

SYNTAX FOR SUBPICTURE SIGNALING IN A VIDEO BITSTREAM

CROSS REFERENCE TO RELATED APPLICATIONS

[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/CN2019/109809, filed on October 2, 2019. For all purposes under the law, the entire disclosure of the aforementioned application is incorporated by reference as part of the disclosure of this application.

TECHNICAL FIELD

[002] This document is related to video and image coding and decoding technologies.

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 disclosed techniques may be used by video or image decoder or encoder embodiments for in which sub-picture based coding or decoding is performed.

[005] In one example aspect a method of video processing is disclosed. The method includes performing a conversion between a video comprising one or more pictures and a bitstream representation of the video. The bitstream representation is required to conform to a format rule that specifies that each picture is coded as one or more slices, and wherein the format rule prohibits samples in a picture from not being covered by any of the one or more slices.

[006] In another example aspect a method of video processing is disclosed. The method includes determining, for a conversion between a picture of a video and a bitstream representation of the video, a manner of signaling information of one or more slices in the picture according to a rule that is associated a number of tiles or a number of bricks in the picture. The method also includes performing the conversion based on the determining.

[007] In another example aspect a method of video processing is disclosed. The method includes performing a conversion between a picture of a video and a bitstream representation of the video according to a rule. The picture is coded in the bitstream representation as one or more slices and the rule specifies whether or how an address of a slice of the picture is included in the bitstream representation.

[008] In another example aspect a method of video processing is disclosed. The method includes determining, for a conversion between a picture of a video and a bitstream representation of the video, whether a syntax element indicating a filter operation that accesses samples across multiple bricks in the picture is enabled is included in the bitstream representation based on a number of tiles or a number of bricks in the picture. The method also includes performing the conversion based on the determining.

[009] In another example aspect a method of video processing is disclosed. The method includes performing a conversion between a picture of a video and a bitstream representation of the video, wherein the picture comprises one or more sub-pictures, and wherein a number of the one or more sub-pictures is indicated by a syntax element in the bitstream representation.

[0010] In another example aspect a method of video processing is disclosed. The method includes performing a conversion between a picture of a video that comprises one or more sub-pictures and a bitstream representation of the video. The bitstream representation conforms to a format rule that specifies that information about a sub-picture is included in the bitstream representation based on at least one of: (1) one or more corner positions of the sub-picture, or (2) a dimension of the sub-picture.

[0011] In another example aspect a method of video processing is disclosed. The method includes determining that a reference picture resampling tool is enabled for a conversion between a picture of a video and a bitstream representation of the video due to the picture being divided into one or more sub-pictures. The method also includes performing the conversion based on the determining.

[0012] In another example aspect a method of video processing is disclosed. The method includes performing a conversion between video comprising a video picture comprising one or more sub-pictures comprising one or more slices and a bitstream representation of the video. The bitstream representation conforms to a format rule that specifies that, for a sub-picture and a slice, in case that an index identifying the sub-picture is included in a header of the slice, an address field for the slice indicates an address of the slice in the sub-picture.

[0013] In another example aspect a method of video processing is disclosed. The method includes determining, for a video block in a first video region of a video, whether a position at which a temporal motion vector predictor is determined for a conversion between the video block and a bitstream representation of the current video block using an affine mode is within a second video region; and performing the conversion based on the determining.

[0014] In another example aspect, another method of video processing is disclosed. The method includes determining, for a video block in a first video region of a video, whether a position at which an integer sample in a reference picture is fetched for a conversion between the video block and a bitstream representation of the current video block is within a second video region, wherein the reference picture is not used in an interpolation process during the conversion; and performing the conversion based on the determining.

[0015] In another example aspect, another method of video processing is disclosed. The method includes determining, for a video block in a first video region of a video, whether a position at which a reconstructed luma sample value is fetched for a conversion between the video block and a bitstream representation of the current video block is within a second video region; and performing the conversion based on the determining.

[0016] In another example aspect, another method of video processing is disclosed. The method includes determining, for a video block in a first video region of a video, whether a position at which a check regarding splitting, depth derivation or split flag signaling for the video block is performed during a conversion between the video block and a bitstream representation of the current video block is within a second video region; and performing the conversion based on the determining.

[0017] In another example aspect, another method of video processing is disclosed. The method includes performing a conversion between a video comprising one or more video pictures comprising one or more video blocks, and a coded representation of the video, wherein the coded representation complies with a coding syntax requirement that the conversion is not to use sub-picture coding/decoding and a dynamic resolution conversion coding/decoding tool or a reference picture resampling tool within a video unit.

[0018] In another example aspect, another method of video processing is disclosed. The method includes performing a conversion between a video comprising one or more video pictures comprising one or more video blocks, and a coded representation of the video, wherein the coded

representation complies with a coding syntax requirement that a first syntax element `subpic_grid_idx[i][j]` is not larger than a second syntax element `max_subpics_minus1`.

[0019] In yet another example aspect, the above-described method may be implemented by a video encoder apparatus that comprises a processor.

[0020] In yet another example aspect, the above-described method may be implemented by a video decoder apparatus that comprises a processor.

[0021] 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.

[0022] These, and other, aspects are further described in the present document.

BRIEF DESCRIPTION OF THE DRAWINGS

[0023] FIG. 1 shows an example of region constraint in temporal motion vector prediction (TMVP) and sub-block TMVP.

[0024] FIG. 2 shows an example of a hierarchical motion estimation scheme.

[0025] FIG. 3 is a block diagram of an example of a hardware platform used for implementing techniques described in the present document.

[0026] FIG. 4 is a flowchart for an example method of video processing.

[0027] FIG. 5 shows an example of a picture with 18 by 12 luma CTUs that is partitioned into 12 tiles and 3 raster-scan slices (informative).

[0028] FIG. 6 shows an example of picture with 18 by 12 luma CTUs that is partitioned into 24 tiles and 9 rectangular slices (informative).

[0029] FIG. 7 shows an example of a picture that is partitioned into 4 tiles, 11 bricks, and 4 rectangular slices (informative).

[0030] FIG. 8 is a block diagram showing an example video processing system in which various techniques disclosed herein may be implemented.

[0031] FIG. 9 is a block diagram that illustrates an example video coding system.

[0032] FIG. 10 is a block diagram that illustrates an encoder in accordance with some embodiments of the present disclosure.

[0033] FIG. 11 is a block diagram that illustrates a decoder in accordance with some embodiments of the present disclosure.

[0034] FIG. 12 is a flowchart representation of a method for video processing in accordance with the present technology.

[0035] FIG. 13 is a flowchart representation of another method for video processing in accordance with the present technology.

[0036] FIG. 14 is a flowchart representation of another method for video processing in accordance with the present technology.

[0037] FIG. 15 is a flowchart representation of another method for video processing in accordance with the present technology.

[0038] FIG. 16 is a flowchart representation of another method for video processing in accordance with the present technology.

[0039] FIG. 17 is a flowchart representation of another method for video processing in accordance with the present technology.

[0040] FIG. 18 is a flowchart representation of another method for video processing in accordance with the present technology.

[0041] FIG. 19 is a flowchart representation of yet another method for video processing in accordance with the present technology.

DETAILED DESCRIPTION

[0042] The present document provides various techniques that can be used by a decoder of image or video bitstreams to improve the quality of decompressed or decoded digital video or images. For brevity, the term “video” is used herein to include both a sequence of pictures (traditionally called video) and individual images. Furthermore, a video encoder may also implement these techniques during the process of encoding in order to reconstruct decoded frames used for further encoding.

[0043] 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.

[0044] 1. Summary

[0045] This document is related to video coding technologies. Specifically, it is related to palette coding with employing base colors based representation 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.

[0046] 2. Initial Discussion

[0047] 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 [1,2]. 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.

[0048] 2.1 The region constraint in TMVP and sub-block TMVP in VVC

[0049] FIG. 1 illustrates example region constraint in TMVP and sub-block TMVP. In TMVP and sub-block TMVP, it is constrained that a temporal MV can only be fetched from the collocated CTU plus a column of 4×4 blocks as shown in FIG. 1.

[0050] 2.2 Example Sub-picture

[0051] In some embodiments, sub-picture-based coding techniques based on flexible tiling approach can be implemented. Summary of the sub-picture-based coding techniques includes the following:

[0052] (1) Pictures can be divided into sub-pictures.

[0053] (2) The indication of existence of sub-pictures is indicated in the SPS, along with other sequence-level information of sub-pictures.

[0054] (3) Whether a sub-picture is treated as a picture in the decoding process (excluding in-loop filtering operations) can be controlled by the bitstream.

[0055] (4) Whether in-loop filtering across sub-picture boundaries is disabled can be controlled by the bitstream for each sub-picture. The DBF, SAO, and ALF processes are updated for controlling of in-loop filtering operations across sub-picture boundaries.

[0056] (5) For simplicity, as a starting point, the sub-picture width, height, horizontal offset, and vertical offset are signalled in units of luma samples in SPS. Sub-picture boundaries are constrained to be slice boundaries.

[0057] (6) Treating a sub-picture as a picture in the decoding process (excluding in-loop filtering operations) is specified by slightly updating the `coding_tree_unit()` syntax, and updates to the following decoding processes:

[0058] - The derivation process for (advanced) temporal luma motion vector prediction

[0059] - The luma sample bilinear interpolation process

[0060] - The luma sample 8-tap interpolation filtering process

[0061] - The chroma sample interpolation process

[0062] (7) Sub-picture IDs are explicitly specified in the SPS and included in the tile group headers to enable extraction of sub-picture sequences without the need of changing VCL NAL units.

[0063] (8) Output sub-picture sets (OSPS) are proposed to specify normative extraction and conformance points for sub-pictures and sets thereof.

[0064] 2.3 Example Sub-pictures in Versatile Video Coding

Sequence parameter set RBSP syntax

<code>seq_parameter_set_rbsp()</code> {	Descriptor
<code>sps_decoding_parameter_set_id</code>	<code>u(4)</code>
<code>sps_video_parameter_set_id</code>	<code>u(4)</code>
...	
<code>pic_width_max_in_luma_samples</code>	<code>ue(v)</code>
<code>pic_height_max_in_luma_samples</code>	<code>ue(v)</code>
<code>subpics_present_flag</code>	<code>u(1)</code>
if(<code>subpics_present_flag</code>) {	
<code>max_subpics_minus1</code>	<code>u(8)</code>
<code>subpic_grid_col_width_minus1</code>	<code>u(v)</code>
<code>subpic_grid_row_height_minus1</code>	<code>u(v)</code>
for(<code>i = 0; i < NumSubPicGridRows; i++</code>)	
for(<code>j = 0; j < NumSubPicGridCols; j++</code>)	
<code>subpic_grid_idx[i][j]</code>	<code>u(v)</code>
for(<code>i = 0; i <= NumSubPics; i++</code>) {	
<code>subpic_treated_as_pic_flag[i]</code>	<code>u(1)</code>
<code>loop_filter_across_subpic_enabled_flag[i]</code>	<code>u(1)</code>
}	
}	
...	
}	

subpics_present_flag equal to 1 indicates that sub-picture parameters are present in the present in the SPS RBSP syntax. **subpics_present_flag** equal to 0 indicates that sub-picture parameters are not present in the present in the SPS RBSP syntax.

NOTE 2 – When a bitstream is the result of a sub-bitstream extraction process and contains only a subset of the sub-pictures of the input bitstream to the sub-bitstream extraction process, it might be required to set the value of **subpics_present_flag** equal to 1 in the RBSP of the SPSs.

max_subpics_minus1 plus 1 specifies the maximum number of sub-pictures that may be present in the CVS. **max_subpics_minus1** shall be in the range of 0 to 254. The value of 255 is reserved for future use by ITU-T | ISO/IEC.

subpic_grid_col_width_minus1 plus 1 specifies the width of each element of the sub-picture identifier grid in units of 4 samples. The length of the syntax element is $\text{Ceil}(\text{Log2}(\text{pic_width_max_in_luma_samples} / 4))$ bits.

The variable NumSubPicGridCols is derived as follows:

$$\text{NumSubPicGridCols} = (\text{pic_width_max_in_luma_samples} + \text{subpic_grid_col_width_minus1} * 4 + 3) / (\text{subpic_grid_col_width_minus1} * 4 + 4) \quad (7-5)$$

subpic_grid_row_height_minus1 plus 1 specifies the height of each element of the sub-picture identifier grid in units of 4 samples. The length of the syntax element is $\text{Ceil}(\text{Log2}(\text{pic_height_max_in_luma_samples} / 4))$ bits.

The variable NumSubPicGridRows is derived as follows:

$$\text{NumSubPicGridRows} = (\text{pic_height_max_in_luma_samples} + \text{subpic_grid_row_height_minus1} * 4 + 3) / (\text{subpic_grid_row_height_minus1} * 4 + 4) \quad (7-6)$$

subpic_grid_idx[i][j] specifies the sub-picture index of the grid position (i, j). The length of the syntax element is $\text{Ceil}(\text{Log2}(\text{max_subpics_minus1} + 1))$ bits.

The variables **SubPicTop[subpic_grid_idx[i][j]]**, **SubPicLeft[subpic_grid_idx[i][j]]**,

SubPicWidth[subpic_grid_idx[i][j]], **SubPicHeight[subpic_grid_idx[i][j]]**, and **NumSubPics** are derived as follows:

NumSubPics = 0

for(i = 0; i < NumSubPicGridRows; i++) {

 for(j = 0; j < NumSubPicGridCols; j++) {

 if (i == 0)

SubPicTop[subpic_grid_idx[i][j]] = 0

 else if(**subpic_grid_idx[i][j]** != **subpic_grid_idx[i - 1][j]**) {

SubPicTop[subpic_grid_idx[i][j]] = i

SubPicHeight[subpic_grid_idx[i - 1][j]] = i - **SubPicTop[subpic_grid_idx[i - 1][j]]**

 }

 if (j == 0)

```

        SubPicLeft[ subpic_grid_idx[ i ][ j ] ] = 0
        else if (subpic_grid_idx[ i ][ j ] != subpic_grid_idx[ i ][ j - 1 ] ) {
            SubPicLeft[ subpic_grid_idx[ i ][ j ] ] = j
            SubPicWidth[ subpic_grid_idx[ i ][ j ] ] = j - SubPicLeft[ subpic_grid_idx[ i ][ j - 1 ] ]
        }
        if ( i == NumSubPicGridRows - 1 )
            SubPicHeight[ subpic_grid_idx[ i ][ j ] ] = i - SubPicTop[ subpic_grid_idx[ i - 1 ][ j ] ] + 1
        if ( j == NumSubPicGridRows - 1 )
            SubPicWidth[ subpic_grid_idx[ i ][ j ] ] = j - SubPicLeft[ subpic_grid_idx[ i ][ j - 1 ] ] + 1
        if ( subpic_grid_idx[ i ][ j ] > NumSubPics )
            NumSubPics = subpic_grid_idx[ i ][ j ]
    }
}

```

subpic_treated_as_pic_flag[i] equal to 1 specifies that the i-th sub-picture of each coded picture in the CVS is treated as a picture in the decoding process excluding in-loop filtering operations. **subpic_treated_as_pic_flag[i]** equal to 0 specifies that the i-th sub-picture of each coded picture in the CVS is not treated as a picture in the decoding process excluding in-loop filtering operations. When not present, the value of **subpic_treated_as_pic_flag[i]** is inferred to be equal to 0.

loop_filter_across_subpic_enabled_flag[i] equal to 1 specifies that in-loop filtering operations may be performed across the boundaries of the i-th sub-picture in each coded picture in the CVS.

loop_filter_across_subpic_enabled_flag[i] equal to 0 specifies that in-loop filtering operations are not performed across the boundaries of the i-th sub-picture in each coded picture in the CVS. When not present, the value of **loop_filter_across_subpic_enabled_flag[i]** is inferred to be equal to 1.

It is a requirement of bitstream conformance that the following constraints apply:

- For any two sub-pictures subpicA and subpicB, when the index of subpicA is less than the index of subpicB, any coded NAL unit of subpicA shall succeed any coded NAL unit of subpicB in decoding order.
- The shapes of the sub-pictures shall be such that each sub-picture, when decoded, shall have its entire left boundary and entire top boundary consisting of picture boundaries or consisting of boundaries of previously decoded sub-pictures.

The list **CtbToSubPicIdx[ctbAddrRs]** for **ctbAddrRs** ranging from 0 to **PicSizeInCtbsY - 1**, inclusive, specifying the conversion from a CTB address in picture raster scan to a sub-picture index, is derived as follows:

```

for( ctbAddrRs = 0; ctbAddrRs < PicSizeInCtbsY; ctbAddrRs++ ) {
    posX = ctbAddrRs % PicWidthInCtbsY * CtbSizeY
    posY = ctbAddrRs / PicWidthInCtbsY * CtbSizeY

```

```

    CtbToSubPicIdx[ ctbAddrRs ] = -1

```

```

for( i = 0; CtbToSubPicIdx[ ctbAddrRs ] < 0 && i < NumSubPics; i++ ) {
    if( ( posX >= SubPicLeft[ i ] * ( subpic_grid_col_width_minus1 + 1 ) * 4 ) &&
        ( posX < ( SubPicLeft[ i ] + SubPicWidth[ i ] ) *
            ( subpic_grid_col_width_minus1 + 1 ) * 4 ) &&
        ( posY >= SubPicTop[ i ] *
            ( subpic_grid_row_height_minus1 + 1 ) * 4 ) &&
        ( posY < ( SubPicTop[ i ] + SubPicHeight[ i ] ) *
            ( subpic_grid_row_height_minus1 + 1 ) * 4 ) )
        CtbToSubPicIdx[ ctbAddrRs ] = i
    }
}

```

num_bricks_in_slice_minus1, when present, specifies the number of bricks in the slice minus 1. The value of num_bricks_in_slice_minus1 shall be in the range of 0 to NumBricksInPic - 1, inclusive. When rect_slice_flag is equal to 0 and single_brick_per_slice_flag is equal to 1, the value of num_bricks_in_slice_minus1 is inferred to be equal to 0. When single_brick_per_slice_flag is equal to 1, the value of num_bricks_in_slice_minus1 is inferred to be equal to 0.

The variable NumBricksInCurrSlice, which specifies the number of bricks in the current slice, and

SliceBrickIdx[i], which specifies the brick index of the i-th brick in the current slice, are derived as follows:

```

if( rect_slice_flag ) {
    sliceIdx = 0
    while( slice_address != slice_id[ sliceIdx ] )
        sliceIdx++
    NumBricksInCurrSlice = NumBricksInSlice[ sliceIdx ]
    brickIdx = TopLeftBrickIdx[ sliceIdx ]
    for( bIdx = 0; brickIdx <= BottomRightBrickIdx[ sliceIdx ]; brickIdx++ )
        if( BricksToSliceMap[ brickIdx ] == sliceIdx )
            SliceBrickIdx[ bIdx++ ] = brickIdx
    }
    else {
        NumBricksInCurrSlice = num_bricks_in_slice_minus1 + 1
        SliceBrickIdx[ 0 ] = slice_address
        for( i = 1; i < NumBricksInCurrSlice; i++ )
            SliceBrickIdx[ i ] = SliceBrickIdx[ i - 1 ] + 1
    }
}

```

The variables SubPicIdx, SubPicLeftBoundaryPos, SubPicTopBoundaryPos, SubPicRightBoundaryPos, and SubPicBotBoundaryPos are derived as follows:

SubPicIdx = CtbToSubPicIdx[CtbAddrBsToRs[FirstCtbAddrBs[SliceBrickIdx[0]]]]

if(subpic_treated_as_pic_flag[SubPicIdx]) {

```

SubPicLeftBoundaryPos = SubPicLeft[ SubPicIdx ] * ( subpic_grid_col_width_minus1 + 1 ) * 4
SubPicRightBoundaryPos      =      ( SubPicLeft[ SubPicIdx ] + SubPicWidth[ SubPicIdx ] ) *
      ( subpic_grid_col_width_minus1 + 1 ) * 4
SubPicTopBoundaryPos        =      SubPicTop[ SubPicIdx ] * ( subpic_grid_row_height_minus1 + 1 ) * 4
SubPicBotBoundaryPos        =      ( SubPicTop[ SubPicIdx ] + SubPicHeight[ SubPicIdx ] ) *
      ( subpic_grid_row_height_minus1 + 1 ) * 4
}
...

```

Derivation process for temporal luma motion vector prediction

Inputs to this process are:

- a luma location (xCb, yCb) of the top-left sample of the current luma coding block relative to the top-left luma sample of the current picture,
- a variable cbWidth specifying the width of the current coding block in luma samples,
- a variable cbHeight specifying the height of the current coding block in luma samples,
- a reference index refIdxLX, with X being 0 or 1.

