{cases} \quad (3)$$

$$\begin{cases} mv_{(x,y)}^h = x' - x = (a-1)x + by + c \\ mv_{(x,y)}^v = y' - y = -bx + (a-1)y + d \end{cases} \quad (4)$$

[0048] Replace $(a-1)$ with a' , then the motion vector can be rewritten as:

$$\begin{cases} mv_{(x,y)}^h = x' - x = a'x + by + c \\ mv_{(x,y)}^v = y' - y = -bx + a'y + d \end{cases} \quad (5)$$

[0049] Suppose motion vectors of the two controls points $(0, 0)$ and $(0, w)$ are known, from Equation (5) we can derive affine parameters,

$$\begin{cases} c = mv_{(0,0)}^h \\ d = mv_{(0,0)}^v \end{cases} \quad (6)$$

[0050] The motion vectors can be rewritten in vector form as:

$$MV(p) = A(P) * MV_C^T \quad (7)$$

[0051] Wherein

$$A(P) = \begin{bmatrix} 1 & x & 0 & y \\ 0 & y & 1 & -x \end{bmatrix} \quad (8)$$

$$MV_C = [mv_{(0,0)}^h \quad a \quad mv_{(0,0)}^v \quad b] \quad (9)$$

[0052] $P = (x, y)$ is the pixel position.

[0053] At encoder, MVD of AF_INTER are derived iteratively. Denote $MV^i(P)$ as the MV derived in the i th iteration for position P and denote dMV_C^i as the delta updated for MV_C in the i th iteration. Then in the $(i+1)$ th iteration,

$$\begin{aligned} MV^{i+1}(P) &= A(P) * \left((MV_C^i)^T + (dMV_C^i)^T \right) \\ &= A(P) * (MV_C^i)^T + A(P) * (dMV_C^i)^T \\ &= MV^i(P) + A(P) * (dMV_C^i)^T \end{aligned} \quad (10)$$

[0054] Denote Pic_{ref} as the reference picture and denote Pic_{cur} as the current picture and denote $Q = P + MV^i(P)$. Suppose we use MSE as the matching criterion, then we need to minimize:

$$\min_P \sum \left(Pic_{cur}(P) - Pic_{ref}(P + MV^{i+1}(P)) \right)^2$$

$$= \min \sum_P \left(Pic_{cur}(P) - Pic_{ref} \left(Q + A(P) * (dMV_C^i)^T \right) \right)^2 \quad (11)$$

[0055] Suppose $(dMV_C^i)^T$ is small enough, we can rewrite $Pic_{ref} \left(Q + A(P) * (dMV_C^i)^T \right)$ approximately as follows with 1th order Taylor expansion.

$$Pic_{ref} \left(Q + A(P) * (dMV_C^i)^T \right) \approx Pic_{ref}(Q) + Pic_{ref}'(Q) * A(P) * (dMV_C^i)^T \quad (12)$$

[0056] Wherein $Pic_{ref}'(Q) = \left[\frac{dPic_{ref}(Q)}{dx} \frac{dPic_{ref}(Q)}{dy} \right]$. Denote $E^{i+1}(P) = Pic_{cur}(P) - Pic_{ref}(Q)$,

$$\begin{aligned} & \min \sum_P \left(Pic_{cur}(P) - Pic_{ref}(Q) - Pic_{ref}'(Q) * A(P) * (dMV_C^i)^T \right)^2 \\ & = \min \sum_P \left(E^{i+1}(P) - Pic_{ref}'(Q) * A(P) * (dMV_C^i)^T \right)^2 \end{aligned} \quad (13)$$

[0057] we can derive dMV_C^i by setting the derivative of the error function to zero. Then can then calculate delta MV of the control points (0, 0) and (0, w) according to $A(P) * (dMV_C^i)^T$,

$$dMV_{(0,0)}^h = dMV_C^i[0] \quad (14)$$

$$dMV_{(0,w)}^h = dMV_C^i[1] * w + dMV_C^i[2] \quad (15)$$

$$dMV_{(0,0)}^v = dMV_C^i[2] \quad (16)$$

$$dMV_{(0,w)}^v = -dMV_C^i[3] * w + dMV_C^i[2] \quad (17)$$

[0058] Suppose such MVD derivation process is iterated by n times, then the final MVD is calculated as follows,

$$fdMV_{(0,0)}^h = \sum_{i=0}^{n-1} dMV_C^i[0] \quad (18)$$

$$fdMV_{(0,w)}^h = \sum_{i=0}^{n-1} dMV_C^i[1] * w + \sum_{i=0}^{n-1} dMV_C^i[2] \quad (19)$$

$$fdMV_{(0,0)}^v = \sum_{i=0}^{n-1} dMV_C^i[2] \quad (20)$$

$$fdMV_{(0,w)}^v = \sum_{i=0}^{n-1} -dMV_C^i[3] * w + \sum_{i=0}^{n-1} dMV_C^i[2] \quad (21)$$

[0059] As an example, e.g., predicting delta MV of control point (0, w), denoted by mvd_1 from delta MV of control point (0, 0), denoted by mvd_0 , now actually only $(\sum_{i=0}^{n-1} dMV_C^i[1] * w, -\sum_{i=0}^{n-1} -dMV_C^i[3] * w)$ is encoded $formvd_1$.

[0060] 2.1.3 AF_MERGE mode

[0061] When a CU is applied in AF_MERGE mode, it gets the first block coded with affine mode from the valid neighbour reconstructed blocks. And the selection order for the candidate block is from left, above, above right, left bottom to above left as shown in FIG. 5A. If the neighbour left bottom block A is coded in affine mode as shown in FIG. 5B, the motion vectors v_2 , v_3 and v_4 of the top left corner, above right corner and left bottom corner of the CU which contains the block A are derived. And the motion vector v_0 of the top left corner on the

current CU is calculated according to v_2 , v_3 and v_4 . Secondly, the motion vector v_1 of the above right of the current CU is calculated.

[0062] After the CPMV of the current CU v_0 and v_1 are derived, according to the simplified affine motion model Equation 1, the MVF of the current CU is generated. In order to identify whether the current CU is coded with AF_MERGE mode, an affine flag is signaled in the bitstream when there is at least one neighbour block is coded in affine mode.

[0063] In an example, which was planned to be adopted into VTM 3.0, an affine merge candidate list is constructed with following steps:

[0064] 1) Insert inherited affine candidates

[0065] Inherited affine candidate means that the candidate is derived from the affine motion model of its valid neighbor affine coded block. In the common base, as shown in FIG. 6, the scan order for the candidate positions is: A1, B1, B0, A0 and B2.

[0066] After a candidate is derived, full pruning process is performed to check whether same candidate has been inserted into the list. If a same candidate exists, the derived candidate is discarded.

[0067] 2) Insert constructed affine candidates

[0068] If the number of candidates in affine merge candidate list is less than MaxNumAffineCand (set to 5 in this contribution), constructed affine candidates are inserted into the candidate list. Constructed affine candidate means the candidate is constructed by combining the neighbor motion information of each control point.

[0069] The motion information for the control points is derived firstly from the specified spatial neighbors and temporal neighbor shown in FIG. 5B. CP k ($k=1, 2, 3, 4$) represents the k -th control point. A0, A1, A2, B0, B1, B2 and B3 are spatial positions for predicting CP k ($k=1, 2, 3$); T is temporal position for predicting CP4.

[0070] The coordinates of CP1, CP2, CP3 and CP4 is (0, 0), (W, 0), (H, 0) and (W, H), respectively, where W and H are the width and height of current block.

[0071] FIG. 6 shows an example of candidates position for affine merge mode

[0072] The motion information of each control point is obtained according to the following priority order:

[0073] For CP1, the checking priority is B2->B3->A2. B2 is used if it is available. Otherwise, if B2 is available, B3 is used. If both B2 and B3 are unavailable, A2 is used. If all the three candidates are unavailable, the motion information of CP1 cannot be obtained.

[0074] For CP2, the checking priority is B1->B0.

[0075] For CP3, the checking priority is A1->A0.

[0076] For CP4, T is used.

[0077] Secondly, the combinations of controls points are used to construct an affine merge candidate.

[0078] Motion information of three control points are needed to construct a 6-parameter affine candidate. The three control points can be selected from one of the following four combinations ({CP1, CP2, CP4}, {CP1, CP2, CP3}, {CP2, CP3, CP4}, {CP1, CP3, CP4}). Combinations {CP1, CP2, CP3}, {CP2, CP3, CP4}, {CP1, CP3, CP4} will be converted to a 6-parameter motion model represented by top-left, top-right and bottom-left control points.

[0079] Motion information of two control points are needed to construct a 4-parameter affine candidate. The two control points can be selected from one of the following six combinations ({CP1, CP4}, {CP2, CP3}, {CP1, CP2}, {CP2, CP4}, {CP1, CP3}, {CP3, CP4}). Combinations {CP1, CP4}, {CP2, CP3}, {CP2, CP4}, {CP1, CP3}, {CP3, CP4} will be converted to a 4-parameter motion model represented by top-left and top-right control points.

[0080] The combinations of constructed affine candidates are inserted into to candidate list as following order:

{CP1, CP2, CP3}, {CP1, CP2, CP4}, {CP1, CP3, CP4}, {CP2, CP3, CP4}, {CP1, CP2}, {CP1, CP3}, {CP2, CP3}, {CP1, CP4}, {CP2, CP4}, {CP3, CP4}

[0081] For reference list X (X being 0 or 1) of a combination, the reference index with highest usage ratio in the control points is selected as the reference index of list X, and motion vectors point to difference reference picture will be scaled.

[0082] After a candidate is derived, full pruning process is performed to check whether same candidate has been inserted into the list. If a same candidate exists, the derived candidate is discarded.

[0083] 3) Padding with zero motion vectors

[0084] If the number of candidates in affine merge candidate list is less than 5, zero motion vectors with zero reference indices are insert into the candidate list, until the list is full.

[0085] **2.2 Affine merge mode with prediction offsets**

[0086] In an example, UMVE is extended to affine merge mode, we will call this UMVE affine mode thereafter. The proposed method selects the first available affine merge candidate as a base predictor. Then it applies a motion vector offset to each control point's motion vector value

from the base predictor. If there's no affine merge candidate available, this proposed method will not be used.

[0087] The selected base predictor's inter prediction direction, and the reference index of each direction is used without change.

[0088] In the current implementation, the current block's affine model is assumed to be a 4-parameter model, only 2 control points need to be derived. Thus, only the first 2 control points of the base predictor will be used as control point predictors.

[0089] For each control point, a zero_MVD flag is used to indicate whether the control point of current block has the same MV value as the corresponding control point predictor. If zero_MVD flag is true, there's no other signaling needed for the control point. Otherwise, a distance index and an offset direction index is signaled for the control point.

[0090] A distance offset table with size of 5 is used as shown in the table below. Distance index is signaled to indicate which distance offset to use. The mapping of distance index and distance offset values is shown in FIG. 7.

Table - Distance offset table

Distance IDX	0	1	2	3	4
Distance-offset	1/2-pel	1-pel	2-pel	4-pel	8-pel

[0091] The direction index can represent four directions as shown below, where only x or y direction may have an MV difference, but not in both directions.

Offset Direction IDX	00	01	10	11
x-dir-factor	+1	-1	0	0
y-dir-factor	0	0	+1	-1

[0092] If the inter prediction is uni-directional, the signaled distance offset is applied on the offset direction for each control point predictor. Results will be the MV value of each control point.

[0093] For example, when base predictor is uni-directional, and the motion vector values of a control point is MVP (v_{px}, v_{py}). When distance offset and direction index are signaled, the motion vectors of current block's corresponding control points will be calculated as below.

$$MV(v_x, v_y) = MVP(v_{px}, v_{py}) + MV(x\text{-dir-factor} * \text{distance-offset}, y\text{-dir-factor} * \text{distance-offset});$$

[0094] If the inter prediction is bi-directional, the signaled distance offset is applied on the signaled offset direction for control point predictor’s L0 motion vector; and the same distance offset with opposite direction is applied for control point predictor’s L1 motion vector. Results will be the MV values of each control point, on each inter prediction direction.

[0095] For example, when base predictor is uni-directional, and the motion vector values of a control point on L0 is $MVP_{L0}(v_{0px}, v_{0py})$, and the motion vector of that control point on L1 is $MVP_{L1}(v_{1px}, v_{1py})$. When distance offset and direction index are signaled, the motion vectors of current block’s corresponding control points will be calculated as below.

$$MV_{L0}(v_{0x}, v_{0y}) = MVP_{L0}(v_{0px}, v_{0py}) + MV(x\text{-dir-factor} * \text{distance-offset}, y\text{-dir-factor} * \text{distance-offset});$$

$$MV_{L1}(v_{0x}, v_{0y}) = MVP_{L1}(v_{0px}, v_{0py}) + MV(- x\text{-dir-factor} * \text{distance-offset}, - y\text{-dir-factor} * \text{distance-offset});$$

[0096] **2.3 Ultimate motion vector expression**

[0097] In an example, ultimate motion vector expression (UMVE) is presented. UMVE is used for either skip or merge modes with a proposed motion vector expression method.

[0098] UMVE re-uses merge candidate as same as those included in the regular merge candidate list in VVC. Among the merge candidates, a base candidate can be selected, and is further expanded by the proposed motion vector expression method.

[0099] UMVE provides a new motion vector difference (MVD) representation method, in which a starting point, a motion magnitude and a motion direction are used to represent a MVD.

[00100] FIG. 8 shows an example of UMVE Search Process.

[00101] FIG. 9 shows examples of UMVE Search Points.

[00102] This proposed technique uses a merge candidate list as it is. But only candidates which are default merge type (MRG_TYPE_DEFAULT_N) are considered for UMVE’s expansion.

[00103] Base candidate index defines the starting point. Base candidate index indicates the best candidate among candidates in the list as follows.

Table 1. Base candidate IDX

Base candidate IDX	0	1	2	3
Nth MVP	1st MVP	2nd MVP	3rd MVP	4th MVP

[00104] If the number of base candidate is equal to 1, Base candidate IDX is not signaled.

[00105] Distance index is motion magnitude information. Distance index indicates the pre-defined distance from the starting point information. Pre-defined distance is as follows:

Table 2a. Distance IDX

Distance IDX	0	1	2	3	4	5	6	7
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	16-pel	32-pel

[00106] The distance IDX is binarized in bins with the truncated unary code in the entropy coding procedure as:

Table 2b: Distance IDX Binarization

Distance IDX		0	1	2	3	4	5	6	7
Bins		0	10	110	1110	11110	111110	1111110	1111111

[00107] In arithmetic coding, the first bin is coded with a probability context, and the following bins are coded with the equal-probability model, a.k.a. by-pass coding.

[00108] Direction index represents the direction of the MVD relative to the starting point. The direction index can represent of the four directions as shown below.

Table 3. Direction IDX

Direction IDX	00	01	10	11
x-axis	+	-	N/A	N/A
y-axis	N/A	N/A	+	-

[00109] UMVE flag is signaled right after sending a skip flag or merge flag. If skip or merge flag is true, UMVE flag is parsed. If UMVE flag is equal to 1, UMVE syntaxes are parsed. But, if not 1, AFFINE flag is parsed. If AFFINE flag is equal to 1, that is AFFINE mode, But, if not 1, skip/merge index is parsed for VTM's skip/merge mode.

[00110] Additional line buffer due to UMVE candidates is not needed. Because a skip/merge candidate of software is directly used as a base candidate. Using input UMVE index, the supplement of MV is decided right before motion compensation. There is no need to hold long line buffer for this.

[00111] In current common test condition, either the first or the second merge candidate in the merge candidate list could be selected as the base candidate.

[00112] UMVE is known as Merge with MVD (MMVD).

[00113] **2.4 Generalized Bi-prediction**

[00114] In conventional bi-prediction, the predictors from L0 and L1 are averaged to generate the final predictor using the equal weight 0.5. The predictor generation formula is shown as in Equ. (3)

$$P_{\text{TraditionalBiPred}} = (P_{L0} + P_{L1} + \text{RoundingOffset}) \gg \text{shiftNum}, \tag{1}$$

[00115] In Equ. (3), $P_{\text{TraditionalBiPred}}$ is the final predictor for the conventional bi-prediction, P_{L0} and P_{L1} are predictors from L0 and L1, respectively, and RoundingOffset and shiftNum are used to normalize the final predictor.

[00116] Generalized Bi-prediction (GBI) is proposed to allow applying different weights to predictors from L0 and L1. The predictor generation is shown in Equ. (4).

$$P_{\text{GBi}} = ((1-w_1) * P_{L0} + w_1 * P_{L1} + \text{RoundingOffset}_{\text{GBi}}) \gg \text{shiftNum}_{\text{GBi}}, \quad (2)$$

[00117] In Equ. (4), P_{GBi} is the final predictor of GBI. $(1-w_1)$ and w_1 are the selected GBI weights applied to the predictors of L0 and L1, respectively. $\text{RoundingOffset}_{\text{GBi}}$ and $\text{shiftNum}_{\text{GBi}}$ are used to normalize the final predictor in GBI.

[00118] The supported weights of w_1 is $\{-1/4, 3/8, 1/2, 5/8, 5/4\}$. One equal-weight set and four unequal-weight sets are supported. For the equal-weight case, the process to generate the final predictor is exactly the same as that in the conventional bi-prediction mode. For the true bi-prediction cases in random access (RA) condition, the number of candidate weight sets is reduced to three.

