(54) Title: CONTROL-POINT BASED INTRA DIRECTION REPRESENTATION FOR INTRA CODING
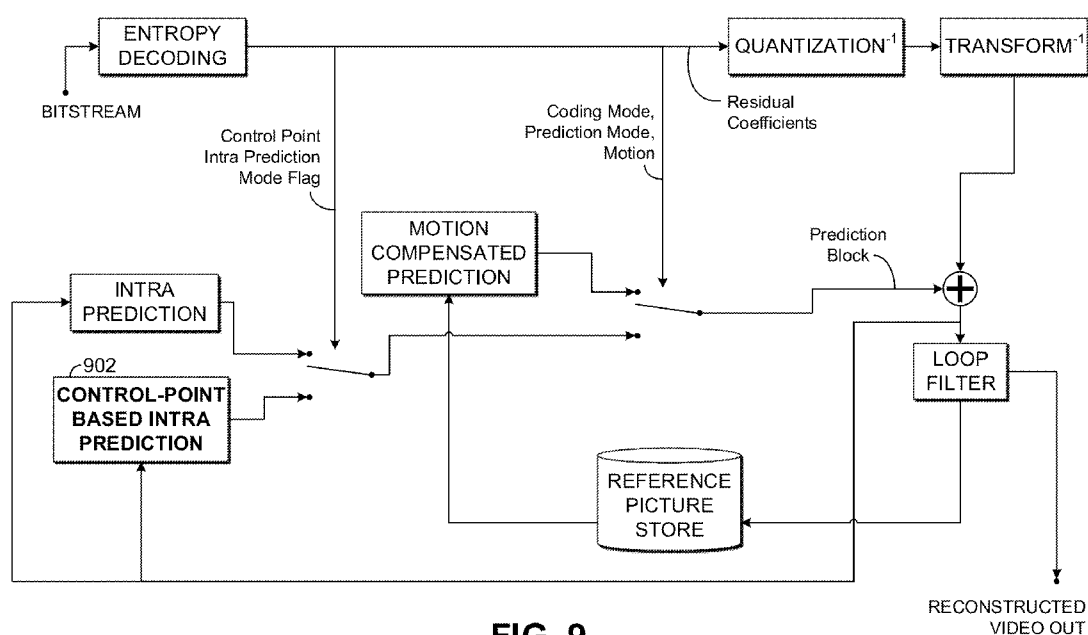
FIG. 9

(57) Abstract: Systems and methods are described for control-point based intra mode for coding a video bitstream. In an exemplary
embodiment, at least two control points in a picture are selected. The control points may be, for example, points at or adjacent to two or
more corners of a current block. For each of the control points, an associated intra prediction direction is identified. The intra prediction
directions may be encoded in the bitstream, e.g. using differential coding. A derived intra prediction direction is interpolated based on
a position of a pixel (or of a block) relative to the control points, and the derived intra prediction direction is used to predict one or more
samples in the video. Different interpolation techniques, such as triangular interpolation or bilinear interpolation may be used.

SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

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

CONTROL-POINT BASED INTRA DIRECTION REPRESENTATION FOR INTRA CODING

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001]   The present application is a non-provisional filing of, and claims benefit under 35 U.S.C. §119(c) from, U.S. Provisional Patent Application Serial No. 62/332,357, filed May 5, 2016, entitled "SYSTEMS AND METHODS FOR INTRA CODING WITH CONTROL-POINT BASED INTRA DIRECTION REPRESENTATION," which is incorporated herein by reference in its entirety.

BACKGROUND

[0002]   Video coding systems are widely used to compress digital video signals to reduce the storage requirements and/or transmission bandwidth of such signals. Among the various types of video coding systems, such as block-based, wavelet-based, and object-based systems, nowadays block-based hybrid video coding systems are the most widely used and deployed. Examples of block-based video coding systems include international video coding standards such as the MPEG1/2/4 part 2, H.264/MPEG-4 part 10 AVC, VC-1, and the latest video coding standard called High Efficiency Video Coding (HEVC), which was developed by JCT-VC (Joint Collaborative Team on Video Coding) of ITU-T/SG16/Q.6/VCEG and ISO/IEC/MPEG.

[0003]   The first version of the HEVC standard was finalized in Oct. 2013 and offers approximately 50% bit-rate saving or equivalent perceptual quality compared to the prior generation video coding standard H.264/MPEG AVC. Although the HEVC standard provides significant coding improvements over its predecessor, there is evidence that superior coding efficiency can be achieved with additional coding tools over HEVC. Recently, both VCEG and MPEG started the exploration work of new coding technologies for future video coding standardization. In Oct. 2015, ITU-T VECG and ISO/IEC MPEG formed the Joint Video Exploration Team (JVET) to begin significant study of advanced technologies that could enable substantial enhancement of coding efficiency over HEVC. In the same month, a software codebase, called Joint Exploration Model (JEM) was established for future video coding exploration work. The JEM reference software was based on HEVC Test Model (HM) that was developed by JCT-VC for HEVC. New developments in coding tools can be integrated into the JEM software and subjected to testing using JVET common test conditions (CTCs). Improved

systems and methods for coding video data to further reduce storage and bandwidth requirements would be desirable for consumers and providers of digital video systems.