Outputs of this process are:

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

The variable currCb specifies the current luma coding block at luma location (xCb, yCb).

The variables mvLXCol and availableFlagLXCol are derived as follows:

- If slice_temporal_mvp_enabled_flag is equal to 0 or (cbWidth * cbHeight) is less than or equal to 32, both components of mvLXCol are set equal to 0 and availableFlagLXCol is set equal to 0.
- Otherwise (slice_temporal_mvp_enabled_flag is equal to 1), the following ordered steps apply:
 1. The bottom right collocated motion vector and the bottom and right boundary sample locations are derived as follows:

$$xColBr = xCb + cbWidth \quad (8-421)$$

$$yColBr = yCb + cbHeight \quad (8-422)$$

$$rightBoundaryPos = subpic_treated_as_pic_flag[SubPicIdx] ?$$

$$SubPicRightBoundaryPos : pic_width_in_luma_samples - 1 \quad (8-423)$$

$$botBoundaryPos = subpic_treated_as_pic_flag[SubPicIdx] ?$$

$$SubPicBotBoundaryPos : pic_height_in_luma_samples - 1 \quad (8-424)$$

- If yCb >> CtbLog2SizeY is equal to yColBr >> CtbLog2SizeY, yColBr is less than or equal to botBoundaryPos and xColBr is less than or equal to rightBoundaryPos, the following applies:
 - The variable colCb specifies the luma coding block covering the modified location given by ((xColBr >> 3) << 3, (yColBr >> 3) << 3) inside the collocated picture specified by ColPic.

- The luma location ($xColCb$, $yColCb$) is set equal to the top-left sample of the collocated luma coding block specified by $colCb$ relative to the top-left luma sample of the collocated picture specified by $ColPic$.
- The derivation process for collocated motion vectors as specified in clause 8.5.2.12 is invoked with $currCb$, $colCb$, ($xColCb$, $yColCb$), $refIdxLX$ and $sbFlag$ set equal to 0 as inputs, and the output is assigned to $mvLXCol$ and $availableFlagLXCol$.

Otherwise, both components of $mvLXCol$ are set equal to 0 and $availableFlagLXCol$ is set equal to 0.

...

Luma sample bilinear interpolation process

Inputs to this process are:

- a luma location in full-sample units ($xInt_L$, $yInt_L$),
- a luma location in fractional-sample units ($xFrac_L$, $yFrac_L$),
- the luma reference sample array $refPicLX_L$.

Output of this process is a predicted luma sample value $predSampleLX_L$.

The variables $shift1$, $shift2$, $shift3$, $shift4$, $offset1$, $offset2$ and $offset3$ are derived as follows:

$$shift1 = BitDepth_Y - 6 \quad (8-453)$$

$$offset1 = 1 \ll (shift1 - 1) \quad (8-454)$$

$$shift2 = 4 \quad (8-455)$$

$$offset2 = 1 \ll (shift2 - 1) \quad (8-456)$$

$$shift3 = 10 - BitDepth_Y \quad (8-457)$$

$$shift4 = BitDepth_Y - 10 \quad (8-458)$$

$$offset4 = 1 \ll (shift4 - 1) \quad (8-459)$$

The variable $picW$ is set equal to $pic_width_in_luma_samples$ and the variable $picH$ is set equal to

$pic_height_in_luma_samples$.

The luma interpolation filter coefficients $fb_L[p]$ for each 1/16 fractional sample position p equal to $xFrac_L$

or $yFrac_L$ are specified in Table 8-10.

The luma locations in full-sample units ($xInt_i$, $yInt_i$) are derived as follows for $i = 0..1$:

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

$$xInt_i = Clip3(SubPicLeftBoundaryPos, SubPicRightBoundaryPos, xInt_L + i) \quad (8-460)$$

$$yInt_i = Clip3(SubPicTopBoundaryPos, SubPicBotBoundaryPos, yInt_L + i) \quad (8-461)$$

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

$$\begin{aligned} xInt_i &= Clip3(0, picW - 1, sps_ref_wraparound_enabled_flag \\ &\quad ClipH((sps_ref_wraparound_offset_minus1 + 1) * MinCbSizeY, picW, (xInt_L + i)) : (8-462) \\ &\quad xInt_L + i) \end{aligned}$$

$$yInt_i = Clip3(0, picH - 1, yInt_L + i) \quad (8-463)$$

...

Derivation process for subblock-based temporal merging candidates

Inputs to this process are:

- a luma location (x_{Cb} , y_{Cb}) of the top-left sample of the current luma coding block relative to the top-left luma sample of the current picture,
- a variable $cbWidth$ specifying the width of the current coding block in luma samples,
- a variable $cbHeight$ specifying the height of the current coding block in luma samples,
- the availability flag $availableFlagA_1$ of the neighbouring coding unit,
- the reference index $refIdxLXA_1$ of the neighbouring coding unit with X being 0 or 1,
- the prediction list utilization flag $predFlagLXA_1$ of the neighbouring coding unit with X being 0 or 1,
- the motion vector in 1/16 fractional-sample accuracy $mvLXA_1$ of the neighbouring coding unit with X being 0 or 1.

Outputs of this process are:

- the availability flag $availableFlagSbCol$,
- the number of luma coding subblocks in horizontal direction $numSbX$ and in vertical direction $numSbY$,
- the reference indices $refIdxL0SbCol$ and $refIdxL1SbCol$,
- the luma motion vectors in 1/16 fractional-sample accuracy $mvL0SbCol[xSbIdx][ySbIdx]$ and $mvL1SbCol[xSbIdx][ySbIdx]$ with $xSbIdx = 0..numSbX - 1$, $ySbIdx = 0..numSbY - 1$,
- the prediction list utilization flags $predFlagL0SbCol[xSbIdx][ySbIdx]$ and $predFlagL1SbCol[xSbIdx][ySbIdx]$ with $xSbIdx = 0..numSbX - 1$, $ySbIdx = 0..numSbY - 1$.

The availability flag $availableFlagSbCol$ is derived as follows.

- If one or more of the following conditions is true, $availableFlagSbCol$ is set equal to 0.
 - $slice_temporal_mvp_enabled_flag$ is equal to 0.
 - $sps_sbtmvp_enabled_flag$ is equal to 0.
 - $cbWidth$ is less than 8.
 - $cbHeight$ is less than 8.
- Otherwise, the following ordered steps apply:
 1. The location (x_{Ctb} , y_{Ctb}) of the top-left sample of the luma coding tree block that contains the current coding block and the location (x_{Ctr} , y_{Ctr}) of the below-right center sample of the current luma coding block are derived as follows:

$$x_{Ctb} = (x_{Cb} \gg CtuLog2Size) \ll CtuLog2Size \quad (8-542)$$

$$y_{Ctb} = (y_{Cb} \gg CtuLog2Size) \ll CtuLog2Size \quad (8-543)$$

$$x_{Ctr} = x_{Cb} + (cbWidth / 2) \quad (8-544)$$

$$y_{Ctr} = y_{Cb} + (cbHeight / 2) \quad (8-545)$$
 2. The luma location ($x_{ColCtrCb}$, $y_{ColCtrCb}$) is set equal to the top-left sample of the collocated luma coding block covering the location given by (x_{Ctr} , y_{Ctr}) inside $ColPic$ relative to the top-left luma sample of the collocated picture specified by $ColPic$.
 3. The derivation process for subblock-based temporal merging base motion data as specified in clause 8.5.5.4 is invoked with the location (x_{Ctb} , y_{Ctb}), the location ($x_{ColCtrCb}$, $y_{ColCtrCb}$), the availability flag

availableFlagA₁, and the prediction list utilization flag predFlagLXA₁, and the reference index refIdxLXA₁, and the motion vector mvLXA₁, with X being 0 and 1 as inputs and the motion vectors ctrMvLX, and the prediction list utilization flags ctrPredFlagLX of the collocated block, with X being 0 and 1, and the temporal motion vector tempMv as outputs.

4. The variable availableFlagSbCol is derived as follows:

- If both ctrPredFlagL0 and ctrPredFlagL1 are equal to 0, availableFlagSbCol is set equal to 0.
- Otherwise, availableFlagSbCol is set equal to 1.

When availableFlagSbCol is equal to 1, the following applies:

– The variables numSbX, numSbY, sbWidth, sbHeight and refIdxLXSbCol are derived as follows:

$$\text{numSbX} = \text{cbWidth} \gg 3 \quad (8-546)$$

$$\text{numSbY} = \text{cbHeight} \gg 3 \quad (8-547)$$

$$\text{sbWidth} = \text{cbWidth} / \text{numSbX} \quad (8-548)$$

$$\text{sbHeight} = \text{cbHeight} / \text{numSbY} \quad (8-549)$$

$$\text{refIdxLXSbCol} = 0 \quad (8-550)$$

– For $\text{xSbIdx} = 0.. \text{numSbX} - 1$ and $\text{ySbIdx} = 0.. \text{numSbY} - 1$, the motion vectors $\text{mvLXSbCol}[\text{xSbIdx}][\text{ySbIdx}]$ and prediction list utilization flags $\text{predFlagLXSbCol}[\text{xSbIdx}][\text{ySbIdx}]$ are derived as follows:

– The luma location (xSb, ySb) specifying the top-left sample of the current coding subblock relative to the top-left luma sample of the current picture is derived as follows:

$$\text{xSb} = \text{xCb} + \text{xSbIdx} * \text{sbWidth} + \text{sbWidth} / 2 \quad (8-551)$$

$$\text{ySb} = \text{yCb} + \text{ySbIdx} * \text{sbHeight} + \text{sbHeight} / 2 \quad (8-552)$$

– The location (xColSb, yColSb) of the collocated subblock inside ColPic is derived as follows.

– The following applies:

$$\begin{aligned} \text{yColSb} &= \text{Clip3}(\text{yCtb}, \\ &\quad \text{Min}(\text{CurPicHeightInSamplesY} - 1, \text{yCtb} + (1 \ll \text{CtbLog2SizeY}) - 1), \\ &\quad \text{ySb} + (\text{tempMv}[1] \gg 4)) \end{aligned} \quad (8-553)$$

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

$$\begin{aligned} \text{xColSb} &= \text{Clip3}(\text{xCtb}, \\ &\quad \text{Min}(\text{SubPicRightBoundaryPos}, \text{xCtb} + (1 \ll \text{CtbLog2SizeY}) + 3), \\ &\quad \text{xSb} + (\text{tempMv}[0] \gg 4)) \end{aligned} \quad (8-554)$$

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

$$\begin{aligned} \text{xColSb} &= \text{Clip3}(\text{xCtb}, \\ &\quad \text{Min}(\text{CurPicWidthInSamplesY} - 1, \text{xCtb} + (1 \ll \text{CtbLog2SizeY}) + 3), \\ &\quad \text{xSb} + (\text{tempMv}[0] \gg 4)) \end{aligned} \quad (8-555)$$

...

Derivation process for subblock-based temporal merging base motion data

Inputs to this process are:

- the location (x_{Ctb} , y_{Ctb}) of the top-left sample of the luma coding tree block that contains the current coding block,
- the location ($x_{ColCtrCb}$, $y_{ColCtrCb}$) of the top-left sample of the collocated luma coding block that covers the below-right center sample.
- the availability flag $availableFlagA_1$ of the neighbouring coding unit,
- the reference index $refIdxLXA_1$ of the neighbouring coding unit,
- the prediction list utilization flag $predFlagLXA_1$ of the neighbouring coding unit,
- the motion vector in 1/16 fractional-sample accuracy $mvLXA_1$ of the neighbouring coding unit.

Outputs of this process are:

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

The variable $tempMv$ is set as follows:

$$tempMv[0] = 0 \quad (8-558)$$

$$tempMv[1] = 0 \quad (8-559)$$

The variable $currPic$ specifies the current picture.

When $availableFlagA_1$ is equal to TRUE, the following applies:

- If all of the following conditions are true, $tempMv$ is set equal to $mvL0A_1$:
 - $predFlagL0A_1$ is equal to 1,
 - $DiffPicOrderCnt(ColPic, RefPicList[0][refIdxL0A_1])$ is equal to 0,
- Otherwise, if all of the following conditions are true, $tempMv$ is set equal to $mvL1A_1$:
 - $slice_type$ is equal to B,
 - $predFlagL1A_1$ is equal to 1,
 - $DiffPicOrderCnt(ColPic, RefPicList[1][refIdxL1A_1])$ is equal to 0.

The location (x_{ColCb} , y_{ColCb}) of the collocated block inside $ColPic$ is derived as follows.

- The following applies:

$$y_{ColCb} = Clip3(y_{Ctb}, Min(CurPicHeightInSamplesY - 1, y_{Ctb} + (1 \ll CtbLog2SizeY) - 1), y_{ColCtrCb} + (tempMv[1] \gg 4)) \quad (8-560)$$

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

$$x_{ColCb} = Clip3(x_{Ctb}, Min(SubPicRightBoundaryPos, x_{Ctb} + (1 \ll CtbLog2SizeY) + 3), x_{ColCtrCb} + (tempMv[0] \gg 4)) \quad (8-561)$$

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

$$x_{ColCb} = Clip3(x_{Ctb}, Min(CurPicWidthInSamplesY - 1, x_{Ctb} + (1 \ll CtbLog2SizeY) + 3), x_{ColCtrCb} + (tempMv[0] \gg 4)) \quad (8-562)$$

...

Luma sample interpolation filtering process

Inputs to this process are:

- a luma location in full-sample units (x_{Int_L} , y_{Int_L}),
- a luma location in fractional-sample units (x_{Frac_L} , y_{Frac_L}),
- a luma location in full-sample units (x_{SbInt_L} , y_{SbInt_L}) specifying the top-left sample of the bounding block for reference sample padding relative to the top-left luma sample of the reference picture,
- the luma reference sample array $refPicLX_L$,
- the half sample interpolation filter index $hpelIdx$,
- a variable $sbWidth$ specifying the width of the current subblock,
- a variable $sbHeight$ specifying the height of the current subblock,
- a luma location (x_{Sb} , y_{Sb}) specifying the top-left sample of the current subblock relative to the top-left luma sample of the current picture,

Output of this process is a predicted luma sample value $predSampleLX_L$.The variables $shift1$, $shift2$ and $shift3$ are derived as follows:

- The variable $shift1$ is set equal to $\text{Min}(4, \text{BitDepth}_Y - 8)$, the variable $shift2$ is set equal to 6 and the variable $shift3$ is set equal to $\text{Max}(2, 14 - \text{BitDepth}_Y)$.
- The variable $picW$ is set equal to $pic_width_in_luma_samples$ and the variable $picH$ is set equal to $pic_height_in_luma_samples$.

The luma interpolation filter coefficients $f_L[p]$ for each 1/16 fractional sample position p equal to x_{Frac_L} or y_{Frac_L} are derived as follows:

- If $\text{MotionModelIdx}[x_{Sb}][y_{Sb}]$ is greater than 0, and $sbWidth$ and $sbHeight$ are both equal to 4, the luma interpolation filter coefficients $f_L[p]$ are specified in Table 8-12.
- Otherwise, the luma interpolation filter coefficients $f_L[p]$ are specified in Table 8-11 depending on $hpelIdx$.

The luma locations in full-sample units (x_{Int_i} , y_{Int_i}) are derived as follows for $i = 0..7$:

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

$$x_{Int_i} = \text{Clip3}(SubPicLeftBoundaryPos, SubPicRightBoundaryPos, x_{Int_L} + i - 3) \quad (8-771)$$

$$y_{Int_i} = \text{Clip3}(SubPicTopBoundaryPos, SubPicBotBoundaryPos, y_{Int_L} + i - 3) \quad (8-772)$$

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

$$x_{Int_i} = \text{Clip3}(0, picW - 1, sps_ref_wraparound_enabled_flag ? \text{ClipH}((sps_ref_wraparound_offset_minus1 + 1) * \text{MinCbSizeY}, picW, x_{Int_L} + i - 3) : (8-773))$$

$$x_{Int_L} + i - 3)$$

$$y_{Int_i} = \text{Clip3}(0, picH - 1, y_{Int_L} + i - 3) \quad (8-774)$$

...

Chroma sample interpolation process

Inputs to this process are:

- a chroma location in full-sample units (x_{Int_C} , y_{Int_C}),

- a chroma location in 1/32 fractional-sample units (x_{FracC} , y_{FracC}),
- a chroma location in full-sample units (x_{SbIntC} , y_{SbIntC}) specifying the top-left sample of the bounding block for reference sample padding relative to the top-left chroma sample of the reference picture,
- a variable sbWidth specifying the width of the current subblock,
- a variable sbHeight specifying the height of the current subblock,
- the chroma reference sample array refPicLXC .

Output of this process is a predicted chroma sample value predSampleLXC

The variables shift1 , shift2 and shift3 are derived as follows:

- The variable shift1 is set equal to $\text{Min}(4, \text{BitDepthC} - 8)$, the variable shift2 is set equal to 6 and the variable shift3 is set equal to $\text{Max}(2, 14 - \text{BitDepthC})$.
- The variable picWC is set equal to $\text{pic_width_in_luma_samples} / \text{SubWidthC}$ and the variable picHC is set equal to $\text{pic_height_in_luma_samples} / \text{SubHeightC}$.

The chroma interpolation filter coefficients $f_c[p]$ for each 1/32 fractional sample position p equal to x_{FracC} or y_{FracC} are specified in Table 8-13.

The variable xOffset is set equal to $(\text{sps_ref_wraparound_offset_minus1} + 1) * \text{MinCbSizeY} / \text{SubWidthC}$.

The chroma locations in full-sample units (x_{Int_i} , y_{Int_i}) are derived as follows for $i = 0..3$:

- If $\text{subpic_treated_as_pic_flag}[\text{SubPicIdx}]$ is equal to 1, the following applies:

$$x_{\text{Int}_i} = \text{Clip3}(\text{SubPicLeftBoundaryPos} / \text{SubWidthC}, \text{SubPicRightBoundaryPos} / \text{SubWidthC}, x_{\text{Int}_L} + i) \quad (8-785)$$

$$y_{\text{Int}_i} = \text{Clip3}(\text{SubPicTopBoundaryPos} / \text{SubHeightC}, \text{SubPicBotBoundaryPos} / \text{SubHeightC}, y_{\text{Int}_L} + i) \quad (8-786)$$
- Otherwise ($\text{subpic_treated_as_pic_flag}[\text{SubPicIdx}]$ is equal to 0), the following applies:

$$x_{\text{Int}_i} = \text{Clip3}(0, \text{picWC} - 1, \text{sps_ref_wraparound_enabled_flag} ? \text{ClipH}(\text{xOffset}, \text{picWC}, x_{\text{IntC}} + i - 1) : (8-787)$$

$$x_{\text{IntC}} + i - 1)$$

$$y_{\text{Int}_i} = \text{Clip3}(0, \text{picHC} - 1, y_{\text{IntC}} + i - 1) \quad (8-788)$$

[0065] 2.4 Example Encoder-only GOP-based temporal filter

[0066] In some embodiments, an encoder-only temporal filter can be implemented. The filtering is done at the encoder side as a pre-processing step. Source pictures before and after the selected picture to encode are read and a block based motion compensation method relative to the selected picture is applied on those source pictures. Samples in the selected picture are temporally filtered using sample values after motion compensation.

[0067] The overall filter strength is set depending on the temporal sub layer of the selected picture as well as the QP. Only pictures at temporal sub layers 0 and 1 are filtered and pictures of layer 0 are filter by a stronger filter than pictures of layer 1. The per sample filter strength is adjusted

depending on the difference between the sample value in the selected picture and the co-located samples in motion compensated pictures so that small differences between a motion compensated picture and the selected picture are filtered more strongly than larger differences.

[0068] GOP Based Temporal Filter

[0069] A temporal filter is introduced directly after reading picture and before encoding. Below are the steps described in more detail.

[0070] Operation 1: Pictures are read by the encoder

[0071] Operation 2: If a picture is low enough in the coding hierarchy, it is filtered before encoding. Otherwise the picture is encoded without filtering. RA pictures with $\text{POC} \% 8 == 0$ are filtered as well as LD pictures with $\text{POC} \% 4 == 0$. AI pictures are never filtered.

[0072] The overall filter strength, s_o , is set according to the equation below for RA.

$$s_o(n) = \begin{cases} 1.5, & n \bmod 16 = 0 \\ 0.95, & n \bmod 16 \neq 0 \end{cases}$$

[0073] where n is the number of pictures read.

[0074] For the LD case, $s_o(n) = 0.95$ is used.

[0075] Operation 3: Two pictures before and/or after the selected picture (referred to as original picture further down) are read. In the edge cases e.g. if is the first picture or close to the last picture, only the available pictures are read.

[0076] Operation 4: Motion of the read pictures before and after, relative to the original picture is estimated per 8x8 picture block.