[00119] For advanced motion vector prediction (AMVP) mode, the weight selection in GBI is explicitly signaled at CU-level if this CU is coded by bi-prediction. For merge mode, the weight selection is inherited from the merge candidate. In this proposal, GBI supports DMVR to generate the weighted average of template as well as the final predictor for BMS-1.0.

[00120] 2.5 Adaptive motion vector difference resolution

[00121] In HEVC, motion vector differences (MVDs) (between the motion vector and predicted motion vector of a PU) are signaled in units of quarter luma samples when $\text{use_integer_mv_flag}$ is equal to 0 in the slice header. In VTM-3.0, a locally adaptive motion vector resolution (LAMVR) is introduced. In the JEM, MVD can be coded in units of quarter luma samples, integer luma samples or four luma samples. The MVD resolution is controlled at the coding unit (CU) level, and MVD resolution flags are conditionally signaled for each CU that has at least one non-zero MVD components.

[00122] For a CU that has at least one non-zero MVD components, a first flag is signaled to indicate whether quarter luma sample MV precision is used in the CU. When the first flag (equal to 1) indicates that quarter luma sample MV precision is not used, another flag is signaled to indicate whether integer luma sample MV precision or four luma sample MV precision is used.

[00123] When the first MVD resolution flag of a CU is zero, or not coded for a CU (meaning all MVDs in the CU are zero), the quarter luma sample MV resolution is used for the CU. When a

CU uses integer-luma sample MV precision or four-luma-sample MV precision, the MVPs in the AMVP candidate list for the CU are rounded to the corresponding precision.

[00124] In arithmetic coding, the first MVD resolution flag is coded with one of three probability contexts: C0, C1 or C2; while the second MVD resolution flag is coded with a fourth probability context: C3. The probability context C_x for the first MVD resolution flag is derived as (L represents the left neighbouring block and A represents the above neighbouring block):

[00125] If L is available, inter-coded, and its first MVD resolution flag is not equal to zero, x_L is set equal to 1; otherwise, x_L is set equal to 0.

[00126] If A is available, inter-coded, and its first MVD resolution flag is not equal to zero, x_A is set equal to 1; otherwise, x_A is set equal to 0.

[00127] x is set equal to x_L+x_A .

[00128] In the encoder, CU-level RD checks are used to determine which MVD resolution is to be used for a CU. That is, the CU-level RD check is performed three times for each MVD resolution. To accelerate encoder speed, the following encoding schemes are applied in the JEM.

[00129] ○ During RD check of a CU with normal quarter luma sample MVD resolution, the motion information of the current CU (integer luma sample accuracy) is stored. The stored motion information (after rounding) is used as the starting point for further small range motion vector refinement during the RD check for the same CU with integer luma sample and 4 luma sample MVD resolution so that the time-consuming motion estimation process is not duplicated three times.

[00130] ○ RD check of a CU with 4 luma sample MVD resolution is conditionally invoked. For a CU, when RD cost integer luma sample MVD resolution is much larger than that of quarter luma sample MVD resolution, the RD check of 4 luma sample MVD resolution for the CU is skipped.

[00131] In VTM-3.0, LAMVR is also known as Integer Motion Vector (IMV).

[00132] 2.6 Current picture referencing

[00133] Decoder aspect:

[00134] In this approach, the current (partially) decoded picture is considered as a reference picture. This current picture is put in the last position of reference picture list 0. Therefore, for a slice using the current picture as the only reference picture, its slice type is considered as a P slice. The bitstream syntax in this approach follows the same syntax structure for inter coding

while the decoding process is unified with inter coding. The only outstanding difference is that the block vector (which is the motion vector pointing to the current picture) always uses integer-pel resolution.

[00135] Changes from block level CPR_flag approach are:

- In encoder search for this mode, both block width and height are smaller than or equal to 16.
- Enable chroma interpolation when luma block vector is an odd integer number.
- Enable adaptive motion vector resolution (AMVR) for CPR mode when the SPS flag is on. In this case, when AMVR is used, a block vector can switch between 1-pel integer and 4-pel integer resolutions at block level.

[00136] Encoder aspect:

[00137] The encoder performs RD check for blocks with either width or height no larger than 16. For non-merge mode, the block vector search is performed using hash-based search first. If there is no valid candidate found from hash search, block matching based local search will be performed.

[00138] In the hash-based search, hash key matching (32-bit CRC) between the current block and a reference block is extended to all allowed block sizes. The hash key calculation for every position in current picture is based on 4x4 blocks. For the current block of a larger size, a hash key matching to a reference block happens when all its 4x4 blocks match the hash keys in the corresponding reference locations. If multiple reference blocks are found to match the current block with the same hash key, the block vector costs of each candidates are calculated and the one with minimum cost is selected.

[00139] In block matching search, the search range is set to be 64 pixels to the left and on top of current block, and the search range is restricted to be within the current CTU.

[00140] 2.7 Merge list design in one example

[00141] There are three different merge list construction processes supported in VVC:

[00142] 1) Sub-block merge candidate list: it includes ATMVP and affine merge candidates. One merge list construction process is shared for both affine modes and ATMVP mode. Here, the ATMVP and affine merge candidates may be added in order. Sub-block merge list size is signaled in slice header, and maximum value is 5.

[00143] 2) Uni-Prediction TPM merge list: For triangular prediction mode, one merge list construction process for the two partitions is shared even two partitions could select their own

merge candidate index. When constructing this merge list, the spatial neighbouring blocks and two temporal blocks of the block are checked. The motion information derived from spatial neighbours and temporal blocks are called regular motion candidates in our IDF. These regular motion candidates are further utilized to derive multiple TPM candidates. Please note the transform is performed in the whole block level, even two partitions may use different motion vectors for generating their own prediction blocks.

[00144] In some embodiments, uni-Prediction TPM merge list size is fixed to be 5.

[00145] 3) Regular merge list: For remaining coding blocks, one merge list construction process is shared. Here, the spatial/temporal/HMVP, pairwise combined bi-prediction merge candidates and zero motion candidates may be inserted in order. Regular merge list size is signaled in slice header, and maximum value is 6.

[00146] Sub-block merge candidate list

[00147] It is suggested that all the sub-block related motion candidates are put in a separate merge list in addition to the regular merge list for non-sub block merge candidates.

[00148] The sub-block related motion candidates are put in a separate merge list is named as 'sub-block merge candidate list'.

[00149] In one example, the sub-block merge candidate list includes affine merge candidates, and ATMVP candidate, and/or sub-block based STMVP candidate.

[00150] In this contribution, the ATMVP merge candidate in the normal merge list is moved to the first position of the affine merge list. Such that all the merge candidates in the new list (i.e., sub-block based merge candidate list) are based on sub-block coding tools.

[00151] An affine merge candidate list is constructed with following steps:

[00152] 1) Insert inherited affine candidates

[00153] Inherited affine candidate means that the candidate is derived from the affine motion model of its valid neighbor affine coded block. The maximum two inherited affine candidates are derived from affine motion model of the neighboring blocks and inserted into the candidate list. For the left predictor, the scan order is {A0, A1}; for the above predictor, the scan order is {B0, B1, B2}.

[00154] 2) Insert constructed affine candidates

[00155] If the number of candidates in affine merge candidate list is less than MaxNumAffineCand (set to 5), constructed affine candidates are inserted into the candidate list.

Constructed affine candidate means the candidate is constructed by combining the neighbor motion information of each control point.

[00156] The motion information for the control points is derived firstly from the specified spatial neighbors and temporal neighbor shown in Fig. 7. CP_k (k=1, 2, 3, 4) represents the k-th control point. A₀, A₁, A₂, B₀, B₁, B₂ and B₃ are spatial positions for predicting CP_k (k=1, 2, 3); T is temporal position for predicting CP₄.

[00157] The coordinates of CP₁, CP₂, CP₃ and CP₄ is (0, 0), (W, 0), (H, 0) and (W, H), respectively, where W and H are the width and height of current block.

[00158] The motion information of each control point is obtained according to the following priority order:

[00159] For CP₁, the checking priority is B₂→B₃→A₂. B₂ is used if it is available. Otherwise, if B₂ is available, B₃ is used. If both B₂ and B₃ are unavailable, A₂ is used. If all the three candidates are unavailable, the motion information of CP₁ cannot be obtained.

[00160] For CP₂, the checking priority is B₁→B₀.

[00161] For CP₃, the checking priority is A₁→A₀.

[00162] For CP₄, T is used.

[00163] Secondly, the combinations of controls points are used to construct an affine merge candidate.

[00164] Motion information of three control points are needed to construct a 6-parameter affine candidate. The three control points can be selected from one of the following four combinations ({CP₁, CP₂, CP₄}, {CP₁, CP₂, CP₃}, {CP₂, CP₃, CP₄}, {CP₁, CP₃, CP₄}). Combinations {CP₁, CP₂, CP₃}, {CP₂, CP₃, CP₄}, {CP₁, CP₃, CP₄} will be converted to a 6-parameter motion model represented by top-left, top-right and bottom-left control points.

[00165] Motion information of two control points are needed to construct a 4-parameter affine candidate. The two control points can be selected from one of the two combinations ({CP₁, CP₂}, {CP₁, CP₃}). The two combinations will be converted to a 4-parameter motion model represented by top-left and top-right control points.

[00166] The combinations of constructed affine candidates are inserted into to candidate list as following order:

[00167] {CP₁, CP₂, CP₃}, {CP₁, CP₂, CP₄}, {CP₁, CP₃, CP₄}, {CP₂, CP₃, CP₄}, {CP₁, CP₂}, {CP₁, CP₃}

[00168] The available combination of motion information of CPs is only added to the affine merge list when the CPs have the same reference index.

[00169] 3) Padding with zero motion vectors

[00170] If the number of candidates in affine merge candidate list is less than 5, zero motion vectors with zero reference indices are insert into the candidate list, until the list is full.

[00171] **2.8 MMVD with affine merge candidate in an example**

[00172] For example the MMVD idea is applied on affine merge candidates (named as Affine merge with prediction offset). It is an extension of A MVD (or named as “distance”, or “offset”) is signaled after an affine merge candidate (known as the is signaled. The all CPMVs are added with the MVD to get the new CPMVs. The distance table is specified as

Distance IDX	0	1	2	3	4
Distance-offset	1/2-pel	1-pel	2-pel	4-pel	8-pel

[00173] In some embodiments, a POC distance based offset mirroring method is used for Bi-prediction. When the base candidate is bi-predicted, the offset applied to L0 is as signaled, and the offset on L1 depends on the temporal position of the reference pictures on list 0 and list 1.

[00174] If both reference pictures are on the same temporal side of the current picture, the same distance offset and same offset directions are applied for CPMVs of both L0 and L1.

[00175] When the two reference pictures are on different sides of the current picture, the CPMVs of L1 will have the distance offset applied on the opposite offset direction.

[00176] **3. Examples of problems solved by the disclosed embodiments**

[00177] There are some potential problems in the design of MMVD:

- The coding/parsing procedure for UMVE information may be not efficient since it uses truncated unary binarization method for coding distance (MVD precision) information and fixed length with bypass coding for direction index. That is based on the assumption that the ¼-pel precision takes the highest percentage. However, it is not true for all kinds of sequences.
- The design of the possible distances may be not efficient.

[00178] **4. Examples of techniques implemented by various embodiments**

[00179] The list below should be considered as examples to explain general concepts. These inventions should not be interpreted in a narrow way. Furthermore, these techniques can be combined in any manner.

Parsing on distance index (e.g., MVD precision index)

1. It is proposed that the distance index (DI) used in UMVE is not binarized with the truncated unary code.
 - a. In one example, DI may be binarized with fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
2. The distance index may be signaled with more than one syntax elements.
3. It is proposed to classify the set of allowed distances to multiple sub-sets, e.g., K sub-sets (K is larger than 1). A sub-set index (1st syntax) is firstly signaled, followed by distance index (2nd syntax) within the sub-set.
 - a. For example, **mmvd_distance_subset_idx** is first signaled, followed by **mmvd_distance_idx_in_subset**.
 - i. In one example, the **mmvd_distance_idx_in_subset** can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
 - (i) Specifically, **mmvd_distance_idx_in_subset** can be binarized as a flag if there are only two possible distances in the sub-set.
 - (ii) Specifically, **mmvd_distance_idx_in_subset** is not signaled if there is only one possible distance in the sub-set.
 - (iii) Specifically, the maximum value is set to be the number of possible distance in the sub-set minus 1 if **mmvd_distance_idx_in_subset** is binarized as a truncated code.
 - b. In one example, there are two sub-sets (e.g., K=2).
 - i. In one example, one sub-set includes all fractional MVD precisions (e.g., 1/4-pel, 1/2-pel). The other sub-set includes all integer MVD precisions (e.g., 1-pel, 2-pel, 4-pel, 8-pel, 16-pel, 32-pel).
 - ii. In one example, one sub-set may only have one distance (e.g., 1/2-pel), and the other sub-set has all the remaining distances.
 - c. In one example, there are three sub-sets (e.g., K=3).
 - i. In one example, a first sub-set includes fractional MVD precisions (e.g., 1/4-pel, 1/2-pel); a second sub-set includes integer MVD precisions less than 4-pel (e.g. 1-pel, 2-pel) ; and a third sub-set includes all other MVD precisions (e.g. 4-pel, 8-pel, 16-pel, 32-pel).

- d. In one example, there are K sub-sets and the number of K is set equal to that allowed MVD precisions in LAMVR.
 - i. Alternatively, furthermore, the signaling of sub-set index may reuse what is done for LAMVR (e.g., reusing the way to derive context offset index; reusing the contexts, etc., al)
 - ii. Distances within a sub-set may be decided by the associated LAMVR index (e.g., AMVR_mode in specifications).
- e. In one example, how to define sub-sets and/or how many sub-sets may be pre-defined or adaptively changed on-the-fly.
- f. In one example, the 1st syntax may be coded with truncated unary, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
- g. In one example, the 2nd syntax may be coded with truncated unary, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
- h. In one example, the sub-set index (e.g., 1st syntax) may be not explicitly coded in the bitstream. Alternatively, furthermore, the sub-set index may be derived on-the-fly, e.g., based on coded information of current block (e.g., block dimension) and/or previously coded blocks.
- i. In one example, the distance index within a sub-set (e.g., 2nd syntax) may be not explicitly coded in the bitstream.
 - i. In one example, when the sub-set only have one distance, there is no need to further signal the distance index.
 - ii. Alternatively, furthermore, the 2nd syntax may be derived on-the-fly, e.g., based on coded information of current block (e.g., block dimension) and/or previously coded blocks.
- j. In one example, a **first resolution bin** is signaled to indicate whether DI is smaller than a predefined number T or not. Alternatively, a **first resolution bin** is signaled to indicate whether the distance is smaller than a predefined number or not.

- i. In one example, two syntax elements are used to represent the distance index. **mmvd_resolution_flag** is first signaled followed by **mmvd_distance_idx_in_subset**.
- ii. In one example, three syntax elements are used to represent the distance index. **mmvd_resolution_flag** is first signaled followed by **mmvd_short_distance_idx_in_subset** when **mmvd_resolution_flag** is equal to 0 or **mmvd_long_distance_idx_in_subset** when **mmvd_resolution_flag** is equal to 1.
- iii. In one example, the distance index number T corresponds to 1-Pel distance. For example, T = 2 with the Table 2a defined in VTM-3.0.
- iv. In one example, the distance index number T corresponds to 1/2-Pel distance. For example, T = 1 with the Table 2a defined in VTM-3.0.
- v. In one example, the distance index number T corresponds to W-Pel distance. For example, T = 3 corresponding to 2-Pel distance with the Table 2a defined in VTM-3.0.
- vi. In one example, the **first resolution bin** is equal to 0 if DI is smaller than T. Alternatively, the **first resolution bin** is equal to 1 if DI is smaller than T.
- vii. In one example, if DI is smaller than T, a code for **short distance index** is further signaled after the **first resolution bin** to indicate the value of DI.
 - (i) In one example, DI is signaled. DI can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
 - a. When DI is binarized as a truncated code such as truncated unary code, the maximum coded value is T-1.
 - (ii) In one example, $S=T-1-DI$ is signaled. T-1-DI can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
 - a. When T-1-DI is binarized as a truncated code such as truncated unary code, the maximum coded value is T-1.

- b. After S is parsed, DI is reconstructed as $DI = T - S - 1$.
- viii. In one example, if DI is not smaller than T, a code for **long distance index** is further signaled after the **first resolution bin** to indicate the value of DI.
- (i) In one example, $B = DI - T$ is signaled. $DI - T$ can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
- a. When $DI - T$ is binarized as a truncated code such as truncated unary code, the maximum coded value is $D_{Max} - T$, where D_{Max} is the maximum allowed distance, such as 7 in VTM-3.0.
- b. After B is parsed, DI is reconstructed as $DI = B + T$.
- (ii) In one example, $B' = D_{Max} - DI$ is signaled, where D_{Max} is the maximum allowed distance, such as 7 in VTM-3.0. $D_{Max} - DI$ can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
- a. When $D_{Max} - DI$ is binarized as a truncated code such as truncated unary code, the maximum coded value is $D_{Max} - T$, where D_{Max} is the maximum allowed distance, such as 7 in VTM-3.0.
- b. After B' is parsed, DI is reconstructed as $DI = D_{Max} - B'$.
- k. In one example, a **first resolution bin** is signaled to indicate whether DI is greater than a predefined number T or not. Alternatively, a **first resolution bin** is signaled to indicate whether the distance is greater than a predefined number or not.
- i. In one example, the distance index number T corresponds to 1-Pel distance. For example, $T = 2$ with the Table 2a defined in VTM-3.0.
- ii. In one example, the distance index number T corresponds to 1/2-Pel distance. For example, $T = 1$ with the Table 2a defined in VTM-3.0.