SUMMARY

[0004] Systems and methods are disclosed herein for intra coding based on control-point representation of prediction directions. The proposed systems and methods may improve the efficiency of intra coding by increasing the precision of intra prediction directions without increasing the overhead of signaling block partitions and intra prediction modes. In an exemplary method, the intra prediction direction of samples inside a video block is derived from a set of controlling intra prediction directions for the current block. Additionally, a flag may be signaled to indicate whether the control-point based intra prediction is enabled or not. In case control-point based intra prediction is enabled, the intra modes of the control-points for the current block may be transmitted to the decoder, and the intra modes of the samples inside the block are derived from the control-point intra directions. In cases where the texture of the samples in the current block changes direction, control-point based intra prediction can provide a good estimate of the intra prediction directions for the samples inside the block. Systems and methods of control-point based intra prediction can make the use of relatively large intra coding units (CUs) and prediction units (PUs), therefore reducing the overhead used for signaling CU/PU partitions and intra direction information.

[0005] Exemplary methods are provided for encoding and/or decoding (collectively "coding") a video in a bitstream, where the video comprises a plurality of frames, each frame comprising blocks of samples. In one exemplary method, for at least a current block, at least a first intra direction is identified for a first control point and a second intra direction is identified for a second control point. For at least one current sample in the current block, an intra prediction direction is derived for the sample by interpolation based on the first intra direction and the second intra direction, and the sample is predicted with intra prediction using the derived intra prediction direction. The interpolation may be performed using bilinear interpolation, triangular interpolation, or other interpolation technique. In some embodiments, the interpolation of the intra prediction direction is further based on a third intra direction for a third control point. In some embodiments, positions of at least the first control point and the second control point are signaled in the bitstream. In some embodiments, at least the first intra prediction direction is signaled in the bitstream.

[0006]   At least the first and second control points may be selected from the group consisting of: a first control point at the top-left corner of the block, a second control point at the top-right corner of the block, a third control point at the bottom-left corner of the block, a fourth control point at the bottom-right corner of the block, a custom control point signaled in the bitstream. In some embodiments, control points at one or more of the corners are shifted by one or more pixels to simplify division operations.

[0007]   In some embodiments, the first control point is at a top-left corner of the current block and the second control point is at the top-right corner or bottom-left corner of the current block. The position of the second control point may be determined based on intra modes in a list of most probable modes for the current block. The second control point is selected to be at a top-right corner of the current block if more of the most probable modes are horizontal than vertical, and the second control point is selected to be at the bottom-left corner of the current block if more of the most probable modes are vertical than horizontal.

[0008]   In some embodiments, the derived intra prediction direction is used for all samples in a sub-block. For example, where the current sample is in a current sub-block of the current block, all samples of the current sub-block may be predicted using the derived intra prediction direction. Different intra prediction directions may be derived for different sub-blocks of the current block.

[0009]   In some embodiments, the first and second intra directions are selected from a list of most probable modes, the list being determined based on intra modes used to code neighboring blocks.

[0010]   In some embodiments, the intra direction of at least one of the control points for the current block is included in a list of most probable modes for intra coding of a neighboring block.

[0011]   In some embodiments, at least the second intra direction is coded in the bitstream using differential coding with respect to the first intra direction.

[0012]   In some embodiments, at least the first intra direction is derived from samples in a template region adjacent to the control point.

[0013]   In some embodiments, the use of control-point based intra coding for the current block is signaled by a flag in the bitstream at the block level.

[0014]   Additional embodiments describe encoders and decoders operative to perform control-point based intra coding.

[0015]   Systems and methods herein provide novel techniques for prediction of pixel values. Such techniques can be used by both encoders and decoders. Prediction of a pixel results in a pixel value that, in an encoding method, can be subtracted from an original pixel input value to determine a residual that is encoded in the bitstream. In a decoding method, a residual can be

decoded from the bitstream and added to the predicted pixel value to obtain a reconstructed pixel that is identical to or approximates the original input pixel. Prediction methods as described herein thus improve the operation of video encoders and decoders by decreasing, in at least some implementations, the number of bits required to encode and decode video. Further benefits of exemplary prediction methods to the operation of video encoders and decoders are described in the Detailed Description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016]   A more detailed understanding may be had from the following description, presented by way of example in conjunction with the accompanying drawings, which are first briefly described below.