[0077] A hierarchical motion estimation scheme is used and the layers L0, L1 and L2, are illustrated in FIG. 2. Subsampled pictures are generated by averaging each 2x2 block for all read pictures and the original picture, e.g. L1 in FIG. 1. L2 is derived from L1 using the same subsampling method.

[0078] FIG. 2 shows examples of different layers of the hierarchical motion estimation. L0 is the original resolution. L1 is a subsampled version of L0. L2 is a subsampled version of L1.

[0079] First, motion estimation is done for each 16x16 block in L2. The squared difference is calculated for each selected motion vector and the motion vector corresponding to the smallest difference is selected. The selected motion vector is then used as initial value when estimating the motion in L1. Then the same is done for estimating motion in L0. As a final step, subpixel motion is estimated for each 8x8 block by using an interpolation filter on L0.

[0080] The VTM 6-tap interpolation filter can used:

0:	0,	0,	64,	0,	0,	0
1:	1,	-3,	64,	4,	-2,	0
2:	1,	-6,	62,	9,	-3,	1
3:	2,	-8,	60,	14,	-5,	1
4:	2,	-9,	57,	19,	-7,	2
5:	3,	-10,	53,	24,	-8,	2
6:	3,	-11,	50,	29,	-9,	2
7:	3,	-11,	44,	35,	-10,	3
8:	1,	-7,	38,	38,	-7,	1
9:	3,	-10,	35,	44,	-11,	3
10:	2,	-9,	29,	50,	-11,	3
11:	2,	-8,	24,	53,	-10,	3
12:	2,	-7,	19,	57,	-9,	2
13:	1,	-5,	14,	60,	-8,	2
14:	1,	-3,	9,	62,	-6,	1
15:	0,	-2,	4,	64,	-3,	1

[0081] Operation 5: Motion compensation is applied on the pictures before and after the original picture according to the best matching motion for each block, e.g., so that the sample coordinates of the original picture in each block have the best matching coordinates in the referenced pictures.

[0082] Operation 6: The samples of the processed one by one for the luma and chroma channels as described in the following steps.

[0083] Operation 7: The new sample value, I_n , is calculated using the following formula.

$$I_n = \frac{I_o + \sum_{i=0}^3 w_r(i, a) I_r(i)}{1 + \sum_{i=0}^3 w_r(i, a)}$$

[0084] Where I_o is the sample value of the original sample, $I_r(i)$ is the intensity of the corresponding sample of motion compensated picture i and $w_r(i, a)$ is the weight of motion compensated picture i when the number of available motion compensated pictures is a .

[0085] In the luma channel, the weights, $w_r(i, a)$, is defined as follows:

$$w_r(i, a) = s_l s_o(n) s_r(i, a) e^{-\frac{\Delta I(i)^2}{2\sigma_l(QP)^2}}$$

[0086] Where

$$s_l = 0.4$$

$$s_r(i, 2) = \begin{cases} 1.2, & i = 0 \\ 1.0, & i = 1 \end{cases}$$

$$s_r(i, 4) = \begin{cases} 0.60, & i = 0 \\ 0.85, & i = 1 \\ 0.85, & i = 2 \\ 0.60, & i = 3 \end{cases}$$

[0087] For all other cases of i , and a : $s_r(i, a) = 0.3$

$$\sigma_l(QP) = 3 * (QP - 10)$$

$$\Delta I(i) = I_r(i) - I_o$$

[0088] For the chroma channels, the weights, $w_r(i, a)$, is defined as follows:

$$w_r(i, a) = s_c s_o(n) s_r(i, a) e^{-\frac{\Delta I(i)^2}{2\sigma_c^2}}$$

[0089] Where $s_c = 0.55$ and $\sigma_c = 30$

[0090] Operation 8: The filter is applied for the current sample. The resulting sample value is stored separately.

[0091] Operation 9: The filtered picture is encoded.

[0092] 2.5 Example Picture partitions (tiles, bricks, slices)

[0093] In some embodiments, a picture is divided into one or more tile rows and one or more tile columns. A tile is a sequence of CTUs that covers a rectangular region of a picture.

[0094] A tile is divided into one or more bricks, each of which consisting of a number of CTU rows within the tile.

[0095] A tile that is not partitioned into multiple bricks is also referred to as a brick. However, a brick that is a true subset of a tile is not referred to as a tile.

[0096] A slice either contains a number of tiles of a picture or a number of bricks of a tile.

[0097] A sub-picture contains one or more slices that collectively cover a rectangular region of a picture.

[0098] Two modes of slices are supported, namely the raster-scan slice mode and the rectangular slice mode. In the raster-scan slice mode, a slice contains a sequence of tiles in a tile raster scan of

a picture. In the rectangular slice mode, a slice contains a number of bricks of a picture that collectively form a rectangular region of the picture. The bricks within a rectangular slice are in the order of brick raster scan of the slice.

[0099] FIG. 5 shows an example of raster-scan slice partitioning of a picture, where the picture is divided into 12 tiles and 3 raster-scan slices.

[00100] FIG. 6 shows an example of rectangular slice partitioning of a picture, where the picture is divided into 24 tiles (6 tile columns and 4 tile rows) and 9 rectangular slices.

[00101] FIG. 7 shows an example of a picture partitioned into tiles, bricks, and rectangular slices, where the picture is divided into 4 tiles (2 tile columns and 2 tile rows), 11 bricks (the top-left tile contains 1 brick, the top-right tile contains 5 bricks, the bottom-left tile contains 2 bricks, and the bottom-right tile contain 3 bricks), and 4 rectangular slices.

Picture parameter set RBSP syntax

pic_parameter_set_rbsp() {	Descriptor
...	
single_tile_in_pic_flag	u(1)
if(!single_tile_in_pic_flag) {	
uniform_tile_spacing_flag	u(1)
if(uniform_tile_spacing_flag) {	
tile_cols_width_minus1	ue(v)
tile_rows_height_minus1	ue(v)
} else {	
num_tile_columns_minus1	ue(v)
num_tile_rows_minus1	ue(v)
for(i = 0; i < num_tile_columns_minus1; i++)	
tile_column_width_minus1[i]	ue(v)
for(i = 0; i < num_tile_rows_minus1; i++)	
tile_row_height_minus1[i]	ue(v)
}	
brick_splitting_present_flag	u(1)
if(uniform_tile_spacing_flag && brick_splitting_present_flag)	
num_tiles_in_pic_minus1	ue(v)
for(i = 0; brick_splitting_present_flag && i <= num_tiles_in_pic_minus1 + 1; i++) {	
if(RowHeight[i] > 1)	
brick_split_flag[i]	u(1)

if(brick_split_flag[i]) {	
if(RowHeight[i] > 2)	
uniform_brick_spacing_flag[i]	u(1)
if(uniform_brick_spacing_flag[i])	
brick_height_minus1[i]	ue(v)
else {	
num_brick_rows_minus2[i]	ue(v)
for(j = 0; j <= num_brick_rows_minus2[i]; j++)	
brick_row_height_minus1[i][j]	ue(v)
}	
}	
}	
single_brick_per_slice_flag	u(1)
if(!single_brick_per_slice_flag)	
rect_slice_flag	u(1)
if(rect_slice_flag && !single_brick_per_slice_flag) {	
num_slices_in_pic_minus1	ue(v)
bottom_right_brick_idx_length_minus1	ue(v)
for(i = 0; i < num_slices_in_pic_minus1; i++) {	
bottom_right_brick_idx_delta[i]	u(v)
brick_idx_delta_sign_flag[i]	u(1)
}	
}	
loop_filter_across_bricks_enabled_flag	u(1)
if(loop_filter_across_bricks_enabled_flag)	
loop_filter_across_slices_enabled_flag	u(1)
}	
if(rect_slice_flag) {	
signalled_slice_id_flag	u(1)
if(signalled_slice_id_flag) {	
signalled_slice_id_length_minus1	ue(v)
for(i = 0; i <= num_slices_in_pic_minus1; i++)	
slice_id[i]	u(v)
}	
}	
...	

slice_header() {	Descriptor
slice_pic_parameter_set_id	ue(v)
if(rect_slice_flag NumBricksInPic > 1)	
slice_address	u(v)
if(!rect_slice_flag && !single_brick_per_slice_flag)	
num_bricks_in_slice_minus1	ue(v)
non_reference_picture_flag	u(1)
slice_type	ue(v)
...	

single_tile_in_pic_flag equal to 1 specifies that there is only one tile in each picture referring to the PPS.

single_tile_in_pic_flag equal to 0 specifies that there is more than one tile in each picture referring to the PPS.

NOTE – In absence of further brick splitting within a tile, the whole tile is referred to as a brick. When a picture contains only a single tile without further brick splitting, it is referred to as a single brick.

It is a requirement of bitstream conformance that the value of **single_tile_in_pic_flag** shall be the same for all PPSs that are referred to by coded pictures within a CVS.

uniform_tile_spacing_flag equal to 1 specifies that tile column boundaries and likewise tile row boundaries are distributed uniformly across the picture and signalled using the syntax elements **tile_cols_width_minus1** and **tile_rows_height_minus1**. **uniform_tile_spacing_flag** equal to 0 specifies that tile column boundaries and likewise tile row boundaries may or may not be distributed uniformly across the picture and signalled using the syntax elements **num_tile_columns_minus1** and **num_tile_rows_minus1** and a list of syntax element pairs **tile_column_width_minus1[i]** and **tile_row_height_minus1[i]**. When not present, the value of **uniform_tile_spacing_flag** is inferred to be equal to 1.

tile_cols_width_minus1 plus 1 specifies the width of the tile columns excluding the right-most tile column of the picture in units of CTBs when **uniform_tile_spacing_flag** is equal to 1. The value of **tile_cols_width_minus1** shall be in the range of 0 to **PicWidthInCtbsY** – 1, inclusive. When not present, the value of **tile_cols_width_minus1** is inferred to be equal to **PicWidthInCtbsY** – 1.

tile_rows_height_minus1 plus 1 specifies the height of the tile rows excluding the bottom tile row of the picture in units of CTBs when **uniform_tile_spacing_flag** is equal to 1. The value of **tile_rows_height_minus1** shall be in the range of 0 to **PicHeightInCtbsY** – 1, inclusive. When not present, the value of **tile_rows_height_minus1** is inferred to be equal to **PicHeightInCtbsY** – 1.

num_tile_columns_minus1 plus 1 specifies the number of tile columns partitioning the picture when **uniform_tile_spacing_flag** is equal to 0. The value of **num_tile_columns_minus1** shall be in the range of 0 to **PicWidthInCtbsY** – 1, inclusive. If **single_tile_in_pic_flag** is equal to 1, the value of **num_tile_columns_minus1** is inferred to be equal to 0. Otherwise, when **uniform_tile_spacing_flag** is equal to 1, the value of **num_tile_columns_minus1** is inferred as specified in clause 6.5.1.

num_tile_rows_minus1 plus 1 specifies the number of tile rows partitioning the picture when **uniform_tile_spacing_flag** is equal to 0. The value of **num_tile_rows_minus1** shall be in the range of 0 to **PicHeightInCtbsY** - 1, inclusive. If **single_tile_in_pic_flag** is equal to 1, the value of **num_tile_rows_minus1** is inferred to be equal to 0. Otherwise, when **uniform_tile_spacing_flag** is equal to 1, the value of **num_tile_rows_minus1** is inferred as specified in clause 6.5.1.

The variable **NumTilesInPic** is set equal to $(\text{num_tile_columns_minus1} + 1) * (\text{num_tile_rows_minus1} + 1)$.

When **single_tile_in_pic_flag** is equal to 0, **NumTilesInPic** shall be greater than 1.

tile_column_width_minus1[*i*] plus 1 specifies the width of the *i*-th tile column in units of CTBs.

tile_row_height_minus1[*i*] plus 1 specifies the height of the *i*-th tile row in units of CTBs.

brick_splitting_present_flag equal to 1 specifies that one or more tiles of pictures referring to the PPS may be divided into two or more bricks. **brick_splitting_present_flag** equal to 0 specifies that no tiles of pictures referring to the PPS are divided into two or more bricks.

num_tiles_in_pic_minus1 plus 1 specifies the number of tiles in each picture referring to the PPS. The value of **num_tiles_in_pic_minus1** shall be equal to **NumTilesInPic** - 1. When not present, the value of **num_tiles_in_pic_minus1** is inferred to be equal to **NumTilesInPic** - 1.

brick_split_flag[*i*] equal to 1 specifies that the *i*-th tile is divided into two or more bricks. **brick_split_flag**[*i*] equal to 0 specifies that the *i*-th tile is not divided into two or more bricks. When not present, the value of **brick_split_flag**[*i*] is inferred to be equal to 0. In some embodiments, PPS parsing dependency on SPS is introduced by adding the syntax condition "if(**RowHeight**[*i*] > 1)" (e.g., similarly for **uniform_brick_spacing_flag**[*i*]).

uniform_brick_spacing_flag[*i*] equal to 1 specifies that horizontal brick boundaries are distributed uniformly across the *i*-th tile and signalled using the syntax element **brick_height_minus1**[*i*]. **uniform_brick_spacing_flag**[*i*] equal to 0 specifies that horizontal brick boundaries may or may not be distributed uniformly across *i*-th tile and signalled using the syntax element **num_brick_rows_minus2**[*i*] and a list of syntax elements **brick_row_height_minus1**[*i*][*j*]. When not present, the value of **uniform_brick_spacing_flag**[*i*] is inferred to be equal to 1.

brick_height_minus1[*i*] plus 1 specifies the height of the brick rows excluding the bottom brick in the *i*-th tile in units of CTBs when **uniform_brick_spacing_flag**[*i*] is equal to 1. When present, the value of **brick_height_minus1** shall be in the range of 0 to **RowHeight**[*i*] - 2, inclusive. When not present, the value of **brick_height_minus1**[*i*] is inferred to be equal to **RowHeight**[*i*] - 1.

num_brick_rows_minus2[*i*] plus 2 specifies the number of bricks partitioning the *i*-th tile when **uniform_brick_spacing_flag**[*i*] is equal to 0. When present, the value of **num_brick_rows_minus2**[*i*] shall be in the range of 0 to **RowHeight**[*i*] - 2, inclusive. If **brick_split_flag**[*i*] is equal to 0, the value of **num_brick_rows_minus2**[*i*] is inferred to be equal to -1. Otherwise, when **uniform_brick_spacing_flag**[*i*] is equal to 1, the value of **num_brick_rows_minus2**[*i*] is inferred as specified in 6.5.1.

brick_row_height_minus1[*i*][*j*] plus 1 specifies the height of the *j*-th brick in the *i*-th tile in units of CTBs when **uniform_tile_spacing_flag** is equal to 0.

The following variables are derived, and, when `uniform_tile_spacing_flag` is equal to 1, the values of `num_tile_columns_minus1` and `num_tile_rows_minus1` are inferred, and, for each `i` ranging from 0 to `NumTilesInPic - 1`, inclusive, when `uniform_brick_spacing_flag[ i ]` is equal to 1, the value of `num_brick_rows_minus2[ i ]` is inferred, by invoking the CTB raster and brick scanning conversion process as specified in clause 6.5.1:

- the list `RowHeight[ j ]` for `j` ranging from 0 to `num_tile_rows_minus1`, inclusive, specifying the height of the `j`-th tile row in units of CTBs,
- the list `CtbAddrRsToBs[ ctbAddrRs ]` for `ctbAddrRs` ranging from 0 to `PicSizeInCtbsY - 1`, inclusive, specifying the conversion from a CTB address in the CTB raster scan of a picture to a CTB address in the brick scan,
- the list `CtbAddrBsToRs[ ctbAddrBs ]` for `ctbAddrBs` ranging from 0 to `PicSizeInCtbsY - 1`, inclusive, specifying the conversion from a CTB address in the brick scan to a CTB address in the CTB raster scan of a picture,
- the list `BrickId[ ctbAddrBs ]` for `ctbAddrBs` ranging from 0 to `PicSizeInCtbsY - 1`, inclusive, specifying the conversion from a CTB address in brick scan to a brick ID,
- the list `NumCtusInBrick[ brickIdx ]` for `brickIdx` ranging from 0 to `NumBricksInPic - 1`, inclusive, specifying the conversion from a brick index to the number of CTUs in the brick,
- the list `FirstCtbAddrBs[ brickIdx ]` for `brickIdx` ranging from 0 to `NumBricksInPic - 1`, inclusive, specifying the conversion from a brick ID to the CTB address in brick scan of the first CTB in the brick.

single_brick_per_slice_flag equal to 1 specifies that each slice that refers to this PPS includes one brick.

`single_brick_per_slice_flag` equal to 0 specifies that a slice that refers to this PPS may include more than one brick.

When not present, the value of `single_brick_per_slice_flag` is inferred to be equal to 1.

rect_slice_flag equal to 0 specifies that bricks within each slice are in raster scan order and the slice information is not signalled in PPS. `rect_slice_flag` equal to 1 specifies that bricks within each slice cover a rectangular region of the picture and the slice information is signalled in the PPS. When `brick_splitting_present_flag` is equal to 1, the value of `rect_slice_flag` shall be equal to 1. When not present, `rect_slice_flag` is inferred to be equal to 1.

num_slices_in_pic_minus1 plus 1 specifies the number of slices in each picture referring to the PPS. The value of `num_slices_in_pic_minus1` shall be in the range of 0 to `NumBricksInPic - 1`, inclusive. When not present and `single_brick_per_slice_flag` is equal to 1, the value of `num_slices_in_pic_minus1` is inferred to be equal to `NumBricksInPic - 1`.

bottom_right_brick_idx_length_minus1 plus 1 specifies the number of bits used to represent the syntax element `bottom_right_brick_idx_delta[ i ]`. The value of `bottom_right_brick_idx_length_minus1` shall be in the range of 0 to $\text{Ceil}(\text{Log2}(\text{NumBricksInPic})) - 1$, inclusive.

bottom_right_brick_idx_delta[i] when `i` is greater than 0 specifies the difference between the brick index of the brick located at the bottom-right corner of the `i`-th slice and the brick index of the bottom-right corner of the `( i - 1 )`-th slice. `bottom_right_brick_idx_delta[ 0 ]` specifies the brick index of the bottom right corner of the 0-th slice. When `single_brick_per_slice_flag` is equal to 1, the value of `bottom_right_brick_idx_delta[ i ]` is inferred to be equal to 1. The value of the `BottomRightBrickIdx[ num_slices_in_pic_minus1 ]` is inferred to be equal to

NumBricksInPic – 1. The length of the bottom_right_brick_idx_delta[i] syntax element is bottom_right_brick_idx_length_minus1 + 1 bits.

brick_idx_delta_sign_flag[i] equal to 1 indicates a positive sign for bottom_right_brick_idx_delta[i]. sign_bottom_right_brick_idx_delta[i] equal to 0 indicates a negative sign for bottom_right_brick_idx_delta[i]. It is a requirement of bitstream conformance that a slice shall include either a number of complete tiles or only a consecutive sequence of complete bricks of one tile.

The variable TopLeftBrickIdx[i], BottomRightBrickIdx[i], NumBricksInSlice[i] and BricksToSliceMap[j], which specify the brick index of the brick located at the top left corner of the i-th slice, the brick index of the brick located at the bottom right corner of the i-th slice, the number of bricks in the i-th slice and the mapping of bricks to slices, are derived as follows:

```

for( j = 0; i == 0 && j < NumBricksInPic; j++ )
    BricksToSliceMap[ j ] = -1
NumBricksInSlice[ i ] = 0
BottomRightBrickIdx[ i ] = bottom_right_brick_idx_delta[ i ] + ( ( i == 0 ) ? 0 :
    ( brick_idx_delta_sign_flag[ i ] ? BottomRightBrickIdx[ i - 1 ] :
    -BottomRightBrickIdx[ i - 1 ] ) )
for( j = BottomRightBrickIdx[ i ]; j >= 0; j-- ) {
    if( BrickColBd[ j ] <= BrickColBd[ BottomRightBrickIdx[ i ] ] &&
        BrickRowBd[ j ] <= BrickRowBd[ BottomRightBrickIdx[ i ] ] &&
        BricksToSliceMap[ j ] == -1 ) {
        TopLeftBrickIdx[ i ] = j
        NumBricksInSlice[ i ]++
        BricksToSliceMap[ j ] = i
    }
}

```

(7-43)

General slice header semantics

When present, the value of each of the slice header syntax elements slice_pic_parameter_set_id, non_reference_picture_flag, colour_plane_id, slice_pic_order_cnt_lsb, recovery_poc_cnt, no_output_of_prior_pics_flag, pic_output_flag, and slice_temporal_mvp_enabled_flag shall be the same in all slice headers of a coded picture.