- iii. In one example, the distance index number T corresponds to W -Pel distance. For example, $T = 3$ corresponding to 2-Pel distance with the Table 2a defined in VTM-3.0.
- iv. In one example, the **first resolution bin** is equal to 0 if DI is greater than T . Alternatively, the **first resolution bin** is equal to 1 if DI is greater than T .
- v. In one example, if DI is not greater than T , a code for **short distanceindex** is further signaled after the **first resolution bin** to indicate the value of DI .
 - (i) In one example, DI is signaled. DI can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
 - a. When DI is binarized as a truncated code such as truncated unary code, the maximum coded value is T .
 - (ii) In one example, $S=T-DI$ is signaled. $T-DI$ can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
 - a. When $T-DI$ is binarized as a truncated code such as truncated unary code, the maximum coded value is T .
 - b. After S is parsed, DI is reconstructed as $DI=T-S$.
- vi. In one example, if DI is greater than T , a code for **long distanceindex** is further signaled after the **first resolution bin** to indicate the value of DI .
 - (i) In one example, $B=DI-1-T$ is signaled. $DI-1-T$ can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
 - a. When $DI-1-T$ is binarized as a truncated code such as truncated unary code, the maximum coded value is $D_{Max} - 1 - T$, where D_{Max} is the maximum allowed distance, such as 7 in VTM-3.0.

- b. After B is parsed, DI is reconstructed as $DI = B+T+1$.
- (ii) In one example, $B' = DMax-DI$ is signaled, where DMax is the maximum allowed distance, such as 7 in VTM-3.0. DMax-DI can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
 - a. When DMax-DI is binarized as a truncated code such as truncated unary code, the maximum coded value is $DMax - 1 - T$, where DMax is the maximum allowed distance, such as 7 in VTM-3.0.
 - b. After B' is parsed, DI is reconstructed as $DI = DMax-B'$.
- I. Several possible binarization methods for the distance index are: (It should be noted that two binarization methods should be considered as identical if a process to change all "1" in one method to be "0" and all "0" in one method to be "1", will produce the same codewords as the other method.)

Distance IDX	0	1	2	3	4	5	6	7
Bins	01	00	10	110	1110	11110	111110	111111
Bins	01	00	11	101	1001	10001	100001	100000
Bins	00	01	10	110	1110	11110	111110	111111
Bins	00	01	11	101	1001	10001	100001	100000
Bins	11	10	00	010	0110	01110	011110	011111
Bins	11	10	01	001	0001	00001	000001	000000
Bins	10	11	00	010	0110	01110	011110	011111
Bins	10	11	01	001	0001	00001	000001	000000
Bins	010	0110	0111	10	110	1110	11110	11111
Bins	0111	0110	010	10	110	1110	11110	11111

- 4. It is proposed that one or more probability contexts are used to code the 1st syntax
 - a. In one example, the 1st syntax is the **first resolution bin** mentioned above.
 - b. In one example, which probability context is used is derived from the **first resolution bins** of neighbouring blocks.

- c. In one example, which probability context is used is derived from the **LAMVR values** of neighbouring blocks (e.g., AMVR_mode values).
5. It is proposed that one or more probability contexts are used to code the 2nd syntax.
- a. In one example, the 2nd syntax is the **short distance index** mentioned above.
 - i. In one example, the first bin to code the **short distance index** is coded with a probability context, and other bins are by-pass coded.
 - ii. In one example, the first N bins to code the **short distance index** are coded with probability contexts, and other bins are by-pass coded.
 - iii. In one example, all bins to code the **short distance index** are coded with probability contexts.
 - iv. In one example, different bins may have different probability contexts.
 - v. In one example, several bins share a single probability context.
 - (i) In one example, these bins are consecutive.
 - vi. In one example, which probability context is used is derived from the **short distance indices** of neighbouring blocks.
 - b. In one example, the 2nd syntax is **long distance index** mentioned above.
 - i. In one example, the first bin to code the **long distance index** is coded with a probability context, and other bins are by-pass coded.
 - ii. In one example, the first N bins to code the **long distance index** are coded with probability contexts, and other bins are by-pass coded.
 - iii. In one example, all bins to code the **long distance index** are coded with probability contexts.
 - iv. In one example, different bins may have different probability contexts.
 - v. In one example, several bins share a single probability context.
 - (i) In one example, these bins are consecutive.
 - vi. In one example, which probability context is used is derived from the **long distance indices** of neighbouring blocks.

Interaction with LAMVR

6. It is proposed that the 1st syntax (e.g., **first resolution bin**) is coded depending on the probability models used to code the LAMVR information.

- a. In one example, the **first resolution bin** is coded with the same manner (e.g., shared context, or same context index derivation method but with neighboring blocks' LAMVR information replaced by MMVD information) as coding the first MVD resolution flag.
 - i. In one example, which probability context is used to code the **first resolution bin** is derived from the LAMVR information of neighbouring blocks.
 - (i) In one example, which probability context is used to code the **first resolution bin** is derived from the first MVD resolution flags of neighbouring blocks.
 - b. Alternatively, the first MVD resolution flag is coded and serve as the **first resolution bin** when the distance index is coded.
 - c. In one example, which probability model is used to code the **first resolution bin** may depend on the coded LAMVR information.
 - i. For example, which probability model is used to code the **first resolution bin** may depend on the MV resolution of neighbouring blocks.
7. It is proposed that the first bin to code the **short distance index** is coded with a probability context.
- a. In one example, the first bin to code the **short distance index** is coded with the same manner (e.g., shared context, or same context index derivation method but with neighboring blocks' LAMVR information replaced by MMVD information) as coding the second MVD resolution flag.
 - b. Alternatively, the second MVD resolution flag is coded and serve as the first bin to code the **short distance index** when the distance index is coded.
 - c. In one example, which probability model is used to code the first bin to code the **short distance index** may depend on the coded LAMVR information.
 - i. For example, which probability model is used to code the first bin to code the **short distance index** may depend on the MV resolution of neighbouring blocks.
8. It is proposed that the first bin to code the **long distance index** is coded with a probability context.

- a. In one example, the first bin to code the **long distance index** is coded with the same manner (e.g., shared context, or same context index derivation method but with neighboring blocks' LAMVR information replaced by MMVD information) as coding the second MVD resolution flag.
 - b. Alternatively, the second MVD resolution flag is coded and serve as the first bin to code the **long distance index** when the distance index is coded.
 - c. In one example, which probability model is used to code the first bin to code the **long distance index** may depend on the coded LAMVR information.
 - i. For example, which probability model is used to code the first bin to code the **long distance index** may depend on the MV resolution of neighbouring blocks.
9. For the LAMVR mode, in arithmetic coding, the first MVD resolution flag is coded with one of three probability contexts: C0, C1 or C2; while the second MVD resolution flag is coded with a forth probability context: C3. Examples to derive the probability context to code the distance index is described as below.
- a. The probability context C_x for the **first resolution bin** is derived as (L represents the left neighbouring block and A represents the above neighbouring block):
 If L is available, inter-coded, and its first MVD resolution flag is not equal to zero, x_L is set equal to 1; otherwise, x_L is set equal to 0.
 If A is available, inter-coded, and its first MVD resolution flag is not equal to zero, x_A is set equal to 1; otherwise, x_A is set equal to 0.
 x is set equal to x_L+x_A.
 - b. The probability context for the first bin to code the **long distance index** is C3.
 - c. The probability context for the first bin to code the **short distance index** is C3.
10. It is proposed that the LAMVR MVD resolution is signaled when MMVD mode is applied.
- a. It is proposed that the syntax used LAMVR MVD resolution signaling is reused when coding the side information of MMVD mode.
 - b. When the signaled LAMVR MVD resolution is 1/4-pel, a **short distance index** is signaled to indicate the MMVD distance in a first sub-set. For example, the **short distance index** can be 0 or 1, to represent the MMVD distance to be 1/4-pel or 1/2 -pel, respectively.

- c. When the signaled LAMVR MVD resolution is 1-pel, a **medium distance index** is signaled to indicate the MMVD distance in a second sub-set. For example, the **medium distance index** can be 0 or 1, to represent the MMVD distance to be 1-pel or 2-pel, respectively.
- d. When the signaled LAMVR MVD resolution is 4-pel in a third sub-set, a **long distance index** is signaled to indicate the MMVD distance. For example, the **medium distance index** can be X, to represent the MMVD distance to be $(4 << X)$ -pel.
- e. In the following disclosure, a sub-set distance index may refer to a **short distance index**, ora **medium distance index**, or a **long distance index**.
 - i. In one example, the sub-set distance index can be binarized with unary code, truncated unary code, fixed-length code, Exponential-Golomb code, truncated Exponential-Golomb code, Rice code, or any other codes.
 - (i) Specifically, a sub-set distance index can be binarized as a flag if there are only two possible distances in the sub-set.
 - (ii) Specifically, a sub-set distance index is not signaled if there is only one possible distance in the sub-set.
 - (iii) Specifically, the maximum value is set to be the number of possible distance in the sub-set minus 1 if a sub-set distance index is binarized as a truncated code.
 - ii. In one example, the first bin to code the sub-set distance index is coded with a probability context, and other bins are by-pass coded.
 - iii. In one example, the first N bins to code the sub-set distance index are coded with probability contexts, and other bins are by-pass coded.
 - iv. In one example, all bins to code the sub-set distance index are coded with probability contexts.
 - v. In one example, different bins may have different probability contexts.
 - vi. In one example, several bins share a single probability context.
 - (i) In one example, these bins are consecutive.
 - vii. It is proposed that one distance cannot appear in two different distance sub-sets.
 - viii. In one example, there may be more distances can be signaled in the short distance sub-set.

- (i) For example, the distances signaled in the short distance sub-set must be in a sub-pel, but not an integer-pel. For example, 5/4-pel, 3/2-pel 7/4-pel may be in the short distance sub-set, but 3-pel cannot be in the short distance sub-set.
 - ix. In one example, there may be more distances can be signaled in the medium distance sub-set.
 - (i) For example, the distances signaled in the medium distance sub-set must be integer-pel but not in a form of $4N$, where N is an integer. For example, 3-pel, 5-pel may be in the medium distance sub-set, but 24-pel cannot be in the medium distance sub-set.
 - x. In one example, there may be more distances can be signaled in the long distance sub-set.
 - (i) For example, the distances signaled in the long distance sub-set must be integer-pel in a form of $4N$, where N is an integer. For example, 4-pel, 8-pel, 16-pel or 24-pel may be in the long distance sub-set.
11. It is proposed that a variable to store the MV resolution of the current block may be decided by the UMVE distance.
- a. In one example, the MV resolution of the current block is set to be 1/4-Pel if the UMVE distance $< T1$ or $\leq T1$.
 - b. In one example, the first and second MVD resolution flags of the current block are set to be 0 if the UMVE distance $< T1$ or $\leq T1$.
 - c. In one example, the MV resolution of the current block is set to be 1-Pel if the UMVE distance $> T1$ or $\geq T1$.
 - d. In one example, the first MVD resolution flag of the current block is set to be 1, and the second MVD resolution flag of the current block is set to be 0, if the UMVE distance $> T1$ or $\geq T1$.
 - e. In one example, the MV resolution of the current block is set to be 4-Pel if the UMVE distance $> T2$ or $\geq T2$.
 - f. In one example, the first and second MVD resolution flags of the current block are set to be 1, if the UMVE distance $> T2$ or $\geq T2$.
 - g. In one example, the MV resolution of the current block is set to be 1-Pel if the UMVE distance $> T1$ or $\geq T1$ and the UMVE distance $< T2$ or $\leq T2$.
 - h. In one example, the first MVD resolution flag of the current block is set to be 1, and the second MVD resolution flag of the current block is set to be 0, if the UMVE distance $> T1$ or $\geq T1$ and the UMVE distance $< T2$ or $\leq T2$.
 - i. $T1$ and $T2$ can be any numbers. For example, $T1 = 1\text{-Pel}$ and $T2=4\text{-Pel}$.

12. It is proposed that a variable to store the MV resolution of the current block may be decided by the UMVE distance index.
- a. In one example, the MV resolution of the current block is set to be $\frac{1}{4}$ -Pel if the UMVE distance index $< T1$ or $\leq T1$.
 - b. In one example, the first and second MVD resolution flags of the current block are set to be 0 if the UMVE distance index $< T1$ or $\leq T1$.
 - c. In one example, the MV resolution of the current block is set to be 1-Pel if the UMVE distance index $> T1$ or $\geq T1$.
 - d. In one example, the first MVD resolution flag of the current block is set to be 1, and the second MVD resolution flag of the current block is set to be 0, if the UMVE distance index $> T1$ or $\geq T1$.
 - e. In one example, the MV resolution of the current block is set to be 4-Pel if the UMVE distance index $> T2$ or $\geq T2$.
 - f. In one example, the first and second MVD resolution flags of the current block are set to be 1, if the UMVE distance index $> T2$ or $\geq T2$.
 - g. In one example, the MV resolution of the current block is set to be 1-Pel if the UMVE distance $> T1$ or $\geq T1$ and the UMVE distance index $< T2$ or $\leq T2$.
 - h. In one example, the first MVD resolution flag of the current block is set to be 1, and the second MVD resolution flag of the current block is set to be 0, if the UMVE distance index $> T1$ or $\geq T1$ and the UMVE distance index $< T2$ or $\leq T2$.
 - i. $T1$ and $T2$ can be any numbers. For example, $T1 = 2$ and $T2 = 3$, or $T1 = 2$ and $T2 = 4$;
13. The variable to store MV resolution of UMVE coded blocks may be utilized for coding following blocks which are coded with LAMVR mode.
- a. Alternatively, the variable to store MV resolution of UMVE coded blocks may be utilized for coding following blocks which are coded with UMVE mode.
 - b. Alternatively, the MV precisions of LAMVR-coded blocks may be utilized for coding following UMVE-coded blocks.
14. The above bullets may be also applicable to coding the direction index.

Mapping between distance index and distance

15. It is proposed that the relationship between the distance index (DI) and the distance may not be the exponential relationship as in VTM-3.0. (Distance = $\frac{1}{4}$ -Pel $\times 2^{DI}$)
- a. In one example, the mapping may be piece-wised.
 - i. For example, Distance = $f_1(DI)$ when $T_0 \leq DI < T_1$, Distance = $f_2(DI)$ when $T_1 \leq DI < T_2$, ... Distance = $f_n(DI)$ when $T_{n-1} \leq DI < T_n$
 - (i) For example, Distance = $\frac{1}{4}$ -Pel $\times 2^{DI}$ when $DI < T_1$, Distance = $a \times DI + b$ when $T_1 \leq DI < T_2$, and Distance = $c \times 2^{DI}$ when $DI \geq T_2$. In one example, $T_1 = 4$, $a = 1$, $b = -1$, $T_2 = 6$, $c = 1/8$.
16. It is proposed that distance table size may be larger than 8, such as 9, 10, 12, 16.
17. It is proposed that distance shorter than 1/4-pel may be included in the distance table, such as 1/8-pel, 1/16-pel or 3/8-pel.
18. It is proposed than distances not in the form of 2^x -pel may be included in the distance table, such as 3-pel, 5-pel, 6-pel, etc.
19. It is proposed that the distance table may be different for different directions.
- a. Correspondingly, the parsing process for the distance index may be different for different directions.
 - b. In one example, four directions with the direction index 0, 1, 2 and 3 have different distance tables.
 - c. In one example, two x-directions with the direction index 0 and 1 have the same distance table.
 - d. In one example, two y-directions with the direction index 2 and 3 have the same distance table.
 - e. In one example, x-directions and y-directions may have two different distance tables.
 - i. Correspondingly, the parsing process for the distance index may be different for x-directions and y-directions.
 - ii. In one example, the distance table for y-directions may have less possible distances than the distance table for x-directions.
 - iii. In one example, the shortest distance in the distance table for y-directions may be shorter than the shortest distance in the distance table for x-directions.

- iv. In one example, the longest distance in the distance table for y-directions may be shorter than the longest distance in the distance table for x-directions.
20. It is proposed that different distance tables may be used for different block width and/or height.
- a. In one example, different distance tables may be used for different block width when the direction is along the x-axis.
 - b. In one example, different distance tables may be used for different block height when the direction is along the y-axis.
21. It is proposed that different distance tables may be used when the POC distance is different. POC difference is calculated as $|\text{POC of the current picture} - \text{POC of the reference picture}|$.
22. It is proposed that different distance tables may be used for different base candidates.
23. It is proposed that for the ratio of two distances (MVD precisions) with consecutive indices is not fixed to be 2.
- a. In one example, the ratio of two distances (MVD precisions) with consecutive indices is fixed to M (e.g., $M = 4$).
 - b. In one example, the delta of two distances (MVD precisions) with consecutive indices (instead of ratios) may be fixed for all indices. Alternatively, the delta of two distances (MVD precisions) with consecutive indices may be different for different indices.
 - c. In one example, the ratio of two distances (MVD precisions) with consecutive indices may be different for different indices.
 - i. In one example, the set of distances such as {1-pel, 2-pel, 4-pel, 8-pel, 16-pel, 32-pel, 48-pel, 64-pel} may be used.
 - ii. In one example, the set of distances such as {1-pel, 2-pel, 4-pel, 8-pel, 16-pel, 32-pel, 64-pel, 96-pel} may be used.
 - iii. In one example, the set of distances such as {1-pel, 2-pel, 3-pel, 4-pel, 5-pel, 16-pel, 32-pel} may be used.
24. The signaling of MMVD side information may be done in the following way:

- a. When the current block is inter and non-merge (which may include e.g., non-skip, non-sub-block, non-triangular, non-MHIntra) mode, a MMVD flag may be firstly signaled, followed by a sub-set index of distance, a distance index within the sub-set, a direction index. Here, MMVD is treated as a different mode from merge mode.
- b. Alternatively, when the current block is merge mode, a MMVD flag may be further signaled, followed by a sub-set index of distance, a distance index within the sub-set, a direction index. Here, MMVD is treated as a special merge mode.