The variable CuQpDeltaVal, specifying the difference between a luma quantization parameter for the coding unit containing cu_qp_delta_abs and its prediction, is set equal to 0. The variables CuQpOffsetCb, CuQpOffsetCr, and CuQpOffsetCbCr, specifying values to be used when determining the respective values of the Qp'Cb, Qp'Cr, and Qp'CbCr quantization parameters for the coding unit containing cu_chroma_qp_offset_flag, are all set equal to 0.

slice_pic_parameter_set_id specifies the value of pps_pic_parameter_set_id for the PPS in use. The value of slice_pic_parameter_set_id shall be in the range of 0 to 63, inclusive.

It is a requirement of bitstream conformance that the value of TemporalId of the current picture shall be greater than or equal to the value of TemporalId of the PPS that has pps_pic_parameter_set_id equal to slice_pic_parameter_set_id.

slice_address specifies the slice address of the slice. When not present, the value of slice_address is inferred to be equal to 0.

If rect_slice_flag is equal to 0, the following applies:

- The slice address is the brick ID as specified by Equation (7-59).
- The length of slice_address is $\text{Ceil}(\text{Log}_2(\text{NumBricksInPic}))$ bits.
- The value of slice_address shall be in the range of 0 to $\text{NumBricksInPic} - 1$, inclusive.

Otherwise (rect_slice_flag is equal to 1), the following applies:

- The slice address is the slice ID of the slice.
- The length of slice_address is $\text{signalled_slice_id_length_minus1} + 1$ bits.
- If **signalled_slice_id_flag** is equal to 0, the value of slice_address shall be in the range of 0 to $\text{num_slices_in_pic_minus1}$, inclusive. Otherwise, the value of slice_address shall be in the range of 0 to $2^{(\text{signalled_slice_id_length_minus1} + 1)} - 1$, inclusive.

It is a requirement of bitstream conformance that the following constraints apply:

- The value of slice_address shall not be equal to the value of slice_address of any other coded slice NAL unit of the same coded picture.
- When rect_slice_flag is equal to 0, the slices of a picture shall be in increasing order of their slice_address values.
- The shapes of the slices of a picture shall be such that each brick, when decoded, shall have its entire left boundary and entire top boundary consisting of a picture boundary or consisting of boundaries of previously decoded brick(s).

num_bricks_in_slice_minus1, when present, specifies the number of bricks in the slice minus 1. The value of num_bricks_in_slice_minus1 shall be in the range of 0 to $\text{NumBricksInPic} - 1$, inclusive. When rect_slice_flag is equal to 0 and single_brick_per_slice_flag is equal to 1, the value of num_bricks_in_slice_minus1 is inferred to be equal to 0. When single_brick_per_slice_flag is equal to 1, the value of num_bricks_in_slice_minus1 is inferred to be equal to 0.

The variable NumBricksInCurrSlice, which specifies the number of bricks in the current slice, and

SliceBrickIdx[i], which specifies the brick index of the i-th brick in the current slice, are derived as follows:

```

if( rect_slice_flag ) {
    sliceIdx = 0
    while( slice_address != slice_id[ sliceIdx ] )
        sliceIdx++
    NumBricksInCurrSlice = NumBricksInSlice[ sliceIdx ]
    brickIdx = TopLeftBrickIdx[ sliceIdx ]
    for( bIdx = 0; brickIdx <= BottomRightBrickIdx[ sliceIdx ]; brickIdx++ )

```

(7-92)

```

        if( BricksToSliceMap[ brickIdx ] == sliceIdx )
            SliceBrickIdx[ bIdx++ ] = brickIdx
    } else {
        NumBricksInCurrSlice = num_bricks_in_slice_minus1 + 1
        SliceBrickIdx[ 0 ] = slice_address
        for( i = 1; i < NumBricksInCurrSlice; i++ )
            SliceBrickIdx[ i ] = SliceBrickIdx[ i - 1 ] + 1
    }

```

The variables SubPicIdx, SubPicLeftBoundaryPos, SubPicTopBoundaryPos, SubPicRightBoundaryPos, and SubPicBotBoundaryPos are derived as follows:

```

SubPicIdx = CtbToSubPicIdx[ CtbAddrBsToRs[ FirstCtbAddrBs[ SliceBrickIdx[ 0 ] ] ] ]
if( subpic_treated_as_pic_flag[ SubPicIdx ] ) {
    SubPicLeftBoundaryPos = SubPicLeft[ SubPicIdx ] * ( subpic_grid_col_width_minus1 + 1 ) * 4
    SubPicRightBoundaryPos = ( SubPicLeft[ SubPicIdx ] + SubPicWidth[ SubPicIdx ] ) *
        ( subpic_grid_col_width_minus1 + 1 ) * 4
    SubPicTopBoundaryPos = SubPicTop[ SubPicIdx ] * ( subpic_grid_row_height_minus1 + 1 ) * 4
    SubPicBotBoundaryPos = ( SubPicTop[ SubPicIdx ] + SubPicHeight[ SubPicIdx ] ) *
        ( subpic_grid_row_height_minus1 + 1 ) * 4
}

```

[00102] 3. Examples of technical problems solved by disclosed embodiments

[00103] (1) There are some designs that can violate the sub-picture constrain.

[00104] A. TMVP in the affine constructed candidates may fetch a MV in the collocated picture out of the range of the current sub-picture.

[00105] B. When deriving gradients in Bi-Directional Optical Flow (BDOF) and Prediction Refinement Optical Flow (PROF), two extended rows and two extended columns of integer reference samples are required to be fetched. These reference samples may be out of the range of the current sub-picture.

[00106] C. When deriving the chroma residual scaling factor in luma mapping chroma scaling (LMCS), the accessed reconstructed luma samples may be out of the range of the current sub-picture.

[00107] D. The neighboring block may be out of the range of the current sub-picture, when deriving the luma intra prediction mode, reference samples for intra prediction, reference samples for CCLM, neighboring block availability for spatial neighboring candidates for merge/AMVP/CIIP/IBC/LMCS, quantization parameters, CABAC initialization process, ctxInc

derivation using left and above syntax elements, and `ctxInc` for the syntax element `mtt_split_cu_vertical_flag`. The representation of sub-picture may lead to sub-picture with incomplete CTUs. The CTU partitions and CU splitting process may need to consider incomplete CTUs.

[00108] (2) The signaled syntax elements related to sub-picture may be arbitrarily large, which may cause an overflow problem.

[00109] (3) The representation of sub-pictures may lead to non-rectangular sub-pictures.

[00110] (4) Currently the sub-picture and sub-picture grid is defined in units of 4 samples. And the length of syntax element is dependent on the picture height divided by 4. However, since the current `pic_width_in_luma_samples` and `pic_height_in_luma_samples` shall be an integer multiple of $\text{Max}(8, \text{MinCbSizeY})$, the sub-picture grid may need to be defined in units of 8 samples.

[00111] (5) The SPS syntax, `pic_width_max_in_luma_samples` and `pic_height_max_in_luma_samples` may need to be restricted to be no smaller than 8.

[00112] (6) Interaction between reference picture resampling/scalability and sub-picture is not considered in the current design.

[00113] (7) In temporal filtering, samples across different sub-pictures may be required.

[00114] (8) When signaling the slices, the information could be inferred without signaling in some cases.

[00115] (9) It is possible that all the defined slices cannot cover the whole picture or sub-picture.

[00116] 4. Example techniques and embodiments

[00117] The detailed listing below should be considered as examples to explain general concepts. These items should not be interpreted in a narrow way. Furthermore, these items can be combined in any manner. Hereinafter, temporal filter is used to represent filters that require samples in other pictures. $\text{Max}(x, y)$ returns the larger one of x and y . $\text{Min}(x, y)$ returns the smaller one of x and y .

1. The position (named position RB) at which a temporal MV predictor is fetched in a picture to generate affine motion candidates (e.g. a constructed affine merge candidate) must be in a required sub-picture, assuming the top-left corner coordinate of the required sub-picture is (x_{TL}, y_{TL}) and bottom-right coordinate of the required sub-picture is (x_{BR}, y_{BR}) .
 - a. In one example, the required sub-picture is the sub-picture covering the current block.

- b. In one example, if position RB with a coordinate (x, y) is out of the required sub-picture, the temporal MV predictor is treated as unavailable.
 - i. In one example, position RB is out of the required sub-picture if $x > x_{BR}$.
 - ii. In one example, position RB is out of the required sub-picture if $y > y_{BR}$.
 - iii. In one example, position RB is out of the required sub-picture if $x < x_{TL}$.
 - iv. In one example, position RB is out of the required sub-picture if $y < y_{TL}$.
 - c. In one example, position RB, if outside of the required sub-picture, a replacement of RB is utilized.
 - i. Alternatively, furthermore, the replacement position shall be in the required sub-picture.
 - d. In one example, position RB is clipped to be in the required sub-picture.
 - i. In one example, x is clipped as $x = \text{Min}(x, x_{BR})$.
 - ii. In one example, y is clipped as $y = \text{Min}(y, y_{BR})$.
 - iii. In one example, x is clipped as $x = \text{Max}(x, x_{TL})$.
 - iv. In one example, y is clipped as $y = \text{Max}(y, y_{TL})$.
 - e. In one example, the position RB may be the bottom right position inside the corresponding block of current block in the collocated picture.
 - f. The proposed method may be utilized in other coding tools which require to access motion information from a picture different than the current picture.
 - g. In one example, whether the above methods are applied (e.g., position RB must be in a required sub-picture (e.g. to do as claimed in 1.a and/or 1.b)) may depend on one or more syntax elements signaled in VPS/DPS/SPS/PPS/APS/slice header/tile group header. For example, the syntax element may be `subpic_treated_as_pic_flag[ SubPicIdx ]`, where SubPicIdx is the sub-picture index of sub-picture covering the current block.
2. The position (named position S) at which an integer sample is fetched in a reference not used in the interpolation process must be in a required sub-picture, assuming the top-left corner coordinate of the required sub-picture is (xTL, yTL) and the bottom-right coordinate of the required sub-picture is (xBR, yBR).
- a. In one example, the required sub-picture is the sub-picture covering the current block.

- b. In one example, if position S with a coordinate (x, y) is out of the required sub-picture, the reference sample is treated as unavailable.
 - i. In one example, position S is out of the required sub-picture if $x > x_{BR}$.
 - ii. In one example, position S is out of the required sub-picture if $y > y_{BR}$.
 - iii. In one example, position S is out of the required sub-picture if $x < x_{TL}$.
 - iv. In one example, position S is out of the required sub-picture if $y < y_{TL}$.
 - c. In one example, position S is clipped to be in the required sub-picture.
 - i. In one example, x is clipped as $x = \text{Min}(x, x_{BR})$.
 - ii. In one example, y is clipped as $y = \text{Min}(y, y_{BR})$.
 - iii. In one example, x is clipped as $x = \text{Max}(x, x_{TL})$.
 - iv. In one example, y is clipped as $y = \text{Max}(y, y_{TL})$.
 - d. In one example, whether position S must be in a required sub-picture (e.g. to do as claimed in 2.a and/or 2.b) may depend on one or more syntax elements signaled in VPS/DPS/SPS/PPS/APS/slice header/tile group header. For example, the syntax element may be `subpic_treated_as_pic_flag[ SubPicIdx ]`, where SubPicIdx is the sub-picture index of sub-picture covering the current block.
 - e. In one example, the fetched integer sample is used to generate gradients in BDOF and/or PORF.
3. The position (named position R) at which the reconstructed luma sample value is fetched may be in a required sub-picture, assuming the top-left corner coordinate of the required sub-picture is (x_{TL}, y_{TL}) and the bottom-right coordinate of the required sub-picture is (x_{BR}, y_{BR}).
- a. In one example, the required sub-picture is the sub-picture covering the current block.
 - b. In one example, if position R with a coordinate (x, y) is out of the required sub-picture, the reference sample is treated as unavailable.
 - i. In one example, position R is out of the required sub-picture if $x > x_{BR}$.
 - ii. In one example, position R is out of the required sub-picture if $y > y_{BR}$.
 - iii. In one example, position R is out of the required sub-picture if $x < x_{TL}$.
 - iv. In one example, position R is out of the required sub-picture if $y < y_{TL}$.

- c. In one example, position R is clipped to be in the required sub-picture.
 - i. In one example, x is clipped as $x = \text{Min}(x, x_{BR})$.
 - ii. In one example, y is clipped as $y = \text{Min}(y, y_{BR})$.
 - iii. In one example, x is clipped as $x = \text{Max}(x, x_{TL})$.
 - iv. In one example, y is clipped as $y = \text{Max}(y, y_{TL})$.
 - d. In one example, whether position R must be in a required sub-picture (e.g. to do as claimed in 4.a and/or 4.b) may depend on one or more syntax elements signaled in VPS/DPS/SPS/PPS/APS/slice header/tile group header. For example, the syntax element may be `subpic_treated_as_pic_flag[ SubPicIdx ]`, where SubPicIdx is the sub-picture index of sub-picture covering the current block.
 - e. In one example, the fetched luma sample is used to derive the scaling factor for the chroma component(s) in LMCS.
4. The position (named position N) at which the picture boundary check for BT/TT/QT splitting, BT/TT/QT depth derivation, and/or the signaling of CU split flag must be in a required sub-picture, assuming the top-left corner coordinate of the required sub-picture is (xTL, yTL) and the bottom-right coordinate of the required sub-picture is (xBR, yBR).
- a. In one example, the required sub-picture is the sub-picture covering the current block.
 - b. In one example, if position N with a coordinate (x, y) is out of the required sub-picture, the reference sample is treated as unavailable.
 - i. In one example, position N is out of the required sub-picture if $x > x_{BR}$.
 - ii. In one example, position N is out of the required sub-picture if $y > y_{BR}$.
 - iii. In one example, position N is out of the required sub-picture if $x < x_{TL}$.
 - iv. In one example, position N is out of the required sub-picture if $y < y_{TL}$.
 - c. In one example, position N is clipped to be in the required sub-picture.
 - i. In one example, x is clipped as $x = \text{Min}(x, x_{BR})$.
 - ii. In one example, y is clipped as $y = \text{Min}(y, y_{BR})$.
 - iii. In one example, x is clipped as $x = \text{Max}(x, x_{TL})$.
 - d. In one example, y is clipped as $y = \text{Max}(y, y_{TL})$. In one example, whether position N must be in a required sub-picture (e.g. to do as claimed in 5.a and/or 5.b) may depend on one or more syntax elements signaled in

VPS/DPS/SPS/PPS/APS/slice header/tile group header. For example, the syntax element may be `subpic_treated_as_pic_flag[ SubPicIdx ]`, where `SubPicIdx` is the sub-picture index of sub-picture covering the current block.

5. History -based Motion Vector Prediction (HMVP) table may be reset before decoding a new sub-picture in one picture.
 - a. In one example, the HMVP table used for IBC coding may be reset
 - b. In one example, the HMVP table used for inter coding may be reset
 - c. In one example, the HMVP table used for intra coding may be reset
6. The sub-picture syntax elements may be defined in units of `N` (such as `N = 8, 32`, and etc) samples.
 - a. In one example, the width of each element of the sub-picture identifier grid in units of `N` samples.
 - b. In one example, the height of each element of the sub-picture identifier grid in units of `N` samples.
 - c. In one example, `N` is set to the width and/or height of CTU.
7. The syntax element of picture width and picture height may be restricted to be no smaller than `K` ($K \geq 8$).
 - a. In one example, the picture width may need to be restricted to be no smaller than 8.
 - b. In one example, the picture height may need to be restricted to be no smaller than 8.
8. A conformance bitstream shall satisfy that sub-picture coding and Adaptive resolution conversion (ARC)/ Dynamic resolution conversion (DRC)/ Reference picture resampling (RPR) are disallowed to be enabled for one video unit (e.g., sequence).
 - a. In one example, signaling of enabling sub-picture coding may be under the conditions of disallowing ARC/DRC/RPR.
 - i. In one example, when sub-picture is enabled, such as **subpics_present_flag** equal to 1, **pic_width_in_luma_samples** for all pictures for which this SPS is active is equal to **max_width_in_luma_samples**.

- b. Alternatively, sub-picture coding and ARC/DRC/RPR may be both enabled for one video unit (e.g., sequence).
 - i. In one example, a conformance bitstream shall satisfy that the donwsampled sub-picture due to ARC/DRC/RPR shall still be in the form of K CTUs in width and M CTUs in height wherein K and M are both integers.
 - ii. In one example, a conformance bitstream shall satisfy that for sub-pictures not located at picture boundaries (e.g., right boundary and/or bottom boundary), the donwsampled sub-picture due to ARC/DRC/RPR shall still be in the form of K CTUs in width and M CTUs in height wherein K and M are both integers.
 - iii. In one example, CTU sizes may be adaptively changed based on the picture resolution.
 - 1) In one example, a max CTU size may be signaled in SPS. For each picture with less resolution, the CTU size may be changed accordingly based on the reduced resolution.
 - 2) In one example, CTU size may be signaled in SPS and PPS, and/or sub-picture level.
9. The syntax element `subpic_grid_col_width_minus1` and `subpic_grid_row_height_minus1` may be constrained.
- a. In one example, `subpic_grid_col_width_minus1` must be no larger (or must be smaller) than T1.
 - b. In one example, `subpic_grid_row_height_minus1` must be no larger (or must be smaller) than T2.
 - c. In one example, in a conformance bit-stream, `subpic_grid_col_width_minus1` and/or `subpic_grid_row_height_minus1` must follow the constraint such as bullet 3.a or 3.b.
 - d. In one example, T1 in 3.a and/or T2 in 3.b may depend on profiles/levels/tiers of a video coding standard.
 - e. In one example, T1 in 3.a may depend on the picture width.

- i. For example, $T1$ is equal to $\text{pic_width_max_in_luma_samples}/4$ or $\text{pic_width_max_in_luma_samples}/4+\text{Off}$. Off may be 1, 2, -1, -2, etc.
 - f. In one example, $T2$ in 3.b may depend on the picture width.
 - i. For example, $T2$ is equal to $\text{pic_height_max_in_luma_samples}/4$ or $\text{pic_height_max_in_luma_samples}/4-1+\text{Off}$. Off may be 1, 2, -1, -2, etc.
- 10. It is constrained that a boundary between two sub-pictures must be a boundary between two CTUs.
 - a. In other words, a CTU cannot be covered by more than one sub-picture.
 - b. In one example, the unit of $\text{subpic_grid_col_width_minus1}$ may be the CTU width (such as 32, 64, 128), instead of 4 as in VVC. The sub-picture grid width should be $(\text{subpic_grid_col_width_minus1}+1) * \text{CTU width}$.
 - c. In one example, the unit of $\text{subpic_grid_col_height_minus1}$ may be the CTU height (such as 32, 64, 128), instead of 4 as in VVC. The sub-picture grid height should be $(\text{subpic_grid_col_height_minus1}+1) * \text{CTU height}$.
 - d. In one example, in a conformance bit-stream, the constraint must be satisfied if the sub-picture approach is applied.
- 11. It is constrained that the shape of a sub-picture must be rectangular.
 - a. In one example, in a conformance bit-stream, the constraint must be satisfied if the sub-picture approach is applied.
 - b. Sub-picture may only contain rectangular slices. For example, in a conformance bit-stream, the constraint must be satisfied if the sub-picture approach is applied.
- 12. It is constrained that two sub-pictures cannot be overlapped.
 - a. In one example, in a conformance bit-stream, the constraint must be satisfied if the sub-picture approach is applied.
 - b. Alternatively, two sub-pictures may be overlapped with each other.
- 13. It is constrained that any position in the picture must be covered by one and only one sub-picture.
 - a. In one example, in a conformance bit-stream, the constraint must be satisfied if the sub-picture approach is applied.
 - b. Alternatively, one sample may not belong to any sub-picture.
 - c. Alternatively, one sample may belong to more than one sub-pictures.

14. It may be constrained that sub-pictures defined in a SPS mapped to every resolution presented in the same sequence should obey the location and/or size constrained mentioned above.
- a. In one example, the width and height of a sub-picture defined in the SPS mapped to a resolution presented in the same sequence, should be integer multiple times of N (such as 8, 16, 32) luma samples.
 - b. In one example, sub-pictures may be defined for certain layer and may be mapped to other layers.
 - i. For example, sub-pictures may be defined for the layer with the highest resolution in the sequence.
 - ii. For example, sub-pictures may be defined for the layer with the lowest resolution in the sequence.
 - iii. Which layer the sub-pictures are defined for may be signaled in SPS/VPS/PPS/slice header.
 - c. In one example, when sub-pictures and different resolutions are both applied, all resolutions (e.g., width or/and height) may be integer multiple of a given resolution.
 - d. In one example, the width and/or height of a sub-picture defined in the SPS may be integer multiple times (e.g., M) of the CTU size.
 - e. Alternatively, sub-pictures and different resolutions in a sequence may not be allowed simultaneously.
15. Sub-pictures may only apply to a certain layer(s)
- a. In one example, sub-pictures defined in a SPS may only apply to the layer with the highest resolution in a sequence.
 - b. In one example, sub-pictures defined in a SPS may only apply to the layer with the lowest temporal id in a sequence.
 - c. Which layer(s) that sub-pictures may be applied to may be indicated by one or multiple syntax elements in SPS/VPS/PPS.
 - d. Which layer(s) that sub-picture cannot be applied to may be indicated by one or multiple syntax elements in SPS/VPS/PPS.