25. The direction of MMVD and the distance of MMVD may be signaled jointly.

- a. In one example, whether and how to signal MMVD distance may depend on MMVD direction.
- b. In one example, whether and how to signal MMVD direction may depend on MMVD distance.
- c. In one example, a joint codeword is signaled with one or more syntax elements. The MMVD distance and MMVD direction can be derived from the code word. For example, the codeword is equal to MMVD distance index + MMVD direction index * 7. In another example, the MMVD a codeword table is designed. Each codeword corresponds to a unique combination of MMVD distance and MMVD direction.

26. Some exemplary UMVE distances tables are listed below:

- a. The distance tables size is 9:

Distance IDX	0	1	2	3	4	5	6	7	8
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	8-pel	16-pel	32-pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	3-pel	4-pel	8-pel	16-pel	32-pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	6-pel	8-pel	16-pel	32-pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	8-pel	12-pel	16-pel	32-pel
Pixel	1/4-	1/2-	1-pel	2-pel	4-pel	8-pel	16-pel	24-pel	32-pel

distance	pel	pel								
-----------------	-----	-----	--	--	--	--	--	--	--	--

b. The distance tables size is 10:

Distance IDX	0	1	2	3	4	5	6	7	8	9
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	8-pel	16- pel	32- pel	64- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	3-pel	4-pel	8-pel	16- pel	32- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	6-pel	8-pel	16- pel	32- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	8-pel	12- pel	16- pel	32- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	8-pel	16- pel	24- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	3-pel	4-pel	6-pel	8-pel	16- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	6-pel	8-pel	12- pel	16- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	3-pel	4-pel	8-pel	12- pel	16- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	3-pel	4-pel	8-pel	16- pel	24- pel	32- pel

a. The distance tables size is 12:

Distance IDX	0	1	2	3	4	5	6	7	8	9	10	11
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1- pel	2- pel	3- pel	4- pel	6- pel	8- pel	16- pel	32- pel	64- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1- pel	2- pel	3- pel	4- pel	8- pel	12- pel	16- pel	24- pel	32- pel
Pixel	1/4- pel	1/2- pel	1-pel	2- pel	3- pel	4- pel	6- pel	8- pel	12- pel	16- pel	24- pel	32- pel

distance	pel	pel		pel	pel	pel	pel	pel	pel	pel	pel	pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	3-pel	4-pel	6-pel	8-pel	12-pel	16-pel	32-pel	64-pel

27. It is proposed that MMVD distance may be signaled in a granularity-signaling method.

The distance is firstly signaled by index with a rough granularity, followed by one or more indices with finer granularities.

- a. For example, a first index F_1 represents distances in an ordered set M_1 ; a second index F_2 represents distances in an ordered set M_2 . The final distance is calculated as, such as $M_1[F_1]+M_2[F_2]$.
- b. For example, a first index F_1 represents distances in an ordered set M_1 ; a second index F_2 represents distances in an ordered set M_2 ; and so on, until a n th index F_n represents distances in an ordered set M_n . The final distance is calculated as $M_1[F_1]+M_2[F_2]+ \dots + M_n[F_n]$
- c. For example, the signaling or binarization of F_k may depend on the signaled F_{k-1} .
 - i. In one example, an entry in $M_k[F_k]$ must be smaller than $M_{k-1}[F_{k-1}+1]- M_{k-1}[F_{k-1}]$ when F_{k-1} is not the largest index referring to M_{k-1} . $1 < k \leq n$.
- d. For example, the signaling or binarization of F_k may depend on the signaled F_s for all $1 \leq s < k$.
 - i. In one example, an entry in $M_k[F_k]$ must be smaller than $M_s [F_s+1]- M_s[F_s]$ for all $1 \leq s < k$, when $1 < k \leq n$.
- e. In one example, if F_k is the largest index referring to M_k , then no further F_{k+1} is signaled, and the final distance is calculated as $M_1[F_1]+M_2[F_2]+ \dots + M_k[F_k]$, where $1 \leq k \leq n$
- f. In one example, the entries in $M_k[F_k]$ may depend on the signaled F_{k-1} .
- g. In one example, the entries in $M_k[F_k]$ may depend on the signaled F_s for all $1 \leq s < k$.
- h. For example, $n=2$. $M_1 = \{1/4\text{-pel}, 1\text{-pel}, 4\text{-pel}, 8\text{-pel}, 16\text{-pel}, 32\text{-pel}, 64\text{-pel}, 128\text{-pel}\}$,
 - i. $M_2 = \{0\text{-pel}, 1/4\text{-pel} \}$ when $F_1 = 0$ ($M_1[F_1]=1/4\text{-pel}$);
 - ii. $M_2 = \{0\text{-pel}, 1\text{-pel}, 2\text{-pel}\}$ when $F_1 = 1$ ($M_1[F_1]=1\text{-pel}$);
 - iii. $M_2 = \{0\text{-pel}, 1\text{-pel}, 2\text{-pel}, 3\text{-pel}\}$ when $F_1 = 2$ ($M_1[F_1]=4\text{-pel}$);

- iv. $M_2 = \{0\text{-pel}, 2\text{-pel}, 4\text{-pel}, 6\text{-pel}\}$ when $F_1 = 3$ ($M_1[F_1]=8\text{-pel}$);
- v. $M_2 = \{0\text{-pel}, 4\text{-pel}, 8\text{-pel}, 12\text{-pel}\}$ when $F_1 = 4$ ($M_1[F_1]=16\text{-pel}$);
- vi. $M_2 = \{0\text{-pel}, 8\text{-pel}, 16\text{-pel}, 24\text{-pel}\}$ when $F_1 = 5$ ($M_1[F_1]=32\text{-pel}$);
- vii. $M_2 = \{0\text{-pel}, 16\text{-pel}, 32\text{-pel}, 48\text{-pel}\}$ when $F_1 = 6$ ($M_1[F_1]=32\text{-pel}$);

Slice/Picture Level control

28. It is proposed that how to signal the MMVD side information (e.g., MMVD distance) and/or how to interpret the signaled MMVD side information (e.g., distance index to the distance), may depend on information signaled or inferred at a level higher than CU level. (e.g. sequence level, or picture level or slice level, or tile group level, such as in VPS/SPS/PPS/slice header/picture header/tile group header).
- a. In one example, a code table index is signaled or inferred at a higher level. A specific code table is determined by the table index. The distance index may be signaled with the methods disclosed in item 1-26. Then the distance is derived by querying the entry with the signaled distance index in the specific code table.
 - b. In one example, a parameter X is signaled or inferred at a higher level. The distance index may be signaled with the methods disclosed in item 1-26. Then a distance D' is derived by querying the entry with the signaled distance index in a code table. Then the final distance D is calculated as $D=f(D', X)$. f can be any function. For example, $f(D', X) = D' \ll X$ or $f(D', X) = D' * X$, or $f(D', X) = D' + X$, or $f(D', X) = D'$ right shifted by X (with rounding or not).
 - c. In one example, valid MV resolution is signaled or inferred at a higher level. Only the MMVD distances with valid MV resolution can be signaled.
 - i. For example, the signaling method of MMVD information at CU level may depend on the valid MV resolutions signaled at a higher level.
 - (i) For example, the signaling method of MMVD distance resolution information at CU level may depend on the valid MV resolutions signaled at a higher level.
 - (ii) For example, the number of distance sub-sets may depend on the valid MV resolutions signaled at a higher level.
 - (iii) For example, the meaning of each sub-set may depend on the valid MV resolutions signaled at a higher level.

- ii. For example, a minimum MV resolution (such as $\frac{1}{4}$ -pel, or 1-pel, or 4-pel) is signaled.
 - (i) For example, when the minimum MV resolution is $\frac{1}{4}$ -pel, the distance index is signaled as disclosed in item 1-26.
 - (ii) For example, when the minimum MV resolution is 1-pel, the flag to signal whether the distance resolution is $\frac{1}{4}$ -pel (such as the first resolution bin in LAMVR) is not signaled. Only medium distance index and long distance index disclosed in Item 10 can be signaled following the LAMVR information.
 - (iii) For example, when the minimum MV resolution is 4-pel, the flag to signal whether the distance resolution is $\frac{1}{4}$ -pel (such as the first resolution bin in LAMVR) is not signaled; and the flag to signal whether the distance resolution is 1-pel (such as the second resolution bin in LAMVR) is not signaled. Only long distance index disclosed in Item 10 can be signaled following the LAMVR information.
 - (iv) For example, when the minimum MV resolution is 1-pel, the distance resolution is signaled in the same manner as when the minimum MV resolution is $\frac{1}{4}$ -pel. But the meanings of distance sub-sets may be different.
 - a. For example, the short distance sub-set represented by the short distance index is redefined as the very-long distance sub-set. For example, two distances that can be signaled within this very-long sub-set are 64-pel and 128-pel.

29A. It is proposed the encoder may decide whether the slice/tile/picture/sequence/group of CTUs/group of blocks is screen content or not by checking the ratio of the block that has one or more similar or identical blocks within the same slice/tile/picture/sequence/group of CTUs/group of blocks.

- a. In one example, if the ratio is larger than a threshold, it is considered as screen content.
- b. In one example, if the ratio is larger than a first threshold and smaller than a second threshold, it is considered as screen content.

- c. In one example, the slice/tile/picture/sequence/group of CTUs/group of blocks may be split into MxN non-overlapped blocks. For each MxN block, the encoder checks whether there is another (or more) MxN block is similar with or identical to it. For example, MxN is equal to 4x4.
- d. In one example, only partial of the blocks are checked when calculating the ratio. For example, only blocks in even row and even columns are checked.
- e. In one example, a key value, e.g., cyclic redundancy check (CRC) code, may be generated for each MxN block, key values of two blocks are compared to check whether the two blocks are identical or not.
 - i. In one example, the key value may be generated using only some color components of the block. For example, the key value is generated by using luma component only.
 - ii. In one example, the key value may be generated using only some pixels of the block. For example, only even rows of the block are used.
- f. In one example, SAD/SATD/SSE or mean-removed SAD/SATD/SSE may be used to measure the similarity of two blocks.
 - i. In one example, SAD/SATD/SSE or mean-removed SAD/SATD/SSE may be only calculated for some pixels. For example, SAD/SATD/SSE or mean-removed SAD/SATD/SSE is only calculated for the even rows.

Affine MMVD

- 29B. It is proposed that the indications of usage of Affine MMVD may be signaled only when the merge index of the sub-block merge list is larger than K (where K = 0 or 1).
- a. Alternatively, when there are separate lists for affine merge list and other merge lists (such as ATMVP list), the indication of usage of affine MMVD may be signaled only when the affine mode is enabled. Alternatively, furthermore, the indication of usage of affine MMVD may be signaled only when the affine mode is enabled and there are more than 1 base affine candidate.
30. It is proposed that the MMVD method can also be applied to other sub-block based coding tools in addition to affine mode, such as ATMVP mode. In one example, if the current CU applies ATMVP, and MMVD on/off flag is set to be 1, MMVD is applied to ATMVP.

- a. In one example, one set of MMVD side information may be applied to all sub-blocks, in this case, one set of MMVD side information signaled. Alternatively, different sub-blocks may choose different sets, in this case, multiple sets of MMVD side information may be signaled.
 - b. In one embodiment, the MV of each sub-block is added with the signaled MVD (a.k.a. offset or distance)
 - c. In one embodiment, the method to signal the MMVD information when the sub-block merge candidate is an ATMVP merge candidate is the same to the method when the sub-block merge candidate is an affine merge candidate.
 - d. In one embodiment, a POC distance based offset mirroring method is used for Bi-prediction to add the MVD on the MV of each sub-block when the sub-block merge candidate is an ATMVP merge candidate.
31. It is proposed that the MV of each sub-block is added with the signaled MVD (a.k.a. offset or distance) when the sub-block merge candidate is an affine merge candidate.
32. It is proposed that the MMVD signaling methods disclosed in bullet 1- 28 can also be applied to signal the MVD used by the affine MMVD mode.
- a. In one embodiment, the LAMVR information used to signal the MMVD information of affine MMVD mode may be different from the LAMVR information used to signal the MMVD information of non-affine MMVD mode.
 - i. For example, the LAMVR information used to signal the MMVD information of affine MMVD mode are also used to signal the MV precision used in affine inter-mode; but the LAMVR information used to signal the MMVD information of non-affine MMVD mode is used to signal the MV precision used in non-affine inter-mode.
33. It is proposed that the MVD information in MMVD mode for sub-block merge candidates should be signaled in the same way as the MVD information in MMVD mode for regular merge candidates.
- a. For example, they share the same distance table;
 - b. For example, they share the same mapping between the a distance index and a distance.
 - c. For example, they share the same directions.
 - d. For example, they share the same binarization methods.

e. For example, they share the same arithmetic coding contexts.

34. It is proposed that the MMVD side information signalling may be dependent on the coded mode, such as affine or normal merge or triangular merge mode or ATMVP mode.

35. It is proposed that the pre-defined MMVD side information may be dependent on the coded mode, such as affine or normal merge or triangular merge mode or ATMVP mode.

36. It is proposed that the pre-defined MMVD side information may be dependent on the color subsampling method (e.g., 4:2:0, 4:2:2, 4:4:4), and/or color component.

Triangle MMVD

37. It is proposed that MMVD can be applied on triangular prediction mode.

a. After a TPM merge candidate is signaled, the MMVD information is signaled. The signaled TPM merge candidate is treated as the base merge candidate.

i. For example, the MMVD information is signaled with the same signaling method as the MMVD for regular merge;

ii. For example, the MMVD information is signaled with the same signaling method as the MMVD for affine merge or other kinds of sub-block merge;

iii. For example, the MMVD information is signaled with a signaling method different to that of the MMVD for regular merge, affine merge or other kinds of sub-block merge;

b. In one example, the MV of each triangle partition is added with the signaled MVD;

c. In one example, the MV of one triangle partition is added with the signaled MVD, and the MV of the other triangle partition is added with $f(\text{signaled MVD})$, f is any function.

i. In one example, f depends on the reference picture POCs or reference indices of the two triangle partitions.

ii. In one example, $f(\text{MVD}) = -\text{MVD}$ if the reference picture of one triangle partition is before the current picture in displaying order and the reference picture of the other triangle partition is after the current picture in displaying order.

38. It is proposed that the MMVD signaling methods disclosed in bullet 1- 28 can also be applied to signal the MVD used by the triangular MMVD mode.

a. In one embodiment, the LAMVR information used to signal the MMVD information of affine MMVD mode may be different from the LAMVR information used to signal the MMVD information of non-affine MMVD mode.

i. For example, the LAMVR information used to signal the MMVD information of affine MMVD mode are also used to signal the MV precision used in affine inter-mode; but the LAMVR information used to signal the MMVD information of non-affine MMVD mode is used to signal the MV precision used in non-affine inter-mode.

39. For all above bullets, the MMVD side information may include e.g., the offset table (distance), and direction information.

[00180] 5. Example embodiments

[00181] This section shows some embodiments for the improved MMVD design.

[00182] 5.1 Embodiment #1 (MMVD distance index coding)

In one embodiment, to code MMVD distance, a first resolution bin is coded. For example, it may be coded with the same probability context as the first flag of MV resolution.

- If the resolution bin is 0, a following flag is coded. For example, it may be coded with another probability context to indicate the short distance index. If the flag is 0, the index is 0; if the flag is 1, the index is 1.
- Otherwise (the resolution bin is 0), the long distance index L is coded as a truncated unary code, with the maximum value MaxDI -2, where MaxDI is the maximum possible distance index, equal to 7 in the embodiment. After paring out L, the distance index is reconstructed as L+2. In an exemplary C-style embodiment:

```

DI = 2;
while( DI<7 ){
    oneBin = parseOneBin();
    if(oneBin == 0){
        break;
    }
    DI++;
}

```

The fist bin of the long distance index is coded with a probability context, and other bins are by-pass coded. In a C-style embodiment:

```

DI = 2;
while( DI<7 ){

```

```

if( DI==2){
    oneBin = parseOneBinWithContext();
}
else{
    oneBin = parseOneBinBypass();
}
if(oneBin == 0){
    break;
}
DI++;
}
    
```

An example of proposed syntax changes are **highlighted**, and the deleted parts are marked with **strikethrough**.

...	ae(v)
mmvd_flag[x0][y0]	
if(mmvd_flag[x0][y0] == 1) {	
mmvd_merge_flag[x0][y0]	ae(v)
mmvd_distance_idx[x0][y0]	ae(v)
mmvd_distance_subset_idx[x0][y0]	ae(v)
mmvd_distance_idx_in_subset[x0][y0]	ae(v)
mmvd_direction_idx[x0][y0]	ae(v)
...	

In one example, the **mmvd_distance_subset_idx** represents the resolution index as mentioned above, **mmvd_distance_idx_in_subset** represents the short or long distance index according to the resolution index. Truncated unary may be used to code **mmvd_distance_idx_in_subset**. In a Random Access test under the common test condition, the embodiment can achieve 0.15% coding gain in average and 0.34% gain on UHD sequences (class A1).

	Y	U	V
Class A1	-0.34%	-0.86%	-0.50%
Class A2	-0.15%	0.30%	0.07%
Class B	-0.09%	0.13%	-0.44%
Class C	-0.08%	0.05%	0.09%
Class E			
Overall	-0.15%	-0.06%	-0.21%
Class D	-0.05%	0.23%	-0.10%
Class F	-0.18%	-0.08%	-0.26%

[00183] 5.2 Embodiment #2 (MMVD side information coding)

MMVD is treated as a separate mode which is not treated as a merge mode. Therefore, MMVD flag may be further coded only when merge flag is 0.

	Descriptor
coding_unit(x0, y0, cbWidth, cbHeight, treeType) {	
if(slice_type != 1) {	
cu_skip_flag[x0][y0]	ae(v)
if(cu_skip_flag[x0][y0] == 0)	
pred_mode_flag	ae(v)
}	
if(CuPredMode[x0][y0] == MODE_INTRA) {	
... }	
} else { /* MODE_INTER */	
if(cu_skip_flag[x0][y0] == 0) {	
merge_flag[x0][y0]	ae(v)
if(merge_flag[x0][y0]) {	
merge_data(x0, y0, cbWidth, cbHeight)	
} else {	
mmvd_flag[x0][y0]	ae(v)
if(mmvd_flag[x0][y0] == 1) {	
mmvd_merge_flag[x0][y0]	ae(v)
mmvd_distance_subset_idx[x0][y0]	ae(v)
mmvd_distance_idx_in_subset[x0][y0]	ae(v)
mmvd_direction_idx[x0][y0]	ae(v)
} else {	
if(slice_type == B)	
inter_pred_idc[x0][y0]	ae(v)
if(sps_affine_enabled_flag && cbWidth >= 16 && cbHeight >= 16) {	
inter_affine_flag[x0][y0]	ae(v)
if(sps_affine_type_flag && inter_affine_flag[x0][y0])	
cu_affine_type_flag[x0][y0]	ae(v)
}	
... }	
}	
}	

	Descriptor
merge_data(x0, y0, cbWidth, cbHeight) {	
 mmvd_flag[x0][y0]	ae(v)
 if(mmvd_flag[x0][y0] == 1) {	
 mmvd_merge_flag[x0][y0]	ae(v)
 mmvd_distance_idx[x0][y0]	ae(v)
 mmvd_direction_idx[x0][y0]	ae(v)
 } else {	
if(MaxNumSubblockMergeCand > 0 && cbWidth >= 8 && cbHeight >= 8)	
merge_subblock_flag[x0][y0]	ae(v)
if(merge_subblock_flag[x0][y0] == 1) {	
if(MaxNumSubblockMergeCand > 1)	
merge_subblock_idx[x0][y0]	ae(v)
} else {	
if(sps_mh_intra_enabled_flag && cu_skip_flag[x0][y0] == 0 && (cbWidth * cbHeight) >= 64 && cbWidth < 128 && cbHeight < 128) {	

mh_intra_flag [x0][y0]	ae(v)
if(mh_intra_flag[x0][y0]) {	
if (cbWidth<= 2 * cbHeight cbHeight<= 2 * cbWidth)	
mh_intra_luma_mpm_flag [x0][y0]	ae(v)
if(mh_intra_luma_mpm_flag[x0][y0])	
mh_intra_luma_mpm_idx [x0][y0]	ae(v)
}	
}	
if(sps_triangle_enabled_flag&&slice_type == B &&cbWidth * cbHeight>= 16)	
merge_triangle_flag [x0][y0]	ae(v)
if(merge_triangle_flag[x0][y0])	
merge_triangle_idx [x0][y0]	ae(v)
else if(MaxNumMergeCand> 1)	
merge_idx [x0][y0]	ae(v)
}	
}	

In one embodiment, the MMVD information is signaled as:

if(mmvd_flag[x0][y0] == 1) {	
mmvd_merge_flag [x0][y0]	ae(v)
amvr_mode [x0][y0]	ae(v)
mmvd_distance_idx_in_subset [x0][y0]	ae(v)
mmvd_direction_idx [x0][y0]	ae(v)
}	

mmvd_distance_idx_in_subset[x0][y0] is binarized as a truncated unary code. The maximum value of the truncated unary code is 1 if **amvr_mode**[x0][y0] <2; Otherwise (**amvr_mode**[x0][y0] is equal to 2), The maximum value is set to be 3.

mmvd_distance_idx[x0][y0] is set equal to **mmvd_distance_idx_in_subset**[x0][y0]+2***amvr_mode**[x0][y0].

Which probability contexts are used by **mmvd_distance_idx_in_subset**[x0][y0] depends on **amvr_mode**[x0][y0].

5.3 Embodiment #3 (MMVD slice level control)

In slice header, a syntax element **mmvd_integer_flag** is signaled.

The syntax change is described as follows, and the newly added parts are highlighted in *italics*

7.3.2.1 Sequence parameter set RBSP syntax

seq_parameter_set_rbsp() {	Descriptor
...	u(1)
sps_gbi_enabled_flag	u(1)
sps_mh_intra_enabled_flag	u(1)
sps_fracmmvd_enabled_flag	u(1)
sps_triangle_enabled_flag	u(1)

sps_ladf_enabled_flag	u(1)
if (sps_ladf_enabled_flag) {	
sps_num_ladf_intervals_minus2	u(2)
sps_ladf_lowest_interval_qp_offset	se(v)
for(i = 0; i < sps_num_ladf_intervals_minus2 + 1; i++) {	
sps_ladf_qp_offset[i]	se(v)
sps_ladf_delta_threshold_minus1[i]	ue(v)
}	
}	
rbbsp_trailing_bits()	
}	

7.3.3.1 General slice header syntax

slice_header() {	Descriptor
...	ue(v)
if (slice_type != I) {	
if(sps_temporal_mvp_enabled_flag)	
slice_temporal_mvp_enabled_flag	u(1)
if(slice_type == B)	
mvd_l1_zero_flag	u(1)
if(slice_temporal_mvp_enabled_flag) {	
if(slice_type == B)	
collocated_from_l0_flag	u(1)
}	
six_minus_max_num_merge_cand	ue(v)
if(sps_affine_enable_flag)	
five_minus_max_num_subblock_merge_cand	ue(v)
if(sps_fracmmvd_enable_flag)	
slice_fracmmvd_flag	u(1)
}	
slice_qp_delta	se(v)
...	
}	

7.4.3.1 Sequence parameter set RBSP semantics

sps_fracmmvd_enabled_flag equal to 1 specifies that *slice_fracmmvd_flag* is present in the slice header syntax for B slices and P slices. *sps_fracmmvd_enabled_flag* equal to 0 specifies that *slice_fracmmvd_flag* is not present in the slice header syntax for B slices and P slices.

7.4.4.1 General slice header semantics

slice_fracmmvd_flag specifies the distance table used to derive *MmvdDistance[x0][y0]*. When not present, the value of *slice_fracmmvd_flag* is inferred to be 1.

coding_unit(x0, y0, cbWidth, cbHeight, treeType) {	Descriptor
if(slice_type != I) {	
cu_skip_flag [x0][y0]	ae(v)
if(cu_skip_flag[x0][y0] == 0)	
pred_mode_flag	ae(v)
}	
if(CuPredMode[x0][y0] == MODE_INTRA) {	
... }	
} else { /* MODE_INTER */	
if(cu_skip_flag[x0][y0] == 0) {	
merge_flag [x0][y0]	ae(v)
if(merge_flag[x0][y0]) {	
merge_data(x0, y0, cbWidth, cbHeight)	
} else {	
mmvd_flag [x0][y0]	ae(v)
if(mmvd_flag[x0][y0] == 1) {	
mmvd_merge_flag [x0][y0]	ae(v)
mmvd_distance_subset_idx [x0][y0]	ae(v)
mmvd_distance_idx_in_subset [x0][y0]	ae(v)
mmvd_direction_idx [x0][y0]	ae(v)
} else {	
if(slice_type == B)	
inter_pred_idc [x0][y0]	ae(v)
if(sps_affine_enabled_flag && cbWidth >= 16 && cbHeight >= 16) {	
inter_affine_flag [x0][y0]	ae(v)
if(sps_affine_type_flag && inter_affine_flag[x0][y0])	
cu_affine_type_flag [x0][y0]	ae(v)
}	
}	
... }	
}	
}	

merge_data(x0, y0, cbWidth, cbHeight) {	Descriptor
 mmvd_flag[x0][y0]	ae(v)
 if(mmvd_flag[x0][y0] == 1) {	
 mmvd_merge_flag[x0][y0]	ae(v)
 mmvd_distance_idx[x0][y0]	ae(v)
 mmvd_direction_idx[x0][y0]	ae(v)
 } else {	
if(MaxNumSubblockMergeCand > 0 && cbWidth >= 8 && cbHeight >= 8)	

merge_subblock_flag [x0][y0]	ae(v)
if(merge_subblock_flag[x0][y0] == 1) {	
if(MaxNumSubblockMergeCand > 1)	
merge_subblock_idx [x0][y0]	ae(v)
} else {	
if(sps_mh_intra_enabled_flag && cu_skip_flag[x0][y0] == 0 && (cbWidth * cbHeight) >= 64 && cbWidth < 128 && cbHeight < 128) {	
mh_intra_flag [x0][y0]	ae(v)
if(mh_intra_flag[x0][y0]) {	
if (cbWidth <= 2 * cbHeight cbHeight <= 2 * cbWidth)	
mh_intra_luma_mpm_flag [x0][y0]	ae(v)
if(mh_intra_luma_mpm_flag[x0][y0])	
mh_intra_luma_mpm_idx [x0][y0]	ae(v)
}	
}	
if(sps_triangle_enabled_flag && slice_type == B && cbWidth * cbHeight >= 16)	
merge_triangle_flag [x0][y0]	ae(v)
if(merge_triangle_flag[x0][y0])	
merge_triangle_idx [x0][y0]	ae(v)
else if(MaxNumMergeCand > 1)	
merge_idx [x0][y0]	ae(v)
}	
}	

In one embodiment, the MMVD information is signaled as:

if(mmvd_flag[x0][y0] == 1) {	
mmvd_merge_flag [x0][y0]	ae(v)
amvr_mode [x0][y0]	ae(v)
mmvd_distance_idx_in_subset [x0][y0]	ae(v)
mmvd_direction_idx [x0][y0]	ae(v)
}	

mmvd_distance_idx_in_subset[x0][y0] is binarized as a truncated unary code. The maximum value of the truncated unary code is 1 if **amvr_mode**[x0][y0] < 2; Otherwise (**amvr_mode**[x0][y0] is equal to 2), The maximum value is set to be 3.

mmvd_distance_idx[x0][y0] is set equal to **mmvd_distance_idx_in_subset**[x0][y0] + 2 * **amvr_mode**[x0][y0]. In one example, the probability contexts are used by **mmvd_distance_idx_in_subset**[x0][y0] depends on **amvr_mode**[x0][y0].

The array indices x0, y0 specify the location (x0, y0) of the top-left luma sample of the considered coding block relative to the top-left luma sample of the picture. The **mmvd_distance_idx**[x0][y0] and the **MmvdDistance**[x0][y0] is as follows:

*Table 7-9 – Specification of $MmvdDistance[x0][y0]$ based on $mmvd_distance_idx[x0][y0]$ when slice **fracmmvd_flag** is equal to 1.*

$mmvd_distance_idx[x0][y0]$	$MmvdDistance[x0][y0]$
0	1
1	2
2	4
3	8
4	16
5	32
6	64
7	128

*Table 7-9 – Specification of $MmvdDistance[x0][y0]$ based on $mmvd_distance_idx[x0][y0]$ when slice **fracmmvd_flag** is equal to 0.*

$mmvd_distance_idx[x0][y0]$	$MmvdDistance[x0][y0]$
0	4
1	8
2	16
3	32
4	64
5	128
6	256
7	512

When $mmvd_integer_flag$ is equal to 1, $mmvd_distance = mmvd_distance \ll 2$.

[00184] FIG. 10 is a flowchart for an example method 1000 for video processing. The method 1000 includes determining(1002), determining, for a current video block which is coded using a merge with motion vector difference (MMVD) mode, a first relationship between a distance and a distance index (DI), wherein the distance is a distance between a motion vector of the current video block and a base candidate selected from a merge candidate list; and performing(1004), based on the first relationship, a conversion between the current video block and a bitstream representation of the current video block.

[00185] FIG. 11 is a flowchart for an example method 1100 for video processing. The method 1100 includes performing(1102) a conversion between a current video block and a bitstream representation of the current video block, wherein the current video block is coded using a merge with motion vector difference (MMVD) mode; wherein the conversion comprises parsing MMVD side information from or writing the MMVD side information into the bitstream representation, wherein the MMVD side information comprises at least one of a MMVD flag indicating whether MMVD syntaxes are parsed, a first syntax element indicating a distance of

MMVD between a motion vector of the current video block and a base candidate selected from a merge candidate list, a second syntax element indicating a direction of MMVD representing a direction of motion vector difference (MVD) relative to the base candidate.

[00186] FIG. 12 is a flowchart for an example method 1200 for video processing. The method 1200 includes determining(1202) at least one distance for motion vector difference(MVD) associated with a current video block, which is coded in a merge with motion vector difference (MMVD) mode, from a first distance with a rough granularity and one or more distances with fine granularities; and performing(1204) a conversion between a current video block and a bitstream representation of the current video block based on the distance for MVD.

[00187] With reference to the methods 1000, 1100, 1200, some examples of motion vector signaling are described in Section 4 of this document, and the aforementioned methods may include the features and steps described below.

[00188] In one aspect, there is disclosed a method for video processing, comprising: determining, for a current video block which is coded using a merge with motion vector difference (MMVD) mode, a first relationship between a distance and a distance index (DI), wherein the distance is a distance between a motion vector of the current video block and a base candidate selected from a merge candidate list; and performing, based on the first relationship, a conversion between the current video block and a bitstream representation of the current video block.

[00189] In one example, the first relationship is different from a single exponential relationship.

[00190] In one example, the single exponential relationship is specified as:

$$\text{distance} = 1/4\text{-pel} \times 2^{\text{DI}}.$$

[00191] In one example, the first relationship is based on a piece-wise mapping.

[00192] In one example, the piece-wise mapping is specified as:

$$\text{distance} = 1/4\text{-pel} \times 2^{\text{DI}} \text{ when } \text{DI} < \text{T1},$$

$$\text{distance} = a \times \text{DI} + b \text{ when } \text{T1} \leq \text{DI} < \text{T2}, \text{ and}$$

$$\text{distance} = c \times 2^{\text{DI}} \text{ when } \text{DI} \geq \text{T2},$$

wherein a, b, c, T1 and T2 are numbers.

[00193] In one example, a = 1, b = -1, c = 1/8, T1 = 4 and T2 = 6.

[00194] In one example, the first relationship is represented as at least one distance table including at least one distance, as an entry, indicated with a distance index.

[00195] In one example, the distance table includes more than 8 entries.

[00196] In one example, the distance table comprises 9, 10, 12 or 16 entries.

[00197] In one example, the distance table comprises one or more entries shorter than 1/4-pel.

[00198] In one example, the one or more entries have one of 1/8-pel, 1/16-pel and 3/8-pel precision.

[00199] In one example, the distance table comprises one or more entries not in a form of 2^X -pel, wherein X is an integer.

[00200] In one example, the one or more entries have one of 3-pel, 5-pel and 6-pel precision.

[00201] In one example, different distance tables are used for different directions with different direction indexes, wherein the direction represents a direction of motion vector difference (MVD) relative to the base candidate.

[00202] In one example, the different directions are two horizontal directions with different direction indexes or two vertical directions with different direction indexes.

[00203] In one example, a same distance table is shared by two horizontal directions with different direction indexes or shared by two vertical directions with different direction indexes.

[00204] In one example, the distance table for a vertical direction has a smaller size than the distance table for a horizontal direction.

[00205] In one example, a minimum entry in the distance table for a vertical direction is smaller than a minimum entry in the distance table for a horizontal direction.

[00206] In one example, a maximum entry in the distance table for a vertical direction is smaller than a maximum entry in the distance table for a horizontal direction.

[00207] In one example, a first distance table is used for the current video block, and a second distance table different from the first distance table is used for a subsequent video block with dimensions different from the current video block.

[00208] In one example, the subsequent video block has a width different from that of the current video block in a horizontal direction.

[00209] In one example, the subsequent video block has a height different from that of the current video block in a vertical direction.

[00210] In one example, a first distance table is used for the current video block, and a second distance table different from the first distance table is used for a subsequent video block with different picture order count (POC) distance.

[00211] In one example, different distance tables are used different base candidates for the current video block.

[00212] In one example, a ratio of two entries with consecutive distance indexes in the distance table is fixed to M, and M is not equal to 2.

[00213] In one example, M=4.

[00214] In one example, a delta of two entries with consecutive distance indexes in the distance table is fixed for all distance indexes.

[00215] In one example, a ratio of two entries with consecutive distance indexes in the distance table is different for different distance indexes.

[00216] In one example, the distance table has one set of entries:

- {1-pel, 2-pel, 4-pel, 8-pel, 16-pel, 32-pel, 48-pel, 64-pel};
- {1-pel, 2-pel, 4-pel, 8-pel, 16-pel, 32-pel, 64-pel, 96-pel}; and
- {1-pel, 2-pel, 3-pel, 4-pel, 5-pel, 16-pel, 32-pel}.