16. In one example, the position and/or dimensions of a sub-picture may be signaled without using `subpic_grid_idx`.
 - a. In one example, the top-left position of a sub-picture may be signaled.
 - b. In one example, the bottom-right position of a sub-picture may be signaled.
 - c. In one example, the width of sub-picture may be signaled.
 - d. In one example, the height of a sub-picture may be signaled.
17. For temporal filter, when performing the temporal filtering of a sample, only samples within the same sub-picture that the current sample belongs to may be used. The required samples may be in the same picture that the current sample belongs to or in other pictures.
18. In one example, whether to and/or how to apply a partitioning method (such as QT, horizontal BT, vertical BT, horizontal TT, vertical TT, or not split, etc.) may depend on whether the current block (or partition) crosses one or multiple boundary of a sub-picture.
 - a. In one example, the picture boundary handling method for partitioning in VVC may also be applied when a picture boundary is replaced by a sub-picture boundary.
 - b. In one example, whether to parse a syntax element (e.g. a flag) which represents a partitioning method (such as QT, horizontal BT, vertical BT, horizontal TT, vertical TT, or not split, etc.) may depend on whether the current block (or partition) crosses one or multiple boundary of a sub-picture.
19. Instead of splitting one picture into multiple sub-pictures with independent coding of each sub-picture, it is proposed to split a picture into at least two sets of sub-regions, with the first set including multiple sub-pictures and the second set including all the remaining samples.
 - a. In one example, a sample in the second set is not in any sub-pictures.
 - b. Alternatively, furthermore, the second set may be encoded/decoded based on the information of the first set.
 - c. In one example, a default value may be utilized to mark whether a sample/MxK sub-region belonging to the second set.
 - i. In one example, the default value may be set equal to $(\text{max_subpics_minus1} + K)$ wherein K is an integer greater than 1.

- ii. The default value may be assigned to `subpic_grid_idx[i][j]` to indicate that grid belongs to the second set.
20. It is proposed that the syntax element `subpic_grid_idx[i][j]` cannot be larger than `max_subpics_minus1`.
- a. For example, it is constrained that in a conformance bit-stream, `subpic_grid_idx[i][j]` cannot be larger than `max_subpics_minus1`.
 - b. For example, the codeword to code `subpic_grid_idx[i][j]` cannot be larger than `max_subpics_minus1`.
21. It is proposed that, any integer number from 0 to `max_subpics_minus1` must be equal to at least one `subpic_grid_idx[i][j]`.
22. IBC virtual buffer may be reset before decoding a new sub-picture in one picture.
- a. In one example, all the samples in the IBC virtual buffer may be reset to -1.
23. Palette entry list may be reset before decoding a new sub-picture in one picture.
- a. In one example, `PredictorPaletteSize` may be set equal to 0 before decoding a new sub-picture in one picture.
24. Whether to signal the information of slices (e.g. number of slices and/or ranges of slices) may depend on the number of tiles and/or the number of bricks.
- a. In one example, if the number of bricks in a picture is one, `num_slices_in_pic_minus1` is not signaled and inferred to be 0.
 - b. In one example, if the number of bricks in a picture is one, the information of slices (e.g. number of slices and/or ranges of slices) may not be signaled.
 - c. In one example, if the number of bricks in a picture is one, the number of slices may be inferred to be one. And the slice covers the whole picture. In one example, if the number of bricks in a picture is one, `single_brick_per_slice_flag` is not signaled and inferred to be one.
 - i. Alternatively, if the number of bricks in a picture is one, `single_brick_per_slice_flag` must be one.
 - d. An exemplary syntax design is as below:

<code>pic_parameter_set_rbsp() {</code>	Descriptor
<code>...</code>	
<code><u>if(NumBricksInPic > 1){</u></code>	

single_brick_per_slice_flag	u(1)
if(!single_brick_per_slice_flag)	
rect_slice_flag	u(1)
if(rect_slice_flag && !single_brick_per_slice_flag) {	
num_slices_in_pic_minus1	ue(v)
bottom_right_brick_idx_length_minus1	ue(v)
for(i = 0; i < num_slices_in_pic_minus1; i++) {	
bottom_right_brick_idx_delta[i]	u(v)
brick_idx_delta_sign_flag[i]	u(1)
}	
}	
1	
loop_filter_across_bricks_enabled_flag	u(1)
if(loop_filter_across_bricks_enabled_flag)	
loop_filter_across_slices_enabled_flag	u(1)
}	
...	

25. Whether to signal slice_address may be decoupled from whether slices are signaled to be rectangles (e.g. whether rect_slice_flag is equal to 0 or 1).

a. An exemplary syntax design is as below:

if([[rect_slice_flag []] NumBricksInPic > 1)	
slice_address	u(v)

26. Whether to signal slice_address may depend on the number of slices when slices are signaled to be rectangles.

<u>if((rect_slice_flag && num_slices_in_pic_minus1 > 0) (!rect_slice_flag && NumBricksInPic > 1))</u>	
slice_address	u(v)

27. Whether to signal num_bricks_in_slice_minus1 may depend on the slice_address and/or the number of bricks in the picture.

a. An exemplary syntax design is as below:

if(!rect_slice_flag && !single_brick_per_slice_flag && <u>slice_address < NumBricksInPic - 1</u>)	
num_bricks_in_slice_minus1	ue(v)

28. Whether to signal loop_filter_across_bricks_enabled_flag may depend on the number of tiles and/or the number of bricks.

- a. In one example, loop_filter_across_bricks_enabled_flag is not signaled if the number of bricks is less than 2.
- b. An exemplary syntax design is as below:

pic_parameter_set_rbsp() {	Descriptor
...	
<u>if(NumBricksInPic > 1)</u>	
loop_filter_across_bricks_enabled_flag	u(1)
if(loop_filter_across_bricks_enabled_flag)	
loop_filter_across_slices_enabled_flag	u(1)
...	

29. It is a requirement of bitstream conformance that all the slices of a picture must cover the whole picture.
 - a. The requirement must be satisfied when slices are signaled to be rectangles (e.g. rect_slice_flag is equal to 1).
30. It is a requirement of bitstream conformance that all the slices of a sub-picture must cover the whole sub-picture.
 - a. The requirement must be satisfied when slices are signaled to be rectangles (e.g. rect_slice_flag is equal to 1).
31. It is a requirement of bitstream conformance that a slice cannot be overlapped with more than one sub-picture.
32. It is a requirement of bitstream conformance that a tile cannot be overlapped with more than one sub-picture.
33. It is a requirement of bitstream conformance that a brick cannot be overlapped with more than one sub-picture.

In the following discussion, a basic unit block (BUB) with dimensions $CW \times CH$ is a rectangle region. For example, a BUB may be a Coding Tree Block (CTB).

34. In one example, the number of sub-pictures (denoted as N) may be signaled.
 - a. It may be required on a conformance bit-stream that there are at least two sub-pictures in a picture if sub-pictures are used (e.g. subpics_present_flag is equal to 1).
 - b. Alternatively, N minus d (i.e., N-d) may be signaled, where d is an integer such as 0, 1, or 2.

- c. For example, $N-d$ may be coded with fixed length coding e.g. $u(x)$.
 - i. In one example, x may be a fixed number such as 8.
 - ii. In one example, x or $x-dx$ may be signaled before $N-d$ is signaled, where dx is an integer such as 0, 1 or 2. The signaled x may not be larger than a maximum value in a conformance bitstream.
 - iii. In one example, x may be derived on-the-fly.
 - 1) For example, x may be derived as a function of the total number (denoted as M) of BUBs in the picture. E.g., $x = \text{Ceil}(\log_2(M+d_0))+d_1$, where d_0 and d_1 are two integers, such as -2, -1, 0, 1, 2, etc.
 - 2) M may be derived as $M = \text{Ceiling}(W/CW) \times \text{Ceiling}(H/CH)$, where W and H represent the width and height of the picture, and CW and CH represent the width and height of a BUB.
 - d. For example, $N-d$ may be coded with a unary code or a truncated unary code.
 - e. In one example, the allowed maximum value of $N-d$ may be a fixed number.
 - i. Alternatively, the allowed maximum value of $N-d$ may be derived as a function of the total number (denoted as M) of BUBs in the picture. E.g., $x = \text{Ceil}(\log_2(M+d_0))+d_1$, where d_0 and d_1 are two integers, such as -2, -1, 0, 1, 2, etc.
35. In one example, a sub-picture may be signaled by indications of one or multiple its selected positions (e.g., top-left/top-right/bottom-left/bottom-right position) and/or its width and/or its height.
- a. In one example, the top-left position of a sub-picture may be signaled in the granularity of a basic unit block (BUB) with dimensions $CW \times CH$.
 - i. For example, the column index (denoted as Col) in terms of BUBs of the top-left BUB of the sub-picture may be signaled.
 - 1) For example, $Col - d$ may be signaled, where d is an integer such as 0, 1, or 2.
 - a) Alternatively, d may be equal to Col of a sub-picture previously coded, added by d_1 , where d_1 is an integer such as -1, 0, or 1.

- b) The sign of Col -d may be signaled.
 - ii. For example, the row index (denoted as Row) in term of BUBs of the top-left BUB of the sub-picture may be signaled.
 - 1) For example, Row - d may be signaled, where d is an integer such as 0, 1, or 2.
 - a) Alternatively, d may be equal to Row of a sub-picture previously coded, added by d1, where d1 is an integer such as -1, 0, or 1.
 - b) The sign of Row -d may be signaled.
 - iii. The row/column index (denoted as Row) mentioned above may be represented in the Coding Tree Block (CTB) unit, e.g., the x or y coordinate relative to the top-left position of a picture may be divided by CTB size and signaled.
 - iv. In one example, whether to signal the position of a sub-picture may depend on the sub-picture index.
 - 1) In one example, for the first sub-picture within a picture, the top-left position may be not signaled.
 - a) Alternatively, furthermore, the top-left position may be inferred, e.g., to be (0, 0).
 - 2) In one example, for the last sub-picture within a picture, the top-left position may be not signaled.
 - a) The top-left position may be inferred depending on information of sub-pictures previously signaled.
 - b. In one example, indications of the width/height/a selected position of a sub-picture may be signaled with truncated unary/truncated binary/unary/fixed length/K-th EG coding (e.g., K=0, 1, 2, 3).
 - c. In one example, the width of a sub-picture may be signaled in the granularity of a BUB with dimensions CW × CH.
 - i. For example, the number of columns of BUBs in the sub-picture (denoted as W) may be signaled.

- ii. For example, $W-d$ may be signaled, where d is an integer such as 0, 1, or 2.
 - 1) Alternatively, d may be equal to W of a sub-picture previously coded, added by $d1$, where $d1$ is an integer such as -1, 0, or 1.
 - 2) The sign of $W-d$ may be signaled.
- d. In one example, the height of a sub-picture may be signaled in the granularity of a BUB with dimensions $CW \times CH$.
 - i. For example, the number of rows of BUBs in the sub-picture (denoted as H) may be signaled.
 - ii. For example, $H-d$ may be signaled, where d is an integer such as 0, 1, or 2.
 - 1) Alternatively, d may be equal to H of a sub-picture previously coded, added by $d1$, where $d1$ is an integer such as -1, 0, or 1.
 - 2) The sign of $H-d$ may be signaled.
- e. In one example, $Col-d$ may be coded with fixed length coding e.g. $u(x)$.
 - i. In one example, x may be a fixed number such as 8.
 - ii. In one example, x or $x-dx$ may be signaled before $Col-d$ is signaled, where dx is an integer such as 0, 1 or 2. The signaled x may not be larger than a maximum value in a conformance bitstream.
 - iii. In one example, x may be derived on-the-fly.
 - 1) For example, x may be derived as a function of the total number (denoted as M) of BUB columns in the picture. E.g., $x = \text{Ceil}(\log_2(M+d0))+d1$, where $d0$ and $d1$ are two integers, such as -2, -1, 0, 1, 2, etc.
 - 2) M may be derived as $M = \text{Ceiling}(W/CW)$, where W represents the width of the picture, and CW represents the width of a BUB.
- f. In one example, $Row-d$ may be coded with fixed length coding e.g. $u(x)$.
 - i. In one example, x may be a fixed number such as 8.
 - ii. In one example, x or $x-dx$ may be signaled before $Row-d$ is signaled, where dx is an integer such as 0, 1 or 2. The signaled x may not be larger than a maximum value in a conformance bitstream.

- iii. In one example, x may be derived on-the-fly.
 - 1) For example, x may be derived as a function of the total number (denoted as M) of BUB rows in the picture. E.g., $x = \text{Ceil}(\log_2(M+d_0))+d_1$, where d_0 and d_1 are two integers, such as -2, -1, 0, 1, 2, etc.
 - 2) M may be derived as $M = \text{Ceiling}(H/CH)$, where H represents the height of the picture, and CH represents the height of a BUB.
- g. In one example, $W-d$ may be coded with fixed length coding e.g. $u(x)$.
 - i. In one example, x may be a fixed number such as 8.
 - ii. In one example, x or $x-dx$ may be signaled before $W-d$ is signaled, where dx is an integer such as 0, 1 or 2. The signaled x may not be larger than a maximum value in a conformance bitstream.
 - iii. In one example, x may be derived on-the-fly.
 - 1) For example, x may be derived as a function of the total number (denoted as M) of BUB columns in the picture. E.g., $x = \text{Ceil}(\log_2(M+d_0))+d_1$, where d_0 and d_1 are two integers, such as -2, -1, 0, 1, 2, etc.
 - 2) M may be derived as $M = \text{Ceiling}(W/CW)$, where W represents the width of the picture, and CW represents the width of a BUB.
- h. In one example, $H-d$ may be coded with fixed length coding e.g. $u(x)$.
 - i. In one example, x may be a fixed number such as 8.
 - ii. In one example, x or $x-dx$ may be signaled before $H-d$ is signaled, where dx is an integer such as 0, 1 or 2. The signaled x may not be larger than a maximum value in a conformance bitstream.
 - iii. In one example, x may be derived on-the-fly.
 - 1) For example, x may be derived as a function of the total number (denoted as M) of BUB rows in the picture. E.g., $x = \text{Ceil}(\log_2(M+d_0))+d_1$, where d_0 and d_1 are two integers, such as -2, -1, 0, 1, 2, etc.
 - 2) M may be derived as $M = \text{Ceiling}(H/CH)$, where H represents the height of the picture, and CH represents the height of a BUB.

- i. Col-d and/or Row-d may be signaled for all sub-pictures.
 - i. Alternatively, Col-d and/or Row-d may not be signaled for all sub-pictures.
 - 1) Col-d and/or Row-d may not be signaled if the number of sub-pictures are less than 2. (equal to 1).
 - 2) For example, Col-d and/or Row-d may not be signaled for the first sub-picture (e.g. with the sub-picture index (or sub-picture ID) equal to 0).
 - a) When they are not signaled, they may be inferred to be 0.
 - 3) For example, Col-d and/or Row-d may not be signaled for the last sub-picture (e.g. with the sub-picture index (or sub-picture ID) equal to NumSubPics-1).
 - a) When they are not signaled, they may be inferred depending on the positions and dimensions of sub-pictures already signaled.
 - j. W-d and/or H-d may be signaled for all sub-pictures.
 - i. Alternatively, W-d and/or H-d may not be signaled for all sub-pictures.
 - 1) W-d and/or H-d may not be signaled if the number of sub-pictures are less than 2. (equal to 1).
 - 2) For example, W-d and/or H-d may not be signaled for the last sub-picture (e.g. with the sub-picture index (or sub-picture ID) equal to NumSubPics-1).
 - a) When they are not signaled, they may be inferred depending on the positions and dimensions of sub-pictures already signaled.
 - k. In the above bullets, a BUB may be a Coding Tree Block (CTB).
36. In one example, the information of sub-pictures should be signaled after information of the CTB size (e.g. log2_ctu_size_minus5) has already been signaled.
37. subpic_treated_as_pic_flag[i] may not be signaled for each sub-pictures. Instead, one subpic_treated_as_pic_flag is signaled to control whether a sub-picture is treated as a picture for all sub-pictures.

38. `loop_filter_across_subpic_enabled_flag[ i ]` may not be signaled for each sub-pictures. Instead, one `loop_filter_across_subpic_enabled_flag` is signaled to control whether loop filters can be applied across sub-pictures for all sub-pictures.
39. `subpic_treated_as_pic_flag[ i ]` (`subpic_treated_as_pic_flag`) and/or `loop_filter_across_subpic_enabled_flag[ i ]` (`loop_filter_across_subpic_enabled_flag`) may be signaled conditionally.
- a. In one example, `subpic_treated_as_pic_flag[ i ]` and/or `loop_filter_across_subpic_enabled_flag[ i ]` may not be signaled if the number of sub-pictures are less than 2. (equal to 1).
40. RPR may be applied when sub-pictures are used.
- a. In one example, the scaling ratio in RPR may be constrained to be a limited set when sub-pictures are used, such as { 1:1, 1:2 and/or 2:1 }, or { 1:1, 1:2 and/or 2:1, 1:4 and/or 4:1 }, { 1:1, 1:2 and/or 2:1, 1:4 and/or 4:1, 1:8 and/or 8:1 }.
- b. In one example, the CTB size of a picture A and the CTB size of a picture B may be different if the resolution of picture A and picture B are different.
- c. In one example, suppose a sub-picture SA with dimensions $SAW \times SAH$ is in picture A and a sub-picture SB with dimensions $SBW \times SBH$ is in picture B, SA corresponds to SB, and the scaling ratios between picture A and picture B are R_w and R_h along the horizontal and vertical directions, then
- i. SAW/SBW or SBW/SAW should be equal to R_w .
- ii. SAH/SBH or SBH/SAH should be equal to R_h .
41. When sub-pictures are used (e.g. `sub_pics_present_flag` is true), a subpicutre index (or sub-picture ID) may be signaled in the slice header, and the slice address is interrupted as the address in a sub-picture instead of the whole picture.

[00118] 5. Embodiments

[00119] In the following embodiments, the newly added texts are bold italicized and the deleted texts are marked by "[]".

[00120] 5.1 Embodiment 1: sub-picture constraint on affine constructed merge candidates

8.5.5.6 Derivation process for constructed affine control point motion vector merging candidates

Inputs to this process are:

- a luma location (x_{Cb} , y_{Cb}) specifying the top-left sample of the current luma coding block relative to the top-left luma sample of the current picture,

- two variables cbWidth and cbHeight specifying the width and the height of the current luma coding block,
- the availability flags availableA₀, availableA₁, availableA₂, availableB₀, availableB₁, availableB₂, availableB₃,
- the sample locations (xNbA₀, yNbA₀), (xNbA₁, yNbA₁), (xNbA₂, yNbA₂), (xNbB₀, yNbB₀), (xNbB₁, yNbB₁), (xNbB₂, yNbB₂) and (xNbB₃, yNbB₃).

Output of this process are:

- the availability flag of the constructed affine control point motion vector merging candidates availableFlagConstK, with K = 1..6,
- the reference indices refIdxLXConstK, with K = 1..6, X being 0 or 1,
- the prediction list utilization flags predFlagLXConstK, with K = 1..6, X being 0 or 1,
- the affine motion model indices motionModelIdxConstK, with K = 1..6,
- the bi-prediction weight indices bcwIdxConstK, with K = 1..6,
- the constructed affine control point motion vectors cpMvLXConstK[cpIdx] with cpIdx = 0..2, K = 1..6 and X being 0 or 1.

...