[00217] In one example, the distance table has a size of 9 entries:

Distance IDX	0	1	2	3	4	5	6	7	8
Pixel distance	1/8-pel	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	3-pel	4-pel	8-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	6-pel	8-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	12-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	16-pel	24-pel	32-pel

[00218] In one example, the distance table has a size of 10 entries:

Distance IDX	0	1	2	3	4	5	6	7	8	9
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	8-pel	16- pel	32- pel	64- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	3-pel	4-pel	8-pel	16- pel	32- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	6-pel	8-pel	16- pel	32- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	8-pel	12- pel	16- pel	32- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	8-pel	16- pel	24- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	3-pel	4-pel	6-pel	8-pel	16- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	4-pel	6-pel	8-pel	12- pel	16- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	3-pel	4-pel	8-pel	12- pel	16- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2-pel	3-pel	4-pel	8-pel	16- pel	24- pel	32- pel

[00219] In one example, the distance table has a size of 11 entries:

Distance IDX	0	1	2	3	4	5	6	7	8	9	10	11
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1- pel	2- pel	3- pel	4- pel	6- pel	8- pel	16- pel	32- pel	64- pel
Pixel distance	1/8- pel	1/4- pel	1/2- pel	1- pel	2- pel	3- pel	4- pel	8- pel	12- pel	16- pel	24- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2- pel	3- pel	4- pel	6- pel	8- pel	12- pel	16- pel	24- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2- pel	3- pel	4- pel	6- pel	8- pel	12- pel	16- pel	32- pel	64- pel

[00220] In one aspect, there is disclosed a method for video processing, comprising:

performing a conversion between a current video block and a bitstream representation of the current video block, wherein the current video block is coded using a merge with motion vector difference (MMVD) mode;

wherein the conversion comprises parsing MMVD side information from or writing the MMVD side information into the bitstream representation, wherein the MMVD side information comprises at least one of an MMVD flag indicating whether MMVD syntaxes are parsed, a first syntax element indicating a distance of MMVD between a motion vector of the current video block and a base candidate selected from a merge candidate list, a second syntax element indicating a direction of MMVD representing a direction of motion vector difference (MVD) relative to the base candidate.

[00221] In one example, distances of MMVD allowed for the current video block are classified into a plurality of subsets, and the first syntax element comprises a subset index and a distance index of MMVD indicating a distance within a subset with the subset index, and the second syntax element comprise a direction index indicating the direction.

[00222] In one example, the MMVD flag may be firstly signaled, followed by the sub-set index of distance, the distance index within the sub-set, a direction index.

[00223] In one example, one of the first and second syntax elements is parsed from or written into the bitstream representation based on the other of the first and second syntax elements.

[00224] In one example, a combination of the first and second syntax elements is represented with at least one codeword.

[00225] In one example, the combination of the first and second syntax elements is a sum of the distance index of MMVD and the direction index of MMVD*7.

[00226] In one example, the at least one codeword is included in a codeword table which comprises a plurality of codewords each corresponds to a unique combination of the first and second syntax elements.

[00227] In one aspect, there is disclosed a method for video processing, comprising:

determining at least one distance for motion vector difference(MVD) associated with a current video block, which is coded in a merge with motion vector difference (MMVD) mode, from a first distance with a rough granularity and one or more distances with fine granularities; and

performing a conversion between a current video block and a bitstream representation of the current video block based on the distance for MVD.

[00228] In one example, the at least one distance is determined as a sum of the first distance with a rough granularity and the one or more distances with fine granularities, $M_1[F_1] + M_2[F_2] + M_i[F_i] \dots + M_n[F_n]$, wherein M_1 represents a first set comprising at least one distance with a rough granularity as an entry, F_1 indicates an entry in M_1 , M_i represents a set comprising at least one distance with a fine granularity as an entry, and F_i indicates an entry in M_i , i is an integer from 2 to n .

[00229] In one example, $M_k[F_k] < M_{k-1}[F_{k-1} + 1] - M_{k-1}[F_{k-1}]$, k is an integer from 2 to n .

[00230] In one example, a binarization of F_k depends on that of F_{k-1} .

[00231] In one example, $M_k[F_k] < M_s[F_s + 1] - M_s[F_s]$, k is an integer from 2 to n , $1 \leq s < k$.

[00232] In one example, a binarization of F_k depends on that of F_s , k is an integer from 2 to n , $1 \leq s < k$.

[00233] In one example, entries in $M_k[F_k]$ depends on F_{k-1} , k is an integer from 2 to n .

[00234] In one example, entries in $M_k[F_k]$ depends on F_s , k is an integer from 2 to n , $1 \leq s < k$.

[00235] In one example, $n=2$, $M_1 = \{1/4\text{-pel}, 1\text{-pel}, 4\text{-pel}, 8\text{-pel}, 16\text{-pel}, 32\text{-pel}, 64\text{-pel}, 128\text{-pel}\}$,

$M_2 = \{0\text{-pel}, 1/4\text{-pel}\}$ when $F_1 = 0$;

$M_2 = \{0\text{-pel}, 1\text{-pel}, 2\text{-pel}\}$ when $F_1 = 1$;

$M_2 = \{0\text{-pel}, 1\text{-pel}, 2\text{-pel}, 3\text{-pel}\}$ when $F_1 = 2$;

$M_2 = \{0\text{-pel}, 2\text{-pel}, 4\text{-pel}, 6\text{-pel}\}$ when $F_1 = 3$;

$M_2 = \{0\text{-pel}, 4\text{-pel}, 8\text{-pel}, 12\text{-pel}\}$ when $F_1 = 4$;

$M_2 = \{0\text{-pel}, 8\text{-pel}, 16\text{-pel}, 24\text{-pel}\}$ when $F_1 = 5$; and

$M_2 = \{0\text{-pel}, 16\text{-pel}, 32\text{-pel}, 48\text{-pel}\}$ when $F_1 = 6$.

[00236] In an example, the conversion includes encoding the current video block into the bitstream representation of the current video block and decoding the current video block from the bitstream representation of the current video block.

[00237] In an aspect, there is disclosed an apparatus in a video system, the apparatus comprising a processor and a non-transitory memory with instructions thereon, wherein the instructions upon execution by the processor, cause the processor to implement the method in any one of examples described as above.

[00238] In an aspect, there is disclosed a computer program product stored on a non-transitory computer readable media, the computer program product including program code for carrying out the method in any one of examples described as above.

[00239] FIG. 13 is a block diagram of a video processing apparatus 1300. The apparatus 1300 may be used to implement one or more of the methods described herein. The apparatus 1300 may be embodied in a smartphone, tablet, computer, Internet of Things (IoT) receiver, and so on. The apparatus 1300 may include one or more processors 1302, one or more memories 1304 and video processing hardware 1306. The processor(s) 1302 may be configured to implement one or more methods described in the present document. The memory (memories) 1304 may be used for storing data and code used for implementing the methods and techniques described herein. The video processing hardware 1306 may be used to implement, in hardware circuitry, some techniques described in the present document, and may be partly or completely be a part of the processors 1302 (e.g., graphics processor core GPU or other signal processing circuitry).

[00240] In the present document, the term “video processing” may refer to video encoding, video decoding, video compression or video decompression. For example, video compression algorithms may be applied during conversion from pixel representation of a video to a corresponding bitstream representation or vice versa. The bitstream representation of a current video block may, for example, correspond to bits that are either co-located or spread in different places within the bitstream, as is defined by the syntax. For example, a macroblock may be encoded in terms of transformed and coded error residual values and also using bits in headers and other fields in the bitstream.

[00241] It will be appreciated that several techniques have been disclosed that will benefit video encoder and decoder embodiments incorporated within video processing devices such as smartphones, laptops, desktops, and similar devices by allowing the use of virtual motion candidates that are constructed based on various rules disclosed in the present document.

[00242] The disclosed and other solutions, examples, embodiments, modules and the functional operations described in this document can be implemented in digital electronic circuitry, or in computer software, firmware, or hardware, including the structures disclosed in this document and their structural equivalents, or in combinations of one or more of them. The disclosed and other embodiments can be implemented as one or more computer program products, i.e., one or more modules of computer program instructions encoded on a computer

readable medium for execution by, or to control the operation of, data processing apparatus. The computer readable medium can be a machine-readable storage device, a machine-readable storage substrate, a memory device, a composition of matter effecting a machine-readable propagated signal, or a combination of one or more them. The term “data processing apparatus” encompasses all apparatus, devices, and machines for processing data, including by way of example a programmable processor, a computer, or multiple processors or computers. The apparatus can include, in addition to hardware, code that creates an execution environment for the computer program in question, e.g., code that constitutes processor firmware, a protocol stack, a database management system, an operating system, or a combination of one or more of them. A propagated signal is an artificially generated signal, e.g., a machine-generated electrical, optical, or electromagnetic signal, that is generated to encode information for transmission to suitable receiver apparatus.

[00243] A computer program (also known as a program, software, software application, script, or code) can be written in any form of programming language, including compiled or interpreted languages, and it can be deployed in any form, including as a stand-alone program or as a module, component, subroutine, or other unit suitable for use in a computing environment. A computer program does not necessarily correspond to a file in a file system. A program can be stored in a portion of a file that holds other programs or data (e.g., one or more scripts stored in a markup language document), in a single file dedicated to the program in question, or in multiple coordinated files (e.g., files that store one or more modules, sub programs, or portions of code). A computer program can be deployed to be executed on one computer or on multiple computers that are located at one site or distributed across multiple sites and interconnected by a communication network.

[00244] The processes and logic flows described in this document can be performed by one or more programmable processors executing one or more computer programs to perform functions by operating on input data and generating output. The processes and logic flows can also be performed by, and apparatus can also be implemented as, special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application specific integrated circuit).

[00245] Processors suitable for the execution of a computer program include, by way of example, both general and special purpose microprocessors, and any one or more processors of any kind of digital computer. Generally, a processor will receive instructions and data from a read only memory or a random-access memory or both. The essential elements of a computer

are a processor for performing instructions and one or more memory devices for storing instructions and data. Generally, a computer will also include, or be operatively coupled to receive data from or transfer data to, or both, one or more mass storage devices for storing data, e.g., magnetic, magneto optical disks, or optical disks. However, a computer need not have such devices. Computer readable media suitable for storing computer program instructions and data include all forms of non-volatile memory, media and memory devices, including by way of example semiconductor memory devices, e.g., EPROM, EEPROM, and flash memory devices; magnetic disks, e.g., internal hard disks or removable disks; magneto optical disks; and CD ROM and DVD-ROM disks. The processor and the memory can be supplemented by, or incorporated in, special purpose logic circuitry.

[00246] While this patent document contains many specifics, these should not be construed as limitations on the scope of any subject matter or of what may be claimed, but rather as descriptions of features that may be specific to particular embodiments of particular techniques. Certain features that are described in this patent document in the context of separate embodiments can also be implemented in combination in a single embodiment. Conversely, various features that are described in the context of a single embodiment can also be implemented in multiple embodiments separately or in any suitable subcombination. Moreover, although features may be described above as acting in certain combinations and even initially claimed as such, one or more features from a claimed combination can in some cases be excised from the combination, and the claimed combination may be directed to a subcombination or variation of a subcombination.

[00247] Similarly, while operations are depicted in the drawings in a particular order, this should not be understood as requiring that such operations be performed in the particular order shown or in sequential order, or that all illustrated operations be performed, to achieve desirable results. Moreover, the separation of various system components in the embodiments described in this patent document should not be understood as requiring such separation in all embodiments.

[00248] Only a few implementations and examples are described and other implementations, enhancements and variations can be made based on what is described and illustrated in this patent document.

CLAIMS

1. A method for video processing, comprising:
 - determining, for a current video block which is coded using a merge with motion vector difference (MMVD) mode, a first relationship between a distance and a distance index (DI), wherein the distance is a distance between a motion vector of the current video block and a base candidate selected from a merge candidate list; and
 - performing, based on the first relationship, a conversion between the current video block and a bitstream representation of the current video block.
2. The method of claim 1, wherein the first relationship is different from a single exponential relationship.
3. The method of claim 2, wherein the single exponential relationship is specified as:

$$\text{distance} = 1/4\text{-pel} \times 2^{\text{DI}}.$$
4. The method of claim 1, wherein the first relationship is based on a piece-wise mapping.
5. The method of claim 4, wherein the piece-wise mapping is specified as:

$$\begin{aligned} \text{distance} &= 1/4\text{-pel} \times 2^{\text{DI}} \text{ when } \text{DI} < \text{T1}, \\ \text{distance} &= a \times \text{DI} + b \text{ when } \text{T1} \leq \text{DI} < \text{T2}, \text{ and} \\ \text{distance} &= c \times 2^{\text{DI}} \text{ when } \text{DI} \geq \text{T2}, \end{aligned}$$
 wherein a, b, c, T1 and T2 are numbers.
6. The method of claim 5, wherein a = 1, b = -1, c = 1/8, T1 = 4 and T2 = 6.
7. The method of any one of claims 1-6, wherein the first relationship is represented as at least one distance table including at least one distance, as an entry, indicated with a distance index.
8. The method of claim 7, wherein the distance table includes more than 8 entries.
9. The method of claim 7, wherein the distance table comprises 9, 10, 12 or 16 entries.
10. The method of claim 7, wherein the distance table comprises one or more entries shorter than 1/4-pel.

11. The method of claim 10, wherein the one or more entries have one of 1/8-pel, 1/16-pel and 3/8-pel precision.
12. The method of claim 7, wherein the distance table comprises one or more entries not in a form of 2^X -pel, wherein X is an integer.
13. The method of claim 12, wherein the one or more entries have one of 3-pel, 5-pel and 6-pel precision.
14. The method of claim 7, wherein different distance tables are used for different directions with different direction indexes, wherein the direction represents a direction of motion vector difference (MVD) relative to the base candidate.
15. The method of claim 14, wherein the different directions are two horizontal directions with different direction indexes or two vertical directions with different direction indexes.
16. The method of claim 7, wherein a same distance table is shared by two horizontal directions with different direction indexes or shared by two vertical directions with different direction indexes.
17. The method of claim 14, wherein the distance table for a vertical direction has a smaller size than the distance table for a horizontal direction.
18. The method of claim 14, wherein a minimum entry in the distance table for a vertical direction is smaller than a minimum entry in the distance table for a horizontal direction.
19. The method of claim 14, wherein a maximum entry in the distance table for a vertical direction is smaller than a maximum entry in the distance table for a horizontal direction.
20. The method of claim 7, wherein a first distance table is used for the current video block, and a second distance table different from the first distance table is used for a subsequent video block with dimensions different from the current video block.
21. The method of claim 20, wherein the subsequent video block has a width different from that of the current video block in a horizontal direction.
22. The method of claim 20, wherein the subsequent video block has a height different from that of the current video block in a vertical direction.

23. The method of claim 7, wherein a first distance table is used for the current video block, and a second distance table different from the first distance table is used for a subsequent video block with different picture order count (POC) distance.

24. The method of claim 7, wherein different distance tables are used for different base candidates for the current video block.

25. The method of claim 7, wherein a ratio of two entries with consecutive distance indexes in the distance table is fixed to M, and M is not equal to 2.

26. The method of claim 25, wherein M=4.

27. The method of claim 7, wherein a delta of two entries with consecutive distance indexes in the distance table is fixed for all distance indexes.

28. The method of claim 7, wherein a ratio of two entries with consecutive distance indexes in the distance table is different for different distance indexes.

29. The method of claim 28, wherein the distance table has one set of entries:

{1-pel, 2-pel, 4-pel, 8-pel, 16-pel, 32-pel, 48-pel, 64-pel};

{1-pel, 2-pel, 4-pel, 8-pel, 16-pel, 32-pel, 64-pel, 96-pel}; and

{1-pel, 2-pel, 3-pel, 4-pel, 5-pel, 16-pel, 32-pel}.

30. The method of claim 7, wherein the distance table has a size of 9 entries:

Distance IDX	0	1	2	3	4	5	6	7	8
Pixel distance	1/8-pel	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	3-pel	4-pel	8-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	6-pel	8-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	12-pel	16-pel	32-pel

Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	16-pel	24-pel	32-pel
----------------	---------	---------	-------	-------	-------	-------	--------	--------	--------

31. The method of claim 7, wherein the distance table has a size of 10 entries:

Distance IDX	0	1	2	3	4	5	6	7	8	9
Pixel distance	1/8-pel	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	16-pel	32-pel	64-pel
Pixel distance	1/8-pel	1/4-pel	1/2-pel	1-pel	2-pel	3-pel	4-pel	8-pel	16-pel	32-pel
Pixel distance	1/8-pel	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	6-pel	8-pel	16-pel	32-pel
Pixel distance	1/8-pel	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	12-pel	16-pel	32-pel
Pixel distance	1/8-pel	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	8-pel	16-pel	24-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	3-pel	4-pel	6-pel	8-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	4-pel	6-pel	8-pel	12-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	3-pel	4-pel	8-pel	12-pel	16-pel	32-pel
Pixel distance	1/4-pel	1/2-pel	1-pel	2-pel	3-pel	4-pel	8-pel	16-pel	24-pel	32-pel

32. The method of claim 7, wherein the distance table has a size of 11 entries:

Distance IDX	0	1	2	3	4	5	6	7	8	9	10	11
Pixel distance	1/8-pel	1/4-pel	1/2-pel	1-pel	2-pel	3-pel	4-pel	6-pel	8-pel	16-pel	32-pel	64-pel
Pixel	1/8-pel	1/4-pel	1/2-pel	1-pel	2-pel	3-pel	4-pel	8-pel	12-pel	16-pel	24-pel	32-pel

distance	pel	pel	pel	pel	pel	pel	pel	pel	pel	pel	pel	pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2- pel	3- pel	4- pel	6- pel	8- pel	12- pel	16- pel	24- pel	32- pel
Pixel distance	1/4- pel	1/2- pel	1-pel	2- pel	3- pel	4- pel	6- pel	8- pel	12- pel	16- pel	32- pel	64- pel

33. A method for video processing, comprising:

performing a conversion between a current video block and a bitstream representation of the current video block, wherein the current video block is coded using a merge with motion vector difference (MMVD) mode;

wherein the conversion comprises parsing MMVD side information from or writing the MMVD side information into the bitstream representation, wherein the MMVD side information comprises at least one of an MMVD flag indicating whether MMVD syntaxes are parsed, a first syntax element indicating a distance of MMVD between a motion vector of the current video block and a base candidate selected from a merge candidate list, a second syntax element indicating a direction of MMVD representing a direction of motion vector difference (MVD) relative to the base candidate.