The fourth (collocated bottom-right) control point motion vector cpMvLXCorner[3], reference index refIdxLXCorner[3], prediction list utilization flag predFlagLXCorner[3] and the availability flag availableFlagCorner[3] with X being 0 and 1 are derived as follows:

- The reference indices for the temporal merging candidate, refIdxLXCorner[3], with X being 0 or 1, are set equal to 0.
- The variables mvLXCol and availableFlagLXCol, with X being 0 or 1, are derived as follows:
 - If slice_temporal_mvp_enabled_flag is equal to 0, both components of mvLXCol are set equal to 0 and availableFlagLXCol is set equal to 0.
 - Otherwise (slice_temporal_mvp_enabled_flag is equal to 1), the following applies:

$$xColBr = xCb + cbWidth \quad (8-601)$$

$$yColBr = yCb + cbHeight \quad (8-602)$$

$$\underline{rightBoundaryPos = subpic_treated_as_pic_flag[SubPicIdx] ?}$$

$$\underline{SubPicRightBoundaryPos : pic_width_in_luma_samples - 1}$$

$$\underline{botBoundaryPos = subpic_treated_as_pic_flag[SubPicIdx] ?}$$

$$\underline{SubPicBotBoundaryPos : pic_height_in_luma_samples - 1}$$

- If yCb >> CtbLog2SizeY is equal to yColBr >> CtbLog2SizeY, yColBr is less than or equal to botBoundaryPos and xColBr is less than or equal to rightBoundaryPos, the following applies:
 - The variable colCb specifies the luma coding block covering the modified location given by ((xColBr >> 3) << 3, (yColBr >> 3) << 3) inside the collocated picture specified by ColPic.
 - The luma location (xColCb, yColCb) is set equal to the top-left sample of the collocated luma coding block specified by colCb relative to the top-left luma sample of the collocated picture specified by ColPic.

- The derivation process for collocated motion vectors as specified in clause 8.5.2.12 is invoked with currCb, colCb, (xColCb, yColCb), refIdxLXCorner[3] and sbFlag set equal to 0 as inputs, and the output is assigned to mvLXCol and availableFlagLXCol.
- Otherwise, both components of mvLXCol are set equal to 0 and availableFlagLXCol is set equal to 0.

...

[00121] 5.2 Embodiment 2: sub-picture constraint on affine constructed merge candidates

8.5.5.6 Derivation process for constructed affine control point motion vector merging candidates

Inputs to this process are:

- a luma location (xCb, yCb) specifying the top-left sample of the current luma coding block relative to the top-left luma sample of the current picture,
- two variables cbWidth and cbHeight specifying the width and the height of the current luma coding block,
- the availability flags availableA₀, availableA₁, availableA₂, availableB₀, availableB₁, availableB₂, availableB₃,
- the sample locations (xNbA₀, yNbA₀), (xNbA₁, yNbA₁), (xNbA₂, yNbA₂), (xNbB₀, yNbB₀), (xNbB₁, yNbB₁), (xNbB₂, yNbB₂) and (xNbB₃, yNbB₃).

Output of this process are:

- the availability flag of the constructed affine control point motion vector merging candidates availableFlagConstK, with K = 1..6,
- the reference indices refIdxLXConstK, with K = 1..6, X being 0 or 1,
- the prediction list utilization flags predFlagLXConstK, with K = 1..6, X being 0 or 1,
- the affine motion model indices motionModelIdxConstK, with K = 1..6,
- the bi-prediction weight indices bcwIdxConstK, with K = 1..6,
- the constructed affine control point motion vectors cpMvLXConstK[cpIdx] with cpIdx = 0..2, K = 1..6 and X being 0 or 1.

...

The fourth (collocated bottom-right) control point motion vector cpMvLXCorner[3], reference index refIdxLXCorner[3], prediction list utilization flag predFlagLXCorner[3] and the availability flag availableFlagCorner[3] with X being 0 and 1 are derived as follows:

- The reference indices for the temporal merging candidate, refIdxLXCorner[3], with X being 0 or 1, are set equal to 0.
- The variables mvLXCol and availableFlagLXCol, with X being 0 or 1, are derived as follows:
 - If slice_temporal_mvp_enabled_flag is equal to 0, both components of mvLXCol are set equal to 0 and availableFlagLXCol is set equal to 0.
 - Otherwise (slice_temporal_mvp_enabled_flag is equal to 1), the following applies:

$$xColBr = xCb + cbWidth \quad (8-601)$$

$$yColBr = yCb + cbHeight \quad (8-602)$$

rightBoundaryPos = subpic_treated_as_pic_flag[SubPicIdx] ?

SubPicRightBoundaryPos : pic_width_in_luma_samples - 1

$botBoundaryPos = subpic_treated_as_pic_flag[SubPicIdx] ?$

$SubPicBotBoundaryPos : pic_height_in_luma_samples - 1$

$xColBr = Min(rightBoundaryPos, xColBr)$

$yColBr = Min(botBoundaryPos, yColBr)$

- If $yCb \gg CtbLog2SizeY$ is equal to $yColBr \gg CtbLog2SizeY$, [[$yColBr$ is less than $pic_height_in_luma_samples$ and $xColBr$ is less than $pic_width_in_luma_samples$, the following applies]]:
 - The variable $colCb$ specifies the luma coding block covering the modified location given by $((xColBr \gg 3) \ll 3, (yColBr \gg 3) \ll 3)$ inside the collocated picture specified by $ColPic$.
 - The luma location $(xColCb, yColCb)$ is set equal to the top-left sample of the collocated luma coding block specified by $colCb$ relative to the top-left luma sample of the collocated picture specified by $ColPic$.
 - The derivation process for collocated motion vectors as specified in clause 8.5.2.12 is invoked with $currCb, colCb, (xColCb, yColCb), refIdxLXCorner[3]$ and $sbFlag$ set equal to 0 as inputs, and the output is assigned to $mvLXCol$ and $availableFlagLXCol$.
- Otherwise, both components of $mvLXCol$ are set equal to 0 and $availableFlagLXCol$ is set equal to 0.

...

[00122] 5.3 Embodiment 3: Fetching integer samples under the sub-picture constraint

8.5.6.3.3 Luma integer sample fetching process

Inputs to this process are:

- a luma location in full-sample units $(xInt_L, yInt_L)$,
- the luma reference sample array $refPicLX_L$.

Output of this process is a predicted luma sample value $predSampleLX_L$.

The variable $shift$ is set equal to $Max(2, 14 - BitDepth_Y)$.

The variable $picW$ is set equal to $pic_width_in_luma_samples$ and the variable $picH$ is set equal to $pic_height_in_luma_samples$.

The luma locations in full-sample units $(xInt, yInt)$ are derived as follows:

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

$xInt = Clip3(SubPicLeftBoundaryPos, SubPicRightBoundaryPos, xInt)$

$yInt = Clip3(SubPicTopBoundaryPos, SubPicBotBoundaryPos, yInt)$

- Otherwise:

$xInt = Clip3(0, picW - 1, sps_ref_wraparound_enabled_flag ?$ (8-782)

$ClipH((sps_ref_wraparound_offset_minus1 + 1) * MinCbSizeY, picW, xInt_L) : xInt_L)$

$yInt = Clip3(0, picH - 1, yInt_L)$ (8-783)

The predicted luma sample value $predSampleLX_L$ is derived as follows:

$predSampleLX_L = refPicLX_L[xInt][yInt] \ll shift3$ (8-784)

[00123] 5.4 Embodiment 4: deriving the variable invAvgLuma in chroma residual scaling of LMCS

8.7.5.3 Picture reconstruction with luma dependent chroma residual scaling process for chroma samples

Inputs to this process are:

- a chroma location (xCurr, yCurr) of the top-left chroma sample of the current chroma transform block relative to the top-left chroma sample of the current picture,
- a variable nCurrSw specifying the chroma transform block width,
- a variable nCurrSh specifying the chroma transform block height,
- a variable tuCbfChroma specifying the coded block flag of the current chroma transform block,
- an (nCurrSw)x(nCurrSh) array predSamples specifying the chroma prediction samples of the current block,
- an (nCurrSw)x(nCurrSh) array resSamples specifying the chroma residual samples of the current block,

Output of this process is a reconstructed chroma picture sample array recSamples.

The variable sizeY is set equal to Min(CtbSizeY, 64).

The reconstructed chroma picture sample recSamples is derived as follows for i = 0..nCurrSw – 1,

j = 0..nCurrSh – 1:

- ...
- Otherwise, the following applies:
 - ...
 - The variable currPic specifies the array of reconstructed luma samples in the current picture.
 - For the derivation of the variable varScale the following ordered steps apply:
 1. The variable invAvgLuma is derived as follows:
 - The array recLuma[i] with i=0..(2 * sizeY – 1) and the variable cnt are derived as follows:
 - The variable cnt is set equal to 0.
 - *The variable rightBoundaryPos and botBoundaryPos are derived as follows:*

$$\text{rightBoundaryPos} = \text{subpic_treated_as_pic_flag}[\text{SubPicIdx}] ?$$

$$\text{SubPicRightBoundaryPos} : \text{pic_width_in_luma_samples} - 1$$

$$\text{botBoundaryPos} = \text{subpic_treated_as_pic_flag}[\text{SubPicIdx}] ?$$

$$\text{SubPicBotBoundaryPos} : \text{pic_height_in_luma_samples} - 1$$
 - When availL is equal to TRUE, the array recLuma[i] with i = 0..sizeY – 1 is set equal to currPic[xCuCb – 1][Min(yCuCb + i, [[pic_height_in_luma_samples – 1]] botBoundaryPos)] with i = 0..sizeY – 1, and cnt is set equal to sizeY
 - When availT is equal to TRUE, the array recLuma[cnt + i] with i = 0..sizeY – 1 is set equal to currPic[Min(xCuCb + i, [[pic_width_in_luma_samples – 1]] rightBoundaryPos)][yCuCb – 1] with i = 0..sizeY – 1, and cnt is set equal to (cnt + sizeY)
 - The variable invAvgLuma is derived as follows:
 - If cnt is greater than 0, the following applies:

$$\text{invAvgLuma} = \text{Clip1}_Y \left(\left(\sum_{k=0}^{\text{cnt}-1} \text{recLuma}[k] + (\text{cnt} >> 1) \right) >> \text{Log2}(\text{cnt}) \right) \quad (8-1013)$$

– Otherwise (cnt is equal to 0), the following applies:

$$\text{invAvgLuma} = 1 << (\text{BitDepth}_Y - 1) \quad (8-1014)$$

[00124] 5.5 Embodiment 5: an example of defining the sub-picture element in unit of N (such as N=8 or 32) other than 4 samples

7.4.3.3 Sequence parameter set RBSP semantics

subpic_grid_col_width_minus1 plus 1 specifies the width of each element of the sub-picture identifier grid in units of $\frac{4N}{4}$ samples. The length of the syntax element is $\text{Ceil}(\text{Log2}(\text{pic_width_max_in_luma_samples} / \frac{4N}{4}))$ bits.

The variable NumSubPicGridCols is derived as follows:

$$\begin{aligned} \text{NumSubPicGridCols} = & (\text{pic_width_max_in_luma_samples} + \text{subpic_grid_col_width_minus1} * [[4+3]] \frac{N}{4+N-1}) \\ & / \\ & (\text{subpic_grid_col_width_minus1} * [[4+3]] \frac{N+N-1}{4+N-1}) \end{aligned} \quad (7-5)$$

subpic_grid_row_height_minus1 plus 1 specifies the height of each element of the sub-picture identifier grid in units of 4 samples. The length of the syntax element is $\text{Ceil}(\text{Log2}(\text{pic_height_max_in_luma_samples} / \frac{4N}{4}))$ bits.

The variable NumSubPicGridRows is derived as follows:

$$\begin{aligned} \text{NumSubPicGridRows} = & (\text{pic_height_max_in_luma_samples} + \text{subpic_grid_row_height_minus1} * \frac{4N+N-1}{4+N-1}) \\ & / \\ & (\text{subpic_grid_row_height_minus1} * [[4+3]] \frac{N+N-1}{4+N-1}) \end{aligned}$$

7.4.7.1 General slice header semantics

The variables SubPicIdx, SubPicLeftBoundaryPos, SubPicTopBoundaryPos, SubPicRightBoundaryPos, and SubPicBotBoundaryPos are derived as follows:

$$\begin{aligned} \text{SubPicIdx} &= \text{CtbToSubPicIdx}[\text{CtbAddrBsToRs}[\text{FirstCtbAddrBs}[\text{SliceBrickIdx}[0]]]] \\ \text{if}(\text{subpic_treated_as_pic_flag}[\text{SubPicIdx}]) & \{ \\ \text{SubPicLeftBoundaryPos} &= \text{SubPicLeft}[\text{SubPicIdx}] * (\text{subpic_grid_col_width_minus1} + 1) * \frac{4N}{4+N-1} \\ \text{SubPicRightBoundaryPos} &= (\text{SubPicLeft}[\text{SubPicIdx}] + \text{SubPicWidth}[\text{SubPicIdx}]) * \\ & (\text{subpic_grid_col_width_minus1} + 1) * \frac{4N}{4+N-1} \quad (7-93) \\ \text{SubPicTopBoundaryPos} &= \text{SubPicTop}[\text{SubPicIdx}] * (\text{subpic_grid_row_height_minus1} + 1) * \frac{4N}{4+N-1} \\ \text{SubPicBotBoundaryPos} &= (\text{SubPicTop}[\text{SubPicIdx}] + \text{SubPicHeight}[\text{SubPicIdx}]) * \\ & (\text{subpic_grid_row_height_minus1} + 1) * \frac{4N}{4+N-1} \\ & \} \end{aligned}$$

[00125] 5.6 Embodiment 6: restrict the picture width and the picture height to be equal or larger than 8

7.4.3.3 Sequence parameter set RBSP semantics

pic_width_max_in_luma_samples specifies the maximum width, in units of luma samples, of each decoded picture referring to the SPS. **pic_width_max_in_luma_samples** shall not be equal to 0 and shall be an integer multiple of $[[\text{MinCbSizeY}]] \text{Max}(8, \text{MinCbSizeY})$.

pic_height_max_in_luma_samples specifies the maximum height, in units of luma samples, of each decoded picture referring to the SPS. **pic_height_max_in_luma_samples** shall not be equal to 0 and shall be an integer multiple of $[[\text{MinCbSizeY}]] \text{ } \underline{\text{Max}(8, \text{MinCbSizeY})}$.

[00126] 5.7 Embodiment 7: sub-picture boundary check for BT/TT/QT splitting, BT/TT/QT depth derivation, and/or the signaling of CU split flag

6.4.2 Allowed binary split process

The variable **allowBtSplit** is derived as follows:

- ...
- Otherwise, if all of the following conditions are true, **allowBtSplit** is set equal to FALSE
 - **btSplit** is equal to **SPLIT_BT_VER**
 - **y0 + cbHeight** is greater than $[[\text{pic_height_in_luma_samples}]] \text{ } \underline{\text{subpic treated as pic flag[SubPicIdx] ? SubPicBotBoundaryPos + 1 : pic height in luma samples.}}$
- Otherwise, if all of the following conditions are true, **allowBtSplit** is set equal to FALSE
 - **btSplit** is equal to **SPLIT_BT_VER**
 - **cbHeight** is greater than **MaxTbSizeY**
 - **x0 + cbWidth** is greater than $[[\text{pic_width_in_luma_samples}]] \text{ } \underline{\text{subpic treated as pic flag[SubPicIdx] ? SubPicRightBoundaryPos + 1 : pic width in luma samples.}}$
- Otherwise, if all of the following conditions are true, **allowBtSplit** is set equal to FALSE
 - **btSplit** is equal to **SPLIT_BT_HOR**
 - **cbWidth** is greater than **MaxTbSizeY**
 - **y0 + cbHeight** is greater than $[[\text{pic_height_in_luma_samples}]] \text{ } \underline{\text{subpic treated as pic flag[SubPicIdx] ? SubPicBotBoundaryPos + 1 : pic height in luma samples.}}$
- Otherwise, if all of the following conditions are true, **allowBtSplit** is set equal to FALSE
 - **x0 + cbWidth** is greater than $[[\text{pic_width_in_luma_samples}]] \text{ } \underline{\text{subpic treated as pic flag[SubPicIdx] ? SubPicRightBoundaryPos + 1 : pic width in luma samples.}}$
 - **y0 + cbHeight** is greater than $[[\text{pic_height_in_luma_samples}]] \text{ } \underline{\text{subpic treated as pic flag[SubPicIdx] ? SubPicBotBoundaryPos + 1 : pic height in luma samples.}}$
 - **cbWidth** is greater than **minQtSize**
- Otherwise, if all of the following conditions are true, **allowBtSplit** is set equal to FALSE
 - **btSplit** is equal to **SPLIT_BT_HOR**
 - **x0 + cbWidth** is greater than $[[\text{pic_width_in_luma_samples}]] \text{ } \underline{\text{subpic treated as pic flag[SubPicIdx] ? SubPicRightBoundaryPos + 1 : pic width in luma samples.}}$
 - **y0 + cbHeight** is less than or equal to $[[\text{pic_height_in_luma_samples}]] \text{ } \underline{\text{subpic treated as pic flag[SubPicIdx] ? SubPicBotBoundaryPos + 1 : pic height in luma samples.}}$

6.4.2 Allowed ternary split process

The variable allowTtSplit is derived as follows:

- If one or more of the following conditions are true, allowTtSplit is set equal to FALSE:
 - cbSize is less than or equal to $2 * \text{MinTtSizeY}$
 - cbWidth is greater than $\text{Min}(\text{MaxTbSizeY}, \text{maxTtSize})$
 - cbHeight is greater than $\text{Min}(\text{MaxTbSizeY}, \text{maxTtSize})$
 - mttDepth is greater than or equal to maxMttDepth
 - $x0 + \text{cbWidth}$ is greater than $[[\text{pic_width_in_luma_samples}]]$ *subpic treated as pic flag SubPicIdx] ? SubPicRightBoundaryPos + 1 : pic width in luma samples*
 - $y0 + \text{cbHeight}$ is greater than $[[\text{pic_height_in_luma_samples}]]$ *subpic treated as pic flag SubPicIdx] ? SubPicBotBoundaryPos + 1 : pic height in luma samples*.
 - treeType is equal to DUAL_TREE_CHROMA and $(\text{cbWidth} / \text{SubWidthC}) * (\text{cbHeight} / \text{SubHeightC})$ is less than or equal to 32
 - treeType is equal to DUAL_TREE_CHROMA and modeType is equal to MODE_TYPE_INTRA
- Otherwise, allowTtSplit is set equal to TRUE.

7.3.8.2 Coding tree unit syntax

dual_tree_implicit_qt_split(x0, y0, cbSize, cqtDepth) {	Descriptor
...	
if(x1 < [[pic_width_in_luma_samples]] <u><i>(subpic treated as pic flag SubPicIdx] ? SubPicRightBoundaryPos + 1 : pic width in luma samples)</i></u>)	
dual_tree_implicit_qt_split(x1, y0, cbSize / 2, cqtDepth + 1)	
if(y1 < [[pic_height_in_luma_samples]] <u><i>(subpic treated as pic flag SubPicIdx] ? SubPicBotBoundaryPos + 1 : pic height in luma samples)</i></u>)	
dual_tree_implicit_qt_split(x0, y1, cbSize / 2, cqtDepth + 1)	
if(x1 < [[pic_width_in_luma_samples]] <u><i>(subpic treated as pic flag SubPicIdx] ? SubPicRightBoundaryPos + 1 : pic width in luma samples)</i></u> && y1 < [[pic_height_in_luma_samples]] <u><i>(subpic treated as pic flag SubPicIdx] ? SubPicBotBoundaryPos + 1 : pic height in luma samples)</i></u>)	
dual_tree_implicit_qt_split(x1, y1, cbSize / 2, cqtDepth + 1)	
} else {	
...	
}	
}	

7.3.8.4 Coding tree syntax

coding_tree(x0, y0, cbWidth, cbHeight, qgOnY, qgOnC, cbSubdiv, cqtDepth, mttDepth, depthOffset, partIdx, treeTypeCurr, modeTypeCurr) {	Descriptor
if((allowSplitBtVer allowSplitBtHor allowSplitTtVer allowSplitTtHor allowSplitQT) &&(x0 + cbWidth <= [[pic_width_in_luma_samples]] <u>(subpic_treated_as_pic_flag[SubPicIdx] ? SubPicRightBoundaryPos + 1 :</u> <u>pic_width_in_luma_samples)</u>) && (y0 + cbHeight <= [[pic_height_in_luma_samples]] <u>(subpic_treated_as_pic_flag[SubPicIdx] ? SubPicBotBoundaryPos +</u> <u>1: pic_height_in_luma_samples)))</u>	
split_cu_flag	ae(v)
if(cu_qp_delta_enabled_flag && qgOnY && cbSubdiv <= cu_qp_delta_subdiv) {	
...	
depthOffset += (y0 + cbHeight > [[pic_height_in_luma_samples]] <u>(subpic_treated_as_pic_flag[SubPicIdx] ? SubPicBotBoundaryPos +</u> <u>1: pic_height_in_luma_samples)) ? 1 : 0</u>	
y1 = y0 + (cbHeight / 2)	
coding_tree(x0, y0, cbWidth, cbHeight / 2, qgOnY, qgOnC, cbSubdiv + 1, cqtDepth, mttDepth + 1, depthOffset, 0, treeType, modeType)	
if(y1 < [[pic_height_in_luma_samples]] <u>(subpic_treated_as_pic_flag[SubPicIdx] ?</u> <u>SubPicBotBoundaryPos + 1: pic_height_in_luma_samples))</u>	
coding_tree(x0, y1, cbWidth, cbHeight / 2, qgOnY, qgOnC, cbSubdiv + 1, cqtDepth, mttDepth + 1, depthOffset, 1, treeType, modeType)	
...	
if(x1 < [[pic_width_in_luma_samples]] <u>(subpic_treated_as_pic_flag[SubPicIdx] ?</u> <u>SubPicRightBoundaryPos + 1: pic_width_in_luma_samples))</u>	
coding_tree(x1, y0, cbWidth / 2, cbHeight / 2, qgOnY, qgOnC, cbSubdiv + 2, cqtDepth + 1, 0, 0, 1, treeType, modeType)	
if(y1 < [[pic_height_in_luma_samples]] <u>(subpic_treated_as_pic_flag[SubPicIdx] ?</u> <u>SubPicBotBoundaryPos + 1: pic_height_in_luma_samples))</u>	
coding_tree(x0, y1, cbWidth / 2, cbHeight / 2, qgOnY, qgOnC, cbSubdiv + 2, cqtDepth + 1, 0, 0, 2, treeType, modeType)	
if(y1 < [[pic_height_in_luma_samples]] <u>(subpic_treated_as_pic_flag[SubPicIdx] ?</u> <u>SubPicBotBoundaryPos + 1: pic_height_in_luma_samples) && x1 <</u> <u>[[pic_width_in_luma_samples]] (subpic_treated_as_pic_flag[SubPicIdx] ?</u> <u>SubPicRightBoundaryPos + 1: pic_width_in_luma_samples))</u>	
coding_tree(x1, y1, cbWidth / 2, cbHeight / 2, qgOnY, qgOnC, cbSubdiv + 2, cqtDepth + 1, 0, 0, 3, treeType, modeType)	

[00127] 5.8 Embodiment 8: an example of defining the sub-pictures

	Descriptor
seq_parameter_set_rbsp() {	
sps_decoding_parameter_set_id	u(4)
...	
pic_width_max_in_luma_samples	ue(v)
pic_height_max_in_luma_samples	ue(v)
[[subpics_present_flag	u(1)
if(subpics_present_flag) {	
max_subpics_minus1	u(8)
subpic_grid_col_width_minus1	u(v)
subpic_grid_row_height_minus1	u(v)
for(i = 0; i < NumSubPicGridRows; i++)	
for(j = 0; j < NumSubPicGridCols; j++)	
subpic_grid_idx[i][j]	u(v)
for(i = 0; i <= NumSubPics; i++) {	
subpic_treated_as_pic_flag[i]	u(1)
loop_filter_across_subpic_enabled_flag[i]	u(1)
}	
}]]	
bit_depth_luma_minus8	ue(v)
...	
log2_ctu_size_minus5	u(2)
...	
<u>subpics_present_flag</u>	<u>u(1)</u>
<u>if(subpics_present_flag) {</u>	
<u>num_subpics_minus1</u>	<u>u(8)</u>
<u>for(i = 0; i <= num_subpics_minus1; i++) {</u>	
<u>subpic_ctb_addr x[i]</u>	<u>u(8)</u>
<u>subpic_ctb_addr y[i]</u>	<u>u(8)</u>
<u>subpic_ctb_width_minus1[i]</u>	<u>u(8)</u>
<u>subpic_ctb_height_minus1[i]</u>	<u>u(8)</u>
<u>subpic_treated_as_pic_flag[i]</u>	<u>u(1)</u>
<u>loop_filter_across_subpic_enabled_flag[i]</u>	<u>u(1)</u>
}	
...	

[00128] 5.9 Embodiment 9: an example of defining the sub-pictures

	Descriptor
seq_parameter_set_rbsp() {	
sps_decoding_parameter_set_id	u(4)
...	
pic_width_max_in_luma_samples	ue(v)
pic_height_max_in_luma_samples	ue(v)
[[subpics_present_flag	u(1)
if(subpics_present_flag) {	
max_subpics_minus1	u(8)
subpic_grid_col_width_minus1	u(v)
subpic_grid_row_height_minus1	u(v)
for(i = 0; i < NumSubPicGridRows; i++)	
for(j = 0; j < NumSubPicGridCols; j++)	
subpic_grid_idx[i][j]	u(v)
for(i = 0; i <= NumSubPics; i++) {	
subpic_treated_as_pic_flag[i]	u(1)
loop_filter_across_subpic_enabled_flag[i]	u(1)
}	
}]]	
bit_depth_luma_minus8	ue(v)
...	
log2_ctu_size_minus5	u(2)
...	
<u>subpics_present_flag</u>	<u>u(1)</u>
<u>if(subpics_present_flag) {</u>	
<u>num_subpics_minus1</u>	<u>ue(v)</u>
<u>for(i = 0; i <= num_subpics_minus1; i++) {</u>	
<u>subpic_ctb_addr x[i]</u>	<u>u(8)</u>
<u>subpic_ctb_addr y[i]</u>	<u>u(8)</u>
<u>subpic_ctb_width_minus1[i]</u>	<u>u(8)</u>
<u>subpic_ctb_height_minus1[i]</u>	<u>u(8)</u>
<u>subpic_treated_as_pic_flag[i]</u>	<u>u(1)</u>
<u>loop_filter_across_subpic_enabled_flag[i]</u>	<u>u(1)</u>
}	
...	

[00129] 5.10 Embodiment 10: an example of defining the sub-pictures

	Descriptor
seq_parameter_set_rbsp() {	
sps_decoding_parameter_set_id	u(4)
...	
pic_width_max_in_luma_samples	ue(v)
pic_height_max_in_luma_samples	ue(v)
[[subpics_present_flag	u(1)
if(subpics_present_flag) {	
max_subpics_minus1	u(8)
subpic_grid_col_width_minus1	u(v)
subpic_grid_row_height_minus1	u(v)
for(i = 0; i < NumSubPicGridRows; i++)	
for(j = 0; j < NumSubPicGridCols; j++)	
subpic_grid_idx[i][j]	u(v)
for(i = 0; i <= NumSubPics; i++) {	
subpic_treated_as_pic_flag[i]	u(1)
loop_filter_across_subpic_enabled_flag[i]	u(1)
}	
}]]	
...	
log2_ctu_size_minus5	u(2)
...	
<u>subpics_present_flag</u>	<u>u(1)</u>
<u>if(subpics_present_flag) {</u>	
<u>num_subpics_minus2</u>	<u>u(v)</u>
subpic_addr_x_length_minus1	<u>ue(v)</u>
subpic_addr_y_length_minus1	<u>ue(v)</u>
<u>for(i = 0; i < NumSubPics; i++) {</u>	
<u>subpic_ctb_addr x[i]</u>	<u>u(v)</u>
<u>subpic_ctb_addr y[i]</u>	<u>u(v)</u>
<u>subpic_ctb_width_minus1[i]</u>	<u>u(v)</u>
<u>subpic_ctb_height_minus1[i]</u>	<u>u(v)</u>
<u>subpic_treated_as_pic_flag[i]</u>	<u>u(1)</u>
<u>loop_filter_across_subpic_enabled_flag[i]</u>	<u>u(1)</u>
}	
...	

[00130] 5.11 Embodiment 11: an example of defining the sub-pictures

	Descriptor
seq_parameter_set_rbsp() {	
sps_decoding_parameter_set_id	u(4)
...	
pic_width_max_in_luma_samples	ue(v)
pic_height_max_in_luma_samples	ue(v)
[[subpics_present_flag	u(1)
if(subpics_present_flag) {	
max_subpics_minus1	u(8)
subpic_grid_col_width_minus1	u(v)
subpic_grid_row_height_minus1	u(v)
for(i = 0; i < NumSubPicGridRows; i++)	
for(j = 0; j < NumSubPicGridCols; j++)	
subpic_grid_idx[i][j]	u(v)
for(i = 0; i <= NumSubPics; i++) {	
subpic_treated_as_pic_flag[i]	u(1)
loop_filter_across_subpic_enabled_flag[i]	u(1)
}	
}]]	
...	
log2_ctu_size_minus5	u(2)
...	
<u>subpics_present_flag</u>	<u>u(1)</u>
<u>if(subpics_present_flag) {</u>	
<u>num_subpics_minus2</u>	<u>u(v)</u>
subpic_addr_x_length_minus1	<u>ue(v)</u>
subpic_addr_y_length_minus1	<u>ue(v)</u>
<u>for(i = 0; i < NumSubPics; i++) {</u>	
<u>if(i < NumSubPics - 1) {</u>	
<u>subpic_ctb_addr_x[i]</u>	<u>u(v)</u>
<u>subpic_ctb_addr_y[i]</u>	<u>u(v)</u>
<u>subpic_ctb_width_minus1[i]</u>	<u>u(v)</u>
<u>subpic_ctb_height_minus1[i]</u>	<u>u(v)</u>
<u>}</u>	
<u>subpic_treated_as_pic_flag[i]</u>	<u>u(1)</u>

<u>loop_filter_across_subpic_enabled_flag[i]</u>	<u>u(1)</u>
}	
...	

NumSubPics = num_subpics_minus2 + 2.

[00131] FIG. 3 is a block diagram of a video processing apparatus 300. The apparatus 300 may be used to implement one or more of the methods described herein. The apparatus 300 may be embodied in a smartphone, tablet, computer, Internet of Things (IoT) receiver, and so on. The apparatus 300 may include one or more processors 302, one or more memories 304 and video processing hardware 306. The processor(s) 302 may be configured to implement one or more methods described in the present document. The memory (memories) 304 may be used for storing data and code used for implementing the methods and techniques described herein. The video processing hardware 306 may be used to implement, in hardware circuitry, some techniques described in the present document.

[00132] FIG. 4 is a flowchart for a method 400 of processing a video. The method 400 includes determining (402), for a video block in a first video region of a video, whether a position at which a temporal motion vector predictor is determined for a conversion between the video block and a bitstream representation of the current video block using an affine mode is within a second video region, and performing (404) the conversion based on the determining.

[00133] The following solutions may be implemented as preferred solutions in some embodiments.

[00134] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 1).

[00135] 1. A method of video processing, comprising: determining, for a video block in a first video region of a video, whether a position at which a temporal motion vector predictor is determined for a conversion between the video block and a bitstream representation of the current video block using an affine mode is within a second video region; and performing the conversion based on the determining.

[00136] 2. The method of solution 1, wherein the video block is covered by the first region and the second region.

[00137] 3. The method of any of solutions 1-2, wherein, in case that the position of the temporal motion vector predictor is outside of the second video region, then the temporal motion vector predictor is marked as unavailable and is unused in the conversion.

[00138] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 2).

[00139] 4. A method of video processing, comprising: determining, for a video block in a first video region of a video, whether a position at which an integer sample in a reference picture is fetched for a conversion between the video block and a bitstream representation of the current video block is within a second video region, wherein the reference picture is not used in an interpolation process during the conversion; and performing the conversion based on the determining.

[00140] 5. The method of solution 4, wherein the video block is covered by the first region and the second region.

[00141] 6. The method of any of solutions 4-5, wherein, in case that the position of the sample is outside of the second video region, then the sample is marked as unavailable and is unused in the conversion.

[00142] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 3).

[00143] 7. A method of video processing, comprising: determining, for a video block in a first video region of a video, whether a position at which a reconstructed luma sample value is fetched for a conversion between the video block and a bitstream representation of the current video block is within a second video region; and performing the conversion based on the determining.

[00144] 8. The method of solution 7, wherein the luma sample is covered by the first region and the second region.

[00145] 9. The method of any of solutions 7-8, wherein, in case that the position of the luma sample is outside of the second video region, then the luma sample is marked as unavailable and is unused in the conversion.

[00146] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 4).

[00147] 10. A method of video processing, comprising: determining, for a video block in a first video region of a video, whether a position at which a check regarding splitting, depth derivation or split flag signaling for the video block is performed during a conversion between the video block and a bitstream representation of the current video block is within a second video region; and performing the conversion based on the determining.

[00148] 11. The method of solution 10, wherein the position is covered by the first region and the second region.

[00149] 12. The method of any of solutions 10-11, wherein, in case that the position is outside of the second video region, then the luma sample is marked as unavailable and is unused in the conversion.

[00150] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 8).

[00151] 13. A method of video processing, comprising: performing a conversion between a video comprising one or more video pictures comprising one or more video blocks, and a coded representation of the video, wherein the coded representation complies with a coding syntax requirement that the conversion is not to use sub-picture coding/decoding and a dynamic resolution conversion coding/decoding tool or a reference picture resampling tool within a video unit.

[00152] 14. The method of solution 13, wherein the video unit corresponds to a sequence of the one or more video pictures.

[00153] 15. The method of any of solutions 13-14, wherein the dynamic resolution conversion coding/decoding tool comprises an adaptive resolution conversion coding/decoding tool.

[00154] 16. The method of any of solutions 13-14, wherein the dynamic resolution conversion coding/decoding tool comprises a dynamic resolution conversion coding/decoding tool.

[00155] 17. The method of any of solutions 13-16, wherein the coded representation indicates that the video unit complies with the coding syntax requirement.

[00156] 18. The method of solution 17, wherein the coded representation indicates that the video unit uses sub-picture coding.

[00157] 19. The method of solution 17, wherein the coded representation indicates that the video unit uses the dynamic resolution conversion coding/decoding tool or the reference picture resampling tool.

[00158] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 10).

[00159] 20. The method of any of solutions 1-19, wherein the second video region comprises a video sub-picture and wherein boundaries of the second video region and another video region is also a boundary between two coding tree units.

[00160] 21. The method of any of solutions 1-19, wherein the second video region comprises a video sub-picture and wherein boundaries of the second video region and another video region is also a boundary between two coding tree units.

[00161] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 11).

[00162] 22. The method of any of solutions 1-21, wherein the first video region and the second video region have rectangular shapes.

[00163] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 12).

[00164] 23. The method of any of solutions 1-22, wherein the first video region and the second video region are non-overlapping.

[00165] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 13).

[00166] 24. The method of any of solutions 1-23, wherein the video picture is divided into video regions such that a pixel in the video picture is covered by one and only one video region.

[00167] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 15).

[00168] 25. The method of any of solutions 1-24, wherein the video picture is split into the first video region and the second video region due to the video picture being in a specific layer of the video sequence.

[00169] The following solutions may be implemented together with additional techniques described in items listed in the previous section (e.g., item 10).

[00170] 26. A method of video processing, comprising: performing a conversion between a video comprising one or more video pictures comprising one or more video blocks, and a coded representation of the video, wherein the coded representation complies with a coding syntax requirement that a first syntax element `subpic_grid_idx[i][j]` is not larger than a second syntax element `max_subpics_minus1`.

[00171] 27. The method of solution 26, wherein a codeword representing the first syntax element is not larger than a codeword representing the second syntax element.

[00172] 28. The method of any of solutions 1-27, wherein the first video region comprises a video sub-picture.

[00173] 29. The method of any of solutions 1-28, wherein the second video region comprises a video sub-picture.

[00174] 30. The method of any of solutions 1 to 29, wherein the conversion comprises encoding the video into the coded representation.

[00175] 31. The method of any of solutions 1 to 29, wherein the conversion comprises decoding the coded representation to generate pixel values of the video.

[00176] 32. A video decoding apparatus comprising a processor configured to implement a method recited in one or more of solutions 1 to 31.

[00177] 33. A video encoding apparatus comprising a processor configured to implement a method recited in one or more of solutions 1 to 31.

[00178] 34. A computer program product having computer code stored thereon, the code, when executed by a processor, causes the processor to implement a method recited in any of solutions 1 to 31.

[00179] 35. A method, apparatus or system described in the present document.

[00180] FIG. 8 is a block diagram showing an example video processing system 800 in which various techniques disclosed herein may be implemented. Various implementations may include some or all of the components of the system 800. The system 800 may include input 802 for receiving video content. The video content may be received in a raw or uncompressed format, e.g., 8 or 10 bit multi-component pixel values, or may be in a compressed or encoded format. The input 802 may represent a network interface, a peripheral bus interface, or a storage interface. Examples of network interface include wired interfaces such as Ethernet, passive optical network (PON), etc. and wireless interfaces such as Wi-Fi or cellular interfaces.

[00181] The system 800 may include a coding component 804 that may implement the various coding or encoding methods described in the present document. The coding component 804 may reduce the average bitrate of video from the input 802 to the output of the coding component 804 to produce a coded representation of the video. The coding techniques are therefore sometimes called video compression or video transcoding techniques. The output of the coding component 804 may be either stored, or transmitted via a communication connected, as represented by the component 806. The stored or communicated bitstream (or coded) representation of the video received at the input 802 may be used by the component 808 for generating pixel values or displayable video that is sent to a display interface 810. The process of generating user-viewable

video from the bitstream representation is sometimes called video decompression. Furthermore, while certain video processing operations are referred to as “coding” operations or tools, it will be appreciated that the coding tools or operations are used at an encoder and corresponding decoding tools or operations that reverse the results of the coding will be performed by a decoder.

[00182] Examples of a peripheral bus interface or a display interface may include universal serial bus (USB) or high definition multimedia interface (HDMI) or Displayport, and so on. Examples of storage interfaces include SATA (serial advanced technology attachment), PCI, IDE interface, and the like. The techniques described in the present document may be embodied in various electronic devices such as mobile phones, laptops, smartphones or other devices that are capable of performing digital data processing and/or video display.

[00183] FIG. 9 is a block diagram that illustrates an example video coding system 100 that may utilize the techniques of this disclosure.

[00184] As shown in FIG. 9, video coding system 100 may include a source device 110 and a destination device 120. Source device 110 generates encoded video data which may be referred to as a video encoding device. Destination device 120 may decode the encoded video data generated by source device 110 which may be referred to as a video decoding device.

[00185] Source device 110 may include a video source 112, a video encoder 114, and an input/output (I/O) interface 116.

[00186] Video source 112 may include a source such as a video capture device, an interface to receive video data from a video content provider, and/or a computer graphics system for generating video data, or a combination of such sources. The video data may comprise one or more pictures. Video encoder 114 encodes the video data from video source 112 to generate a bitstream. The bitstream may include a sequence of bits that form a coded representation of the video data. The bitstream may include coded pictures and associated data. The coded picture is a coded representation of a picture. The associated data may include sequence parameter sets, picture parameter sets, and other syntax structures. I/O interface 116 may include a modulator/demodulator (modem) and/or a transmitter. The encoded video data may be transmitted directly to destination device 120 via I/O interface 116 through network 130a. The encoded video data may also be stored onto a storage medium/server 130b for access by destination device 120.

[00187] Destination device 120 may include an I/O interface 126, a video decoder 124, and a

display device 122.

[00188] I/O interface 126 may include a receiver and/or a modem. I/O interface 126 may acquire encoded video data from the source device 110 or the storage medium/ server 130b. Video decoder 124 may decode the encoded video data. Display device 122 may display the decoded video data to a user. Display device 122 may be integrated with the destination device 120, or may be external to destination device 120 which be configured to interface with an external display device.

[00189] Video encoder 114 and video decoder 124 may operate according to a video compression standard, such as the High Efficiency Video Coding (HEVC) standard, Versatile Video Coding (VVC) standard and other current and/or further standards.

[00190] FIG. 10 is a block diagram illustrating an example of video encoder 200, which may be video encoder 114 in the system 100 illustrated in FIG. 9.

[00191] Video encoder 200 may be configured to perform any or all of the techniques of this disclosure. In the example of FIG. 10, video encoder 200 includes a plurality of functional components. The techniques described in this disclosure may be shared among the various components of video encoder 200. In some examples, a processor may be configured to perform any or all of the techniques described in this disclosure.

[00192] The functional components of video encoder 200 may include a partition unit 201, a predication unit 202 which may include a mode select unit 203, a motion estimation unit 204, a motion compensation unit 205 and an intra prediction unit 206, a residual generation unit 207, a transform unit 208, a quantization unit 209, an inverse quantization unit 210, an inverse transform unit 211, a reconstruction unit 212, a buffer 213, and an entropy encoding unit 214.

[00193] In other examples, video encoder 200 may include more, fewer, or different functional components. In an example, predication unit 202 may include an intra block copy (IBC) unit. The IBC unit may perform predication in an IBC mode in which at least one reference picture is a picture where the current video block is located.

[00194] Furthermore, some components, such as motion estimation unit 204 and motion compensation unit 205 may be highly integrated, but are represented in the example of FIG. 5 separately for purposes of explanation.