34. The method of claim 33, wherein distances of MMVD allowed for the current video block are classified into a plurality of subsets, and the first syntax element comprises a subset index and a distance index of MMVD indicating a distance within a subset with the subset index, and the second syntax element comprise a direction index indicating the direction.

35. The method of claim 34, wherein the MMVD flag is firstly signaled, followed by the sub-set index of distance, the distance index within the sub-set, a direction index.

36. The method of claim 33 or 34, wherein one of the first and second syntax elements is parsed from or written into the bitstream representation based on the other of the first and second syntax elements.

37. The method of claim 33 or 34, wherein a combination of the first and second syntax elements is represented with at least one codeword.

38. The method of claim 37, wherein the combination of the first and second syntax elements is a sum of the distance index of MMVD and the direction index of MMVD*7.

39. The method of 37, wherein the at least one codeword is included in a codeword table which comprises a plurality of codewords each corresponds to a unique combination of the first and second syntax elements.

40. A method for video processing, comprising:

determining at least one distance for motion vector difference (MVD) associated with a current video block, which is coded in a merge with motion vector difference (MMVD) mode, from a first distance with a rough granularity and one or more distances with fine granularities; and

performing a conversion between a current video block and a bitstream representation of the current video block based on the distance for MVD.

41. The method of claim 40, wherein the at least one distance is determined as a sum of the first distance with a rough granularity and the one or more distances with fine granularities, $M_1[F_1]+M_2[F_2]+M_i[F_i] \dots + M_n[F_n]$, wherein M_1 represents a first set comprising at least one distance with a rough granularity as an entry, F_1 indicates an entry in M_1 , M_i represents a set comprising at least one distance with a fine granularity as an entry, and F_i indicates an entry in M_i , i is an integer from 2 to n .

42. The method of claim 41, wherein $M_k[F_k] < M_{k-1}[F_{k-1}+1]-M_{k-1}[F_{k-1}]$, k is an integer from 2 to n .

43. The method of claim 42, wherein a binarization of F_k depends on that of F_{k-1} .

44. The method of claim 41, wherein $M_k[F_k] < M_s[F_s+1]-M_s[F_s]$, k is an integer from 2 to n , $1 \leq s < k$.

45. The method of claim 44, wherein a binarization of F_k depends on that of F_s , k is an integer from 2 to n , $1 \leq s < k$.

46. The method of claim 41, wherein entries in $M_k[F_k]$ depends on F_{k-1} , k is an integer from 2 to n .

47. The method of claim 41, wherein entries in $M_k[F_k]$ depends on F_s , k is an integer from 2 to n , $1 \leq s < k$,

48. The method of claim 46, wherein $n=2$, $M_1 = \{1/4\text{-pel}, 1\text{-pel}, 4\text{-pel}, 8\text{-pel}, 16\text{-pel}, 32\text{-pel}, 64\text{-pel}, 128\text{-pel}\}$,

$M_2 = \{0\text{-pel}, 1/4\text{-pel}\}$ when $F_1 = 0$;

$M_2 = \{0\text{-pel}, 1\text{-pel}, 2\text{-pel}\}$ when $F_1 = 1$;

$M_2 = \{0\text{-pel}, 1\text{-pel}, 2\text{-pel}, 3\text{-pel}\}$ when $F_1 = 2$;

$M_2 = \{0\text{-pel}, 2\text{-pel}, 4\text{-pel}, 6\text{-pel}\}$ when $F_1 = 3$;

$M_2 = \{0\text{-pel}, 4\text{-pel}, 8\text{-pel}, 12\text{-pel}\}$ when $F_1 = 4$;

$M_2 = \{0\text{-pel}, 8\text{-pel}, 16\text{-pel}, 24\text{-pel}\}$ when $F_1 = 5$; and

$M_2 = \{0\text{-pel}, 16\text{-pel}, 32\text{-pel}, 48\text{-pel}\}$ when $F_1 = 6$.

49. The method of any of claims 1 to 48, wherein the conversion includes encoding the current video block into the bitstream representation of the current video block and decoding the current video block from the bitstream representation of the current video block.

50. An apparatus in a video system comprising a processor and a non-transitory memory with instructions thereon, wherein the instructions upon execution by the processor, cause the processor to implement the method in any one of claims 1 to 49.

51. A computer program product stored on a non-transitory computer readable media, the computer program product including program code for carrying out the method in any one of claims 1 to 49.

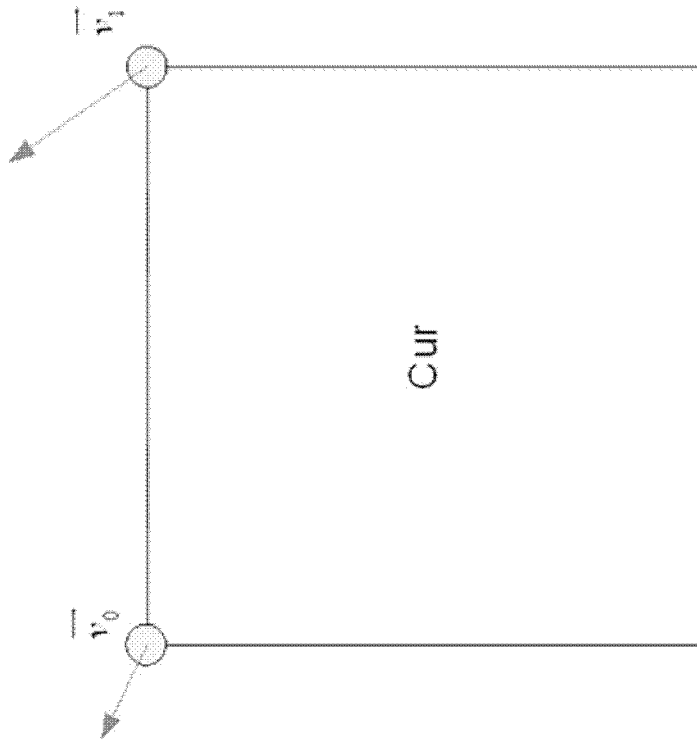


FIG. 1

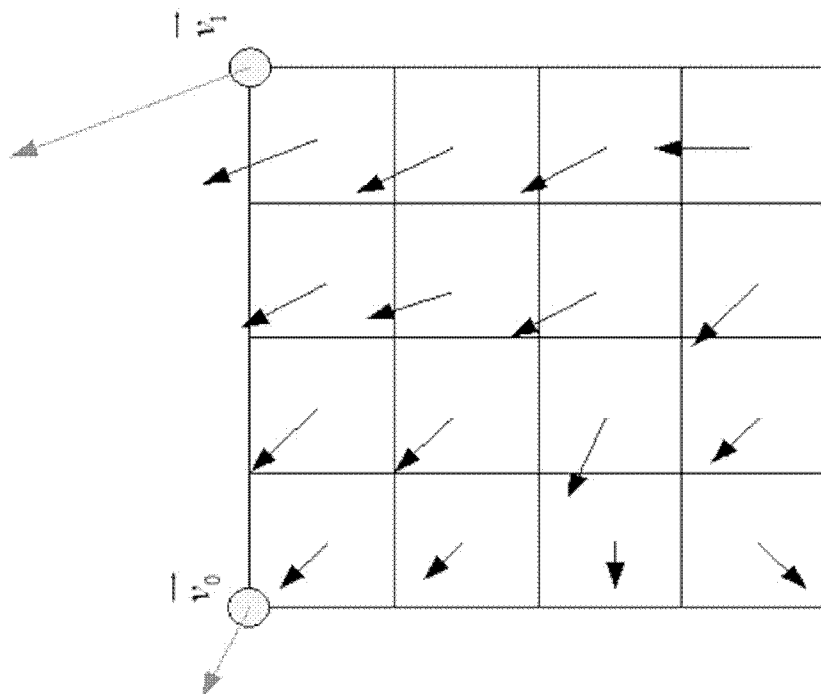


FIG. 2

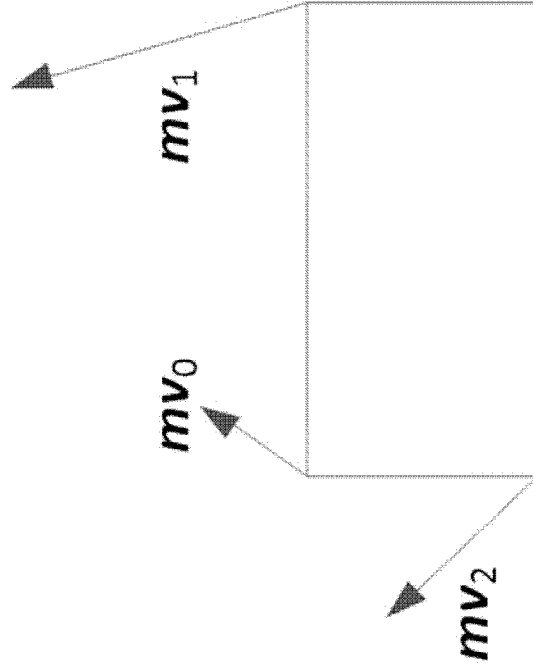


FIG. 3B

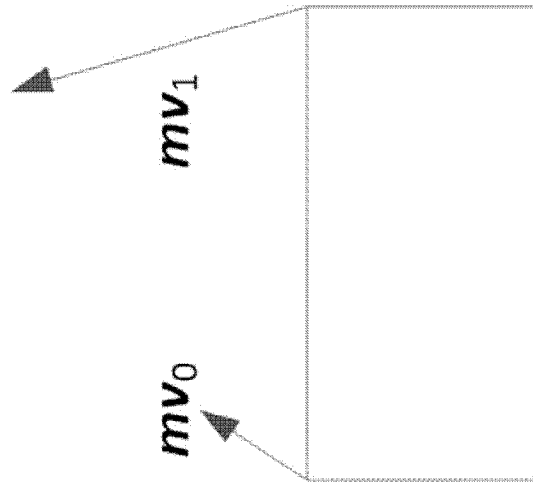


FIG. 3A

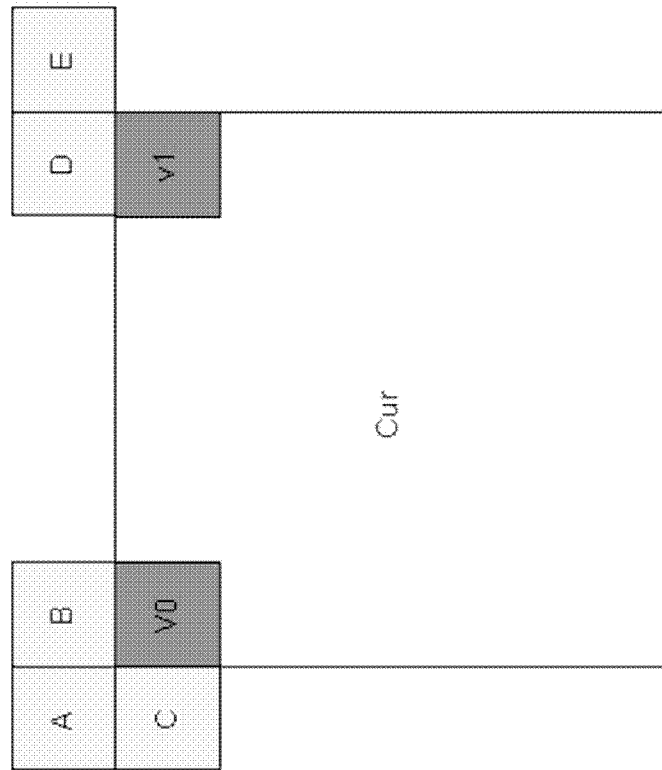


FIG. 4

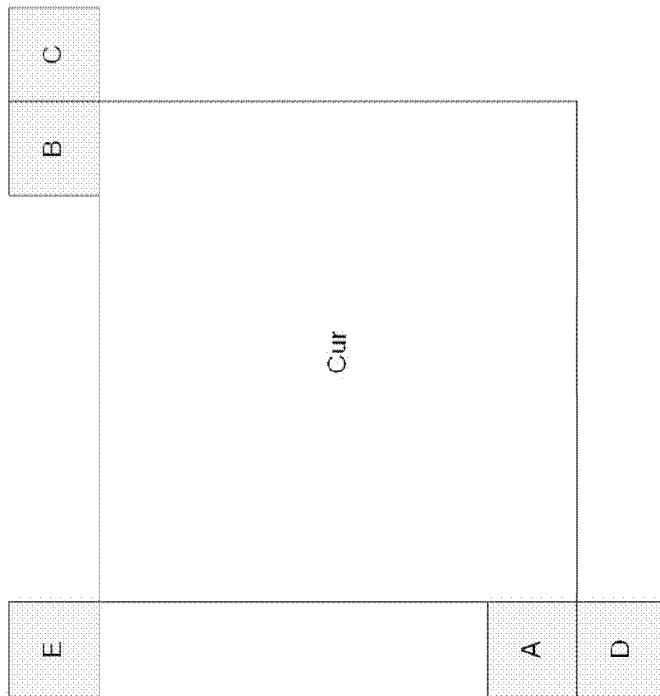


FIG. 5A

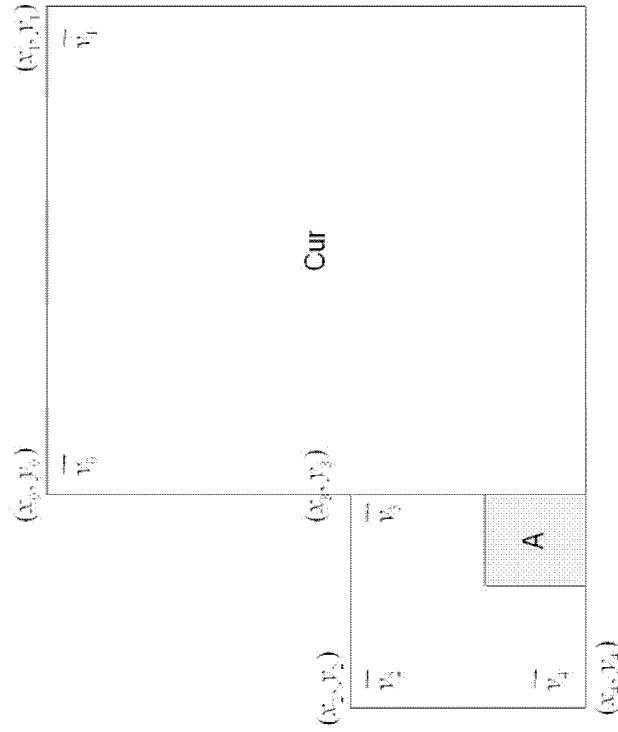


FIG. 5B

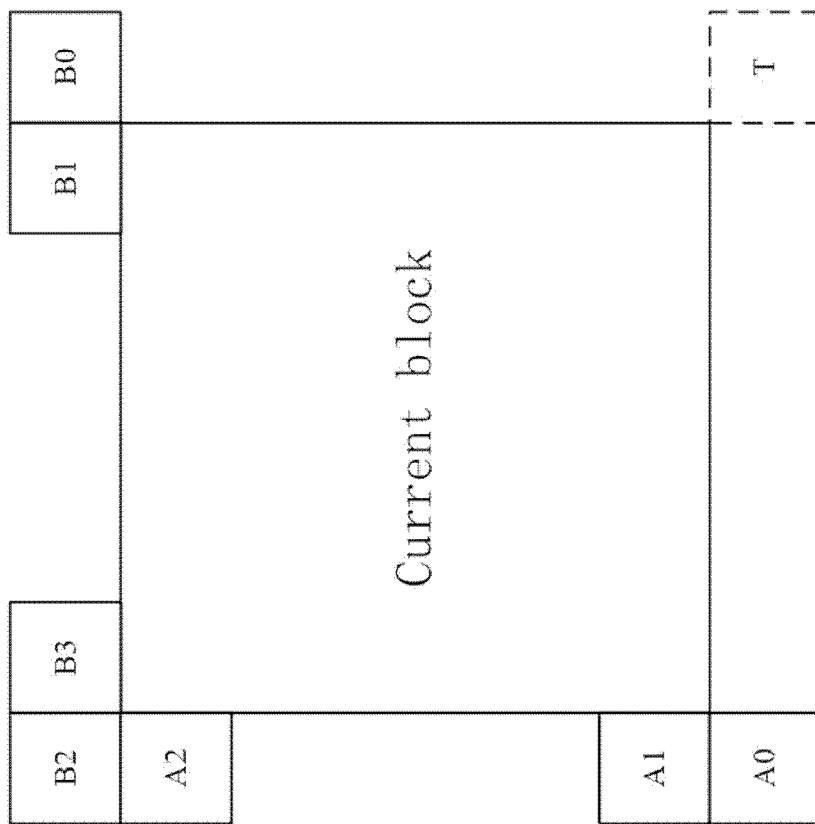


FIG. 6

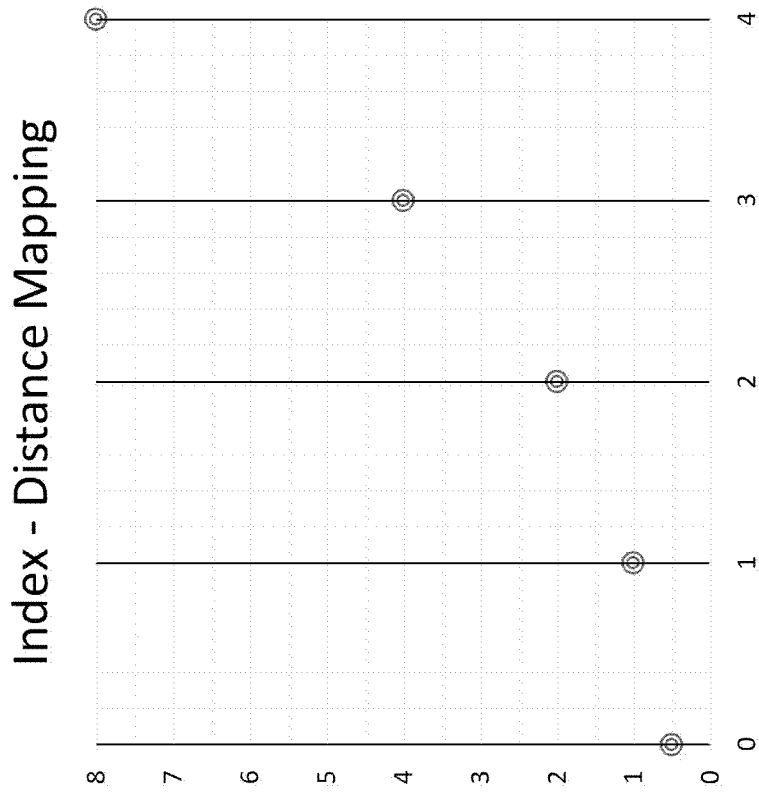


FIG. 7

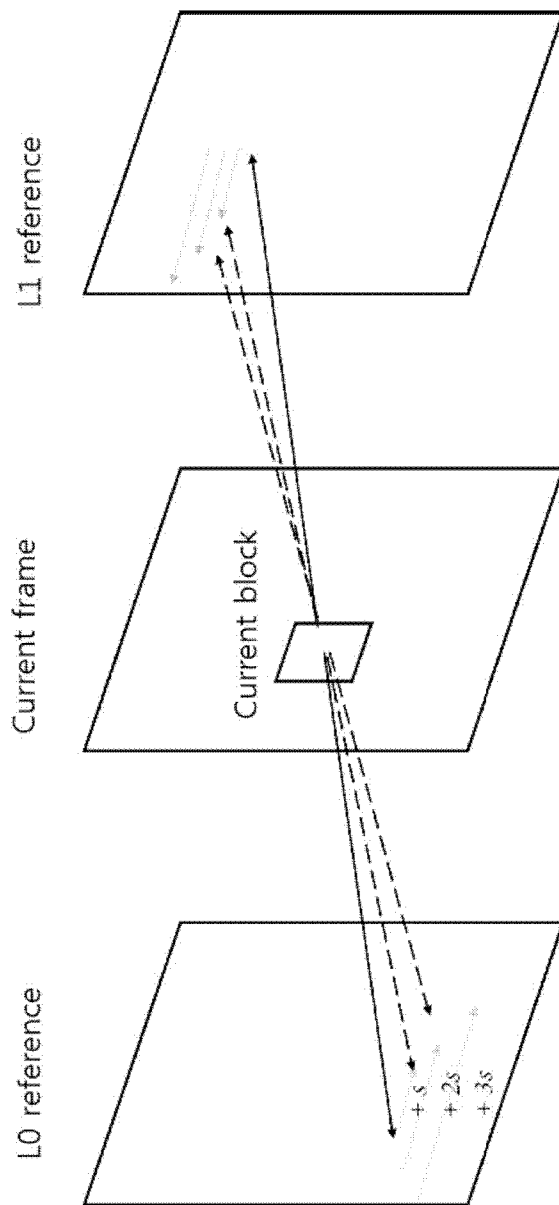


FIG. 8

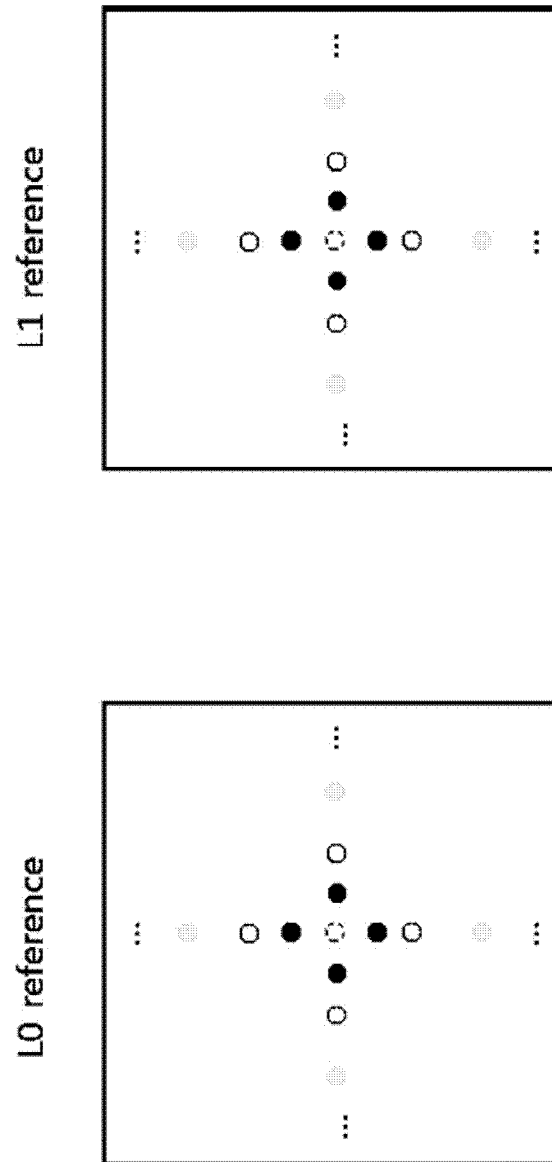


FIG. 9

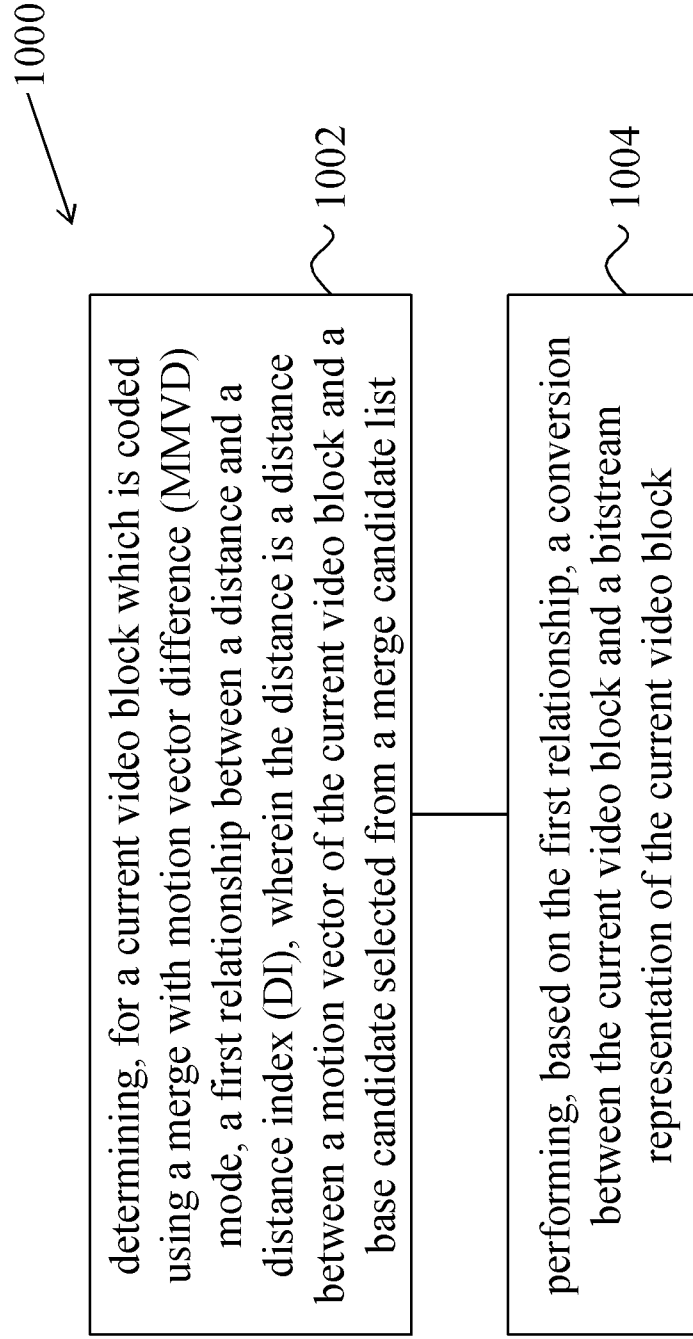


FIG. 10

performing a conversion between a current video block and a bitstream representation of the current video block, wherein the current video block is coded using a merge with motion vector difference (MMVD) mode; wherein the conversion comprises parsing MMVD side information from or writing the MMVD side information into the bitstream representation, wherein the MMVD side information comprises at least one of an MMVD flag indicating whether MMVD syntaxes are parsed, a first syntax element indicating a distance of MMVD between a motion vector of the current video block and a base candidate selected from a merge candidate list, a second syntax element indicating a direction of MMVD representing a direction of motion vector difference (MVD) relative to the base candidate

1100



1102



FIG. 11

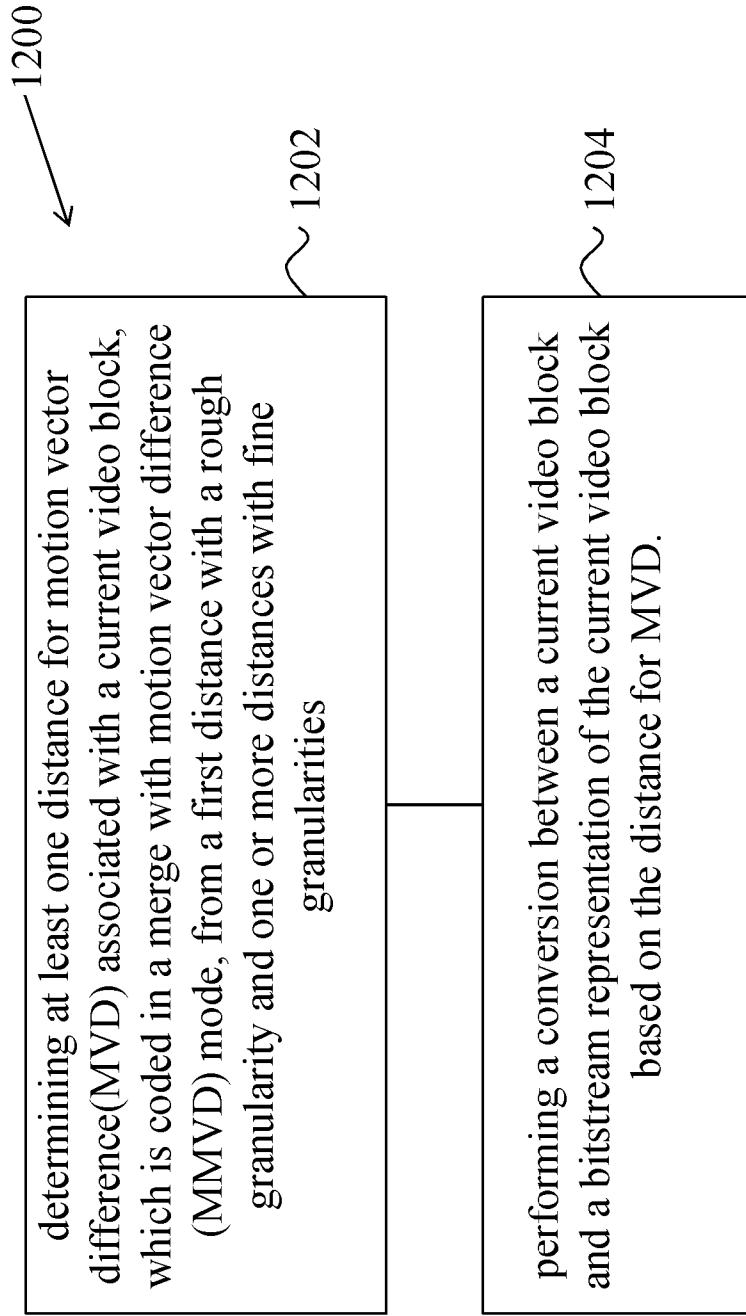


FIG. 12

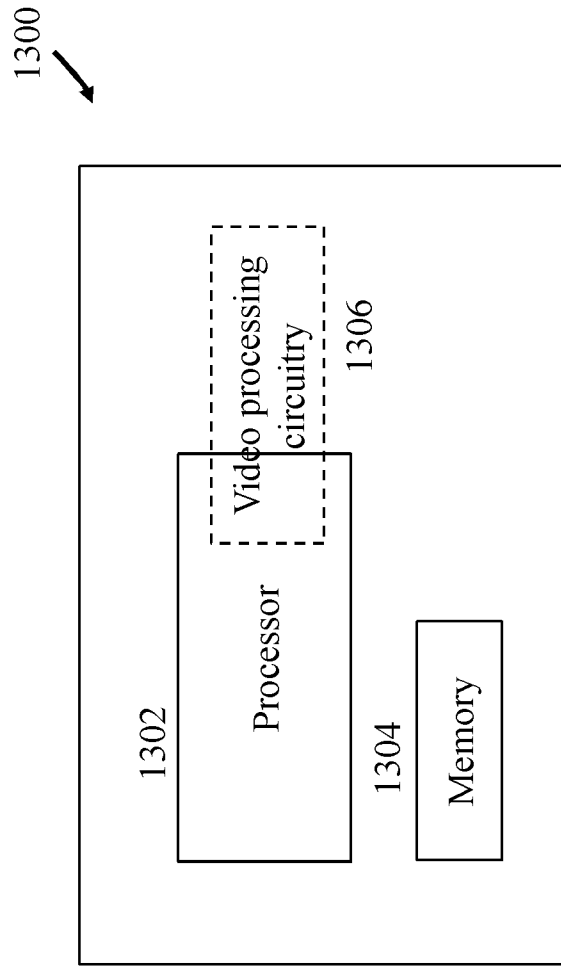


FIG. 13

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/130725

A. CLASSIFICATION OF SUBJECT MATTER		
H04N 19/52(2014.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
H04N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
WPI,EPODOC,CNKI,CNPAT,ITU,IEEE:MMVD,MVD,merge,motion,vector,distance,index,difference,current,candidate,block,video,rough,granularity,fine		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JEONG Seungsoo. "CE4 Ultimate motion vector expression (Test 4.5.4)" <i>Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11</i> , 03 October 2018 (2018-10-03), pages 1-6	1-51
A	JEONG Seungsoo. "Proposed WD for CE4 Ultimate motion vector expression (Test 4.5.4)" <i>Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11</i> , 03 October 2018 (2018-10-03), the whole document	1-51
A	US 2017006302 A1 (KT CORPORATION) 05 January 2017 (2017-01-05) the whole document	1-51
A	US 2017339425 A1 (SAMSUNG ELECTRONICS CO., LTD.) 23 November 2017 (2017-11-23) the whole document	1-51
A	CN 108886618 A (LG ELECTRONICS INC.) 23 November 2018 (2018-11-23) the whole document	1-51
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
12 March 2020		27 March 2020
Name and mailing address of the ISA/CN		Authorized officer
National Intellectual Property Administration, PRC 6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088 China		LANG, Yihong
Facsimile No. (86-10)62019451		Telephone No. 53961711

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/130725

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2017006302	A1	05 January 2017	KR	20150109282	A	01 October 2015
				CN	108632629	A	09 October 2018
				WO	2015142054	A1	24 September 2015
				CN	106105191	A	09 November 2016
				CN	108683922	A	19 October 2018
				CN	108965888	A	07 December 2018
				CN	108989815	A	11 December 2018
US	2017339425	A1	23 November 2017	WO	2016068685	A1	06 May 2016
				JP	2017537529	A	14 December 2017
				EP	3200461	A1	02 August 2017
				KR	20170078672	A	07 July 2017
				CN	107113440	A	29 August 2017
				MX	2017005651	A	29 June 2017
				IN	201727014814	A	30 June 2017
CN	108886618	A	23 November 2018	EP	3432578	A1	23 January 2019
				KR	20180123041	A	14 November 2018
				MX	2018011412	A	10 January 2019
				IN	201817034799	A	28 December 2018
				VN	61310	A	25 January 2019
				US	2019110061	A1	11 April 2019
				WO	2017164441	A1	28 September 2017