[00195] Partition unit 201 may partition a picture into one or more video blocks. Video encoder 200 and video decoder 300 may support various video block sizes.

[00196] Mode select unit 203 may select one of the coding modes, intra or inter, e.g., based on

error results, and provide the resulting intra- or inter-coded block to a residual generation unit 207 to generate residual block data and to a reconstruction unit 212 to reconstruct the encoded block for use as a reference picture. In some example, Mode select unit 203 may select a combination of intra and inter predication (CIIP) mode in which the predication is based on an inter predication signal and an intra predication signal. Mode select unit 203 may also select a resolution for a motion vector (e.g., a sub-pixel or integer pixel precision) for the block in the case of inter-predication.

[00197] To perform inter prediction on a current video block, motion estimation unit 204 may generate motion information for the current video block by comparing one or more reference frames from buffer 213 to the current video block. Motion compensation unit 205 may determine a predicted video block for the current video block based on the motion information and decoded samples of pictures from buffer 213 other than the picture associated with the current video block.

[00198] Motion estimation unit 204 and motion compensation unit 205 may perform different operations for a current video block, for example, depending on whether the current video block is in an I slice, a P slice, or a B slice.

[00199] In some examples, motion estimation unit 204 may perform uni-directional prediction for the current video block, and motion estimation unit 204 may search reference pictures of list 0 or list 1 for a reference video block for the current video block. Motion estimation unit 204 may then generate a reference index that indicates the reference picture in list 0 or list 1 that contains the reference video block and a motion vector that indicates a spatial displacement between the current video block and the reference video block. Motion estimation unit 204 may output the reference index, a prediction direction indicator, and the motion vector as the motion information of the current video block. Motion compensation unit 205 may generate the predicted video block of the current block based on the reference video block indicated by the motion information of the current video block.

[00200] In other examples, motion estimation unit 204 may perform bi-directional prediction for the current video block, motion estimation unit 204 may search the reference pictures in list 0 for a reference video block for the current video block and may also search the reference pictures in list 1 for another reference video block for the current video block. Motion estimation unit 204 may then generate reference indexes that indicate the reference pictures in list 0 and list 1 containing the reference video blocks and motion vectors that indicate spatial displacements

between the reference video blocks and the current video block. Motion estimation unit 204 may output the reference indexes and the motion vectors of the current video block as the motion information of the current video block. Motion compensation unit 205 may generate the predicted video block of the current video block based on the reference video blocks indicated by the motion information of the current video block.

[00201] In some examples, motion estimation unit 204 may output a full set of motion information for decoding processing of a decoder.

[00202] In some examples, motion estimation unit 204 may do not output a full set of motion information for the current video. Rather, motion estimation unit 204 may signal the motion information of the current video block with reference to the motion information of another video block. For example, motion estimation unit 204 may determine that the motion information of the current video block is sufficiently similar to the motion information of a neighboring video block.

[00203] In one example, motion estimation unit 204 may indicate, in a syntax structure associated with the current video block, a value that indicates to the video decoder 300 that the current video block has the same motion information as the another video block.

[00204] In another example, motion estimation unit 204 may identify, in a syntax structure associated with the current video block, another video block and a motion vector difference (MVD). The motion vector difference indicates a difference between the motion vector of the current video block and the motion vector of the indicated video block. The video decoder 300 may use the motion vector of the indicated video block and the motion vector difference to determine the motion vector of the current video block.

[00205] As discussed above, video encoder 200 may predictively signal the motion vector. Two examples of predictive signaling techniques that may be implemented by video encoder 200 include advanced motion vector predication (AMVP) and merge mode signaling.

[00206] Intra prediction unit 206 may perform intra prediction on the current video block. When intra prediction unit 206 performs intra prediction on the current video block, intra prediction unit 206 may generate prediction data for the current video block based on decoded samples of other video blocks in the same picture. The prediction data for the current video block may include a predicted video block and various syntax elements.

[00207] Residual generation unit 207 may generate residual data for the current video block by subtracting (e.g., indicated by the minus sign) the predicted video block(s) of the current video

block from the current video block. The residual data of the current video block may include residual video blocks that correspond to different sample components of the samples in the current video block.

[00208] In other examples, there may be no residual data for the current video block for the current video block, for example in a skip mode, and residual generation unit 207 may not perform the subtracting operation.

[00209] Transform processing unit 208 may generate one or more transform coefficient video blocks for the current video block by applying one or more transforms to a residual video block associated with the current video block.

[00210] After transform processing unit 208 generates a transform coefficient video block associated with the current video block, quantization unit 209 may quantize the transform coefficient video block associated with the current video block based on one or more quantization parameter (QP) values associated with the current video block.

[00211] Inverse quantization unit 210 and inverse transform unit 211 may apply inverse quantization and inverse transforms to the transform coefficient video block, respectively, to reconstruct a residual video block from the transform coefficient video block. Reconstruction unit 212 may add the reconstructed residual video block to corresponding samples from one or more predicted video blocks generated by the predication unit 202 to produce a reconstructed video block associated with the current block for storage in the buffer 213.

[00212] After reconstruction unit 212 reconstructs the video block, loop filtering operation may be performed reduce video blocking artifacts in the video block.

[00213] Entropy encoding unit 214 may receive data from other functional components of the video encoder 200. When entropy encoding unit 214 receives the data, entropy encoding unit 214 may perform one or more entropy encoding operations to generate entropy encoded data and output a bitstream that includes the entropy encoded data.

[00214] FIG. 11 is a block diagram illustrating an example of video decoder 300 which may be video decoder 114 in the system 100 illustrated in FIG. 9.

[00215] The video decoder 300 may be configured to perform any or all of the techniques of this disclosure. In the example of FIG. 11, the video decoder 300 includes a plurality of functional components. The techniques described in this disclosure may be shared among the various components of the video decoder 300. In some examples, a processor may be configured to

perform any or all of the techniques described in this disclosure.

[00216] In the example of FIG. 11, video decoder 300 includes an entropy decoding unit 301, a motion compensation unit 302, an intra prediction unit 303, an inverse quantization unit 304, an inverse transformation unit 305, and a reconstruction unit 306 and a buffer 307. Video decoder 300 may, in some examples, perform a decoding pass generally reciprocal to the encoding pass described with respect to video encoder 200 (e.g., FIG. 10).

[00217] Entropy decoding unit 301 may retrieve an encoded bitstream. The encoded bitstream may include entropy coded video data (e.g., encoded blocks of video data). Entropy decoding unit 301 may decode the entropy coded video data, and from the entropy decoded video data, motion compensation unit 302 may determine motion information including motion vectors, motion vector precision, reference picture list indexes, and other motion information. Motion compensation unit 302 may, for example, determine such information by performing the AMVP and merge mode.

[00218] Motion compensation unit 302 may produce motion compensated blocks, possibly performing interpolation based on interpolation filters. Identifiers for interpolation filters to be used with sub-pixel precision may be included in the syntax elements.

[00219] Motion compensation unit 302 may use interpolation filters as used by video encoder 200 during encoding of the video block to calculate interpolated values for sub-integer pixels of a reference block. Motion compensation unit 302 may determine the interpolation filters used by video encoder 200 according to received syntax information and use the interpolation filters to produce predictive blocks.

[00220] Motion compensation unit 302 may use some of the syntax information to determine sizes of blocks used to encode frame(s) and/or slice(s) of the encoded video sequence, partition information that describes how each macroblock of a picture of the encoded video sequence is partitioned, modes indicating how each partition is encoded, one or more reference frames (and reference frame lists) for each inter-encoded block, and other information to decode the encoded video sequence.

[00221] Intra prediction unit 303 may use intra prediction modes for example received in the bitstream to form a prediction block from spatially adjacent blocks. Inverse quantization unit 303 inverse quantizes, i.e., de-quantizes, the quantized video block coefficients provided in the bitstream and decoded by entropy decoding unit 301. Inverse transform unit 303 applies an inverse

transform.

[00222] Reconstruction unit 306 may sum the residual blocks with the corresponding prediction blocks generated by motion compensation unit 202 or intra-prediction unit 303 to form decoded blocks. If desired, a deblocking filter may also be applied to filter the decoded blocks in order to remove blockiness artifacts. The decoded video blocks are then stored in buffer 307, which provides reference blocks for subsequent motion compensation/intra prediction and also produces decoded video for presentation on a display device.

[00223] FIG. 12 is a flowchart representation of a method for video processing in accordance with the present technology. The method 1200 includes, at operation 1210, performing a conversion between a video comprising one or more pictures and a bitstream representation of the video. The bitstream representation is required to conform to a format rule that specifies that each picture is coded as one or more slices. The format rule prohibits samples in a picture from not being covered by any of the one or more slices.

[00224] In some embodiments, the picture comprises multiple sub-pictures, and the format rule further specifies that multiple slices of a sub-picture of the picture cover an entirety of the sub-picture. In some embodiments, the bitstream representation includes a syntax element indicating that multiple slices have a rectangular shape. In some embodiments, a slice overlaps with at most one sub-picture of the picture. In some embodiments, a tile of the video overlaps with at most one sub-picture of the picture. In some embodiments, a brick of the video overlaps with at most one sub-picture of the picture.

[00225] FIG. 13 is a flowchart representation of a method for video processing in accordance with the present technology. The method 1300 includes, at operation 1310, determining, for a conversion between a picture of a video and a bitstream representation of the video, a manner of signaling information of one or more slices in the picture according to a rule that is associated a number of tiles or a number of bricks in the picture. The method 1300 includes, at operation 1320, performing the conversion based on the determining.

[00226] In some embodiments, the information of the one or more slices of the picture comprises at least a number of the one or more slices or a range of the one or more slices. In some embodiments, the rule specifies that, in case the number of bricks in the picture is one, the information of the one or more slices in the picture is excluded in the bitstream representation. In some embodiments, the number of the one or more slices in the picture is deemed to be 1. In some

embodiments, the slice covers an entirety of the picture. In some embodiments, a number of bricks per slice is deemed to be 1.

[00227] FIG. 14 is a flowchart representation of a method for video processing in accordance with the present technology. The method 1400 includes, at operation 1410, performing a conversion between a picture of a video and a bitstream representation of the video according to a rule. The picture is coded in the bitstream representation as one or more slices. The rule specifies whether or how an address of a slice of the picture is included in the bitstream representation.

[00228] In some embodiments, the rule specifies that signaling of the address of the slice is independent from signaling of whether the slice has a rectangular shape. In some embodiments, the rule specifies that signaling of the address of the slice is dependent on a number of slices in the picture in case the slice as a rectangular shape. In some embodiments, whether to signal a number of bricks in the slice is at least in part based on the address of the slice. In some embodiments, whether to signal the number of bricks in the slice is further based on a number of bricks in the picture.

[00229] FIG. 15 is a flowchart representation of a method for video processing in accordance with the present technology. The method 1500 includes, at operation 1510, determining, for a conversion between a picture of a video and a bitstream representation of the video, whether a syntax element indicating a filter operation that accesses samples across multiple bricks in the picture is enabled is included in the bitstream representation based on a number of tiles or a number of bricks in the picture. The method 1500 includes, at operation 1520, performing the conversion based on the determining. In some embodiments, the syntax element is excluded in the bitstream representation in case the number of bricks in the picture is less than 2.

[00230] FIG. 16 is a flowchart representation of a method for video processing in accordance with the present technology. The method 1600 includes, at operation 1610, performing a conversion between a picture of a video and a bitstream representation of the video. The picture comprises one or more sub-pictures, and a number of the one or more sub-pictures is indicated by a syntax element in the bitstream representation.

[00231] In some embodiments, the number of the multiple sub-pictures is represented as N , and the syntax element includes a value of $(N-d)$, N and d being integers. In some embodiments, d is 0, 1, or 2. In some embodiments, the value of $(N-d)$ is coded using a unary coding scheme or a truncated unary coding scheme. In some embodiments, the value of $(N-d)$ is coded to have a fixed

length using a fixed length coding scheme. In some embodiments, the fixed length is 8 bits. In some embodiments, the fixed length is represented as $x-dx$, x and dx being positive integers. x is equal to or smaller than a maximum value determined based on a conformance rule, and dx is 0, 1, or 2. In some embodiments, $x-dx$ is signaled. In some embodiments, the fixed length is determined based on a number of basic unit blocks in the picture. In some embodiments, the fixed length is represented as x , the number of basic unit blocks is represented as M , and $x = \text{Ceil}(\log_2(M+d_0)) + d_1$, d_0 and d_1 being integers.

[00232] In some embodiments, a maximum value of $(N-d)$ is predetermined. In some embodiments, a maximum value of $(N-d)$ is determined based on a number of basic unit blocks in the picture. In some embodiments, the number of basic unit blocks in the picture is represented as M , M being an integer, and the maximum value of $N-d$ is $M-d$. In some embodiments, a basic unit block comprises a slice. In some embodiments, a basic unit block comprises a coding tree unit. In some embodiments, at least one of d_0 or d_1 is -2, -1, 0, 1, or 2.

[00233] FIG. 17 is a flowchart representation of a method for video processing in accordance with the present technology. The method 1700 includes, at operation 1710, performing a conversion between a picture of a video that comprises one or more sub-pictures and a bitstream representation of the video. The bitstream representation conforms to a format rule that specifies that information about a sub-picture is included in the bitstream representation based on at least one of: (1) one or more corner positions of the sub-picture, or (2) a dimension of the sub-picture.

[00234] In some embodiments, the format rule further specifies that the information about the sub-picture is positioned after information about a size of a coding tree block in the bitstream representation. In some embodiments, the one or more corner positions of the sub-picture is signaled in the bitstream representation using a granularity of a block in the sub-picture. In some embodiments, the block comprises a basic unit block or a coding tree block.

[00235] In some embodiments, an index of a block is used to indicate a corner position of the sub-picture. In some embodiments, the index comprises a column index or a row index. In some embodiments, the index of the block is represented as I , and a value of $(I - d)$ is signaled in the bitstream representation, d being 0, 1, or 2. In some embodiments, d is determined based on an index of previously coded sub-picture and another integer d_1 , d_1 being equal to -1, 0, or 1. In some embodiments, a sign of $(I-d)$ is signaled in the bitstream representation.

[00236] In some embodiments, whether a position of the sub-picture is signaled in the bitstream

representation is based on an index of the sub-picture. In some embodiments, the position of the sub-picture is omitted in the bitstream representation in case the sub-picture is a first sub-picture in the picture. In some embodiments, a top-left position of the sub-picture is determined to be (0, 0). In some embodiments, the position of the sub-picture is omitted in the bitstream representation in case the sub-picture is a last sub-picture in the picture. In some embodiments, a top-left position of the sub-picture is determined based on information about a sub-picture previously converted. In some embodiments, the information about the sub-picture comprises a top-left position and a dimension of the sub-picture. The information is coded using a truncated unary coding scheme, a truncated binary coding scheme, a unary coding scheme, a fixed length coding scheme, or a K-th EG coding scheme.

[00237] In some embodiments, a dimension of the sub-picture is signaled in the bitstream representation using a granularity of a block in the sub-picture. In some embodiments, the dimension comprises a width or a height of the sub-picture. In some embodiments, the dimension is represented as a number of columns or a number of rows of blocks in the sub-picture. In some embodiments, the number of columns or the number of rows of blocks is represented as I , and a value of $(I - d)$ is signaled in the bitstream representation, d being 0, 1, or 2. In some embodiments, d is determined based on a dimension of a block of a previously coded sub-picture and another integer d_1 , d_1 being equal to -1, 0, or 1. In some embodiments, a sign of $(I - d)$ is signaled in the bitstream representation.

[00238] In some embodiments, the value of $(I - d)$ is coded to have a fixed length using a fixed length coding scheme. In some embodiments, the fixed length is 8 bits. In some embodiments, the fixed length is represented as $x - dx$, x and dx being positive integers. x is equal to or smaller than a maximum value determined based on a conformance rule, and dx is 0, 1, or 2. In some embodiments, the fixed length is determined based on a number of basic unit blocks in the picture. In some embodiments, the fixed length is represented as x , the number of basic unit blocks is represented as M , and $x = \text{Ceil}(\log_2(M + d_0)) + d_1$, d_0 and d_1 being integers. In some embodiments, at least one of d_0 or d_1 is -2, -1, 0, 1, or 2. In some embodiments, $(I - d)$ is signaled for all sub-pictures of the picture. In some embodiments, $(I - d)$ is signaled for a subset of sub-pictures of the picture. In some embodiments, the value of $(I - d)$ is omitted in the bitstream representation in case a number of sub-pictures is equal to 1. In some embodiments, the value of $(I - d)$ is omitted in the bitstream representation in case the sub-picture is a first sub-picture of the picture. In some

embodiments, the value of (I-d) is determined to be 0. In some embodiments, the value of (I-d) is omitted in the bitstream representation in case the sub-picture is a last sub-picture of the picture. In some embodiments, the value of (I-d) is determined based on information about a sub-picture previously converted.

[00239] In some embodiments, a single syntax element is conditionally included in the bitstream representation to indicate whether all sub-pictures of the picture are processed as a picture. In some embodiments, a single syntax element is conditionally included in the bitstream representation to indicate whether a loop filter is applicable across sub-pictures of the picture for all sub-pictures. In some embodiments, the single syntax element is omitted in the bitstream representation in case a number of sub-pictures of the picture is equal to 1.

[00240] FIG. 18 is a flowchart representation of a method for video processing in accordance with the present technology. The method 1800 includes, at operation 1810, determining that a reference picture resampling tool is enabled for a conversion between a picture of a video and a bitstream representation of the video due to the picture being divided into one or more sub-pictures. The method 1800 also includes, at operation 1820, performing the conversion based on the determining.

[00241] In some embodiments, a scaling ratio used in the reference picture resampling coding tool is determined based on a set of ratios. In some embodiments, the set of ratios comprises at least one of: {1:1, 1:2}, {1:1, 2:1}, {1:1, 1:2, 1:4}, {1:1, 1:2, 4:1}, {1:1, 2:1, 1:4}, {1:1, 2:1, 4:1}, {1:1, 1:2, 1:4, 1:8}, {1:1, 1:2, 1:4, 8:1}, {1:1, 1:2, 4:1, 1:8}, {1:1, 1:2, 4:1, 8:1}, {1:1, 2:1, 1:4, 1:8}, {1:1, 2:1, 1:4, 8:1}, {1:1, 2:1, 4:1, 1:8}, or {1:1, 2:1, 4:1, 8:1}. In some embodiments, a size of a coding tree block of the picture is different than a size of a coding tree block of a second picture in case a resolution of the picture is different than a resolution of the second picture. In some embodiments, a sub-picture of the picture has a dimension of $SAW \times SAH$ and a sub-picture of a second picture of the video has a dimension of $SBW \times SBH$. Scaling ratios between the picture and the second picture are R_w and R_h along a horizontal direction and a vertical direction. SAW/SBW or SBW/SAW is equal to R_w , and SAH/SBH or SBH/SAH is equal to R_h .

[00242] FIG. 19 is a flowchart representation of a method for video processing in accordance with the present technology. The method 1900 includes, at operation 1910, performing a conversion between video comprising a video picture comprising one or more sub-pictures comprising one or more slices and a bitstream representation of the video. The bitstream representation conforms to a format rule that specifies that, for a sub-picture and a slice, in case that an index identifying the

sub-picture is included in a header of the slice, an address field for the slice indicates an address of the slice in the sub-picture.

[00243] In some embodiments, the conversion generates the video from the bitstream representation. In some embodiments, the conversion generates the bitstream representation from the video.

[00244] Some embodiments of the disclosed technology include making a decision or determination to enable a video processing tool or mode. In an example, when the video processing tool or mode is enabled, the encoder will use or implement the tool or mode in the processing of a block of video, but may not necessarily modify the resulting bitstream based on the usage of the tool or mode. That is, a conversion from the block of video to the bitstream representation of the video will use the video processing tool or mode when it is enabled based on the decision or determination. In another example, when the video processing tool or mode is enabled, the decoder will process the bitstream with the knowledge that the bitstream has been modified based on the video processing tool or mode. That is, a conversion from the bitstream representation of the video to the block of video will be performed using the video processing tool or mode that was enabled based on the decision or determination.

[00245] Some embodiments of the disclosed technology include making a decision or determination to disable a video processing tool or mode. In an example, when the video processing tool or mode is disabled, the encoder will not use the tool or mode in the conversion of the block of video to the bitstream representation of the video. In another example, when the video processing tool or mode is disabled, the decoder will process the bitstream with the knowledge that the bitstream has not been modified using the video processing tool or mode that was enabled based on the decision or determination.

[00246] 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, e.g., 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.

[00247] 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.

[00248] 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).

[00249] 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.

[00250] 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.

[00251] 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.

[00252] 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.