[0017]   FIG. 1 is a functional block diagram illustrating an example of a block-based video encoder.

[0018]   FIG. 2 is a functional block diagram illustrating an example of a block-based video decoder.

[0019]   FIG. 3 is a diagram illustrating the notation used in the present disclosure to describe reference samples in intra prediction.

[0020]   FIG. 4 is a diagram illustrating an example of 33 directional intra prediction modes and two non-directional intra prediction modes in HEVC.

[0021]   FIG. 5 is a diagram illustrating prediction directions for intra mode in JEM. The solid lines correspond to the HEVC intra mode directions, which are a subset of the JEM modes. The numbers are the mode numbers for HEVC.

[0022]   FIG. 6 is a schematic representation of a position-dependent intra prediction combination.

[0023]   FIGs. 7A-7D are schematic illustrations comparing different intra mode partition methods. FIG. 7A: original block; FIG. 7B: one CU partition; FIG. 7C: four CU partitions; FIG. 7D: CU + PU partition.

[0024]   FIG. 8 is a block diagram of a hybrid video encoder with the inclusion of a control-point based intra prediction module.

[0025]   FIG. 9 is a block diagram of a hybrid video decoder with the inclusion of a control-point based intra prediction module.

[0026]   FIG. 10 is a schematic illustration of an exemplary control-point representation of intra prediction directions.

[0027]   FIGs. 11A-11D are schematic illustration of selections of control points in embodiments in which two controlling intra prediction directions are applied. FIG. 11A: horizontal control-point selection; FIG. 11B: vertical control-point selection; FIG. 11C: diagonal control-point selection; FIG. 11D: anti-diagonal control-point selection.

[0028]   FIGs. 12A and 12B are schematic illustrations of derivation of intra prediction directions in embodiments in which three control-points are applied. FIG. 12A: derivation at the PU-level; FIG. 12B: derivation at the TU-level.

[0029]   FIG. 13 is a schematic illustration of decoder-side derivation of control intra prediction directions in an exemplary embodiment.

[0030]   FIG. 14 is a schematic illustration of generation of intra prediction samples using a weighted combination of intra prediction signals using the intra directions of control-points.

[0031]   FIG. 15 is a diagram illustrating an example of a coded bitstream structure.

[0032]   FIG. 16 is a diagram illustrating an example communication system.

[0033]   FIG. 17 is a diagram illustrating an example wireless transmit/receive unit (WTRU), which may be used as an encoder or decoder in some embodiments.

DETAILED DESCRIPTION

Exemplary Block-Based Encoding and Decoding Systems and Methods.

[0034]   Embodiments disclosed herein, like the HEVC and JEM software, are built upon a block-based hybrid video coding framework.

[0035]   FIG. 1 is a functional block diagram of a block-based hybrid video encoding system. The input video signal 102 is processed block by block. In HEVC, extended block sizes (called a "coding unit" or CU) are used to efficiently compress high resolution (1080p and beyond) video signals. In HEVC, a CU can be up to 64x64 pixels, and bigger block size up to 256x256 for the encoding of large resolution videos (e.g. 720p and beyond) is allowed in JEM. A CU can be further partitioned into prediction units (PU), for which separate prediction methods are applied. For each input video block (MB or CU), spatial prediction (160) and/or temporal prediction (162) may be performed. Spatial prediction (or "intra prediction") uses pixels from the already coded neighboring blocks in the same video picture/slice to predict the current video block. Spatial

prediction reduces spatial redundancy inherent in the video signal. Temporal prediction (also referred to as "inter prediction" or "motion compensated prediction") uses reconstructed pixels from the already coded video pictures to predict the current video block. Temporal prediction reduces temporal redundancy inherent in the video signal. A temporal prediction signal for a given video block is usually signaled by one or more motion vectors which indicate the amount and the direction of motion between the current block and its reference block. Also, if multiple reference pictures are supported (as is the case for the recent video coding standards such as H.264/AVC or HEVC), then for each video block, its reference picture index is sent additionally; and the reference index is used to identify from which reference picture in the reference picture store (164) the temporal prediction signal comes. After spatial and/or temporal prediction, the mode decision block (180) in the encoder chooses the best prediction mode, for example based on the rate-distortion optimization method. The prediction block is then subtracted from the current video block (116); and the prediction residual is de-correlated using transform (104) and quantized (106) to achieve the target bit-rate. The quantized residual coefficients are inverse quantized (110) and inverse transformed (112) to form the reconstructed residual, which is then added back to the prediction block (126) to form the reconstructed video block. Further in-loop filtering such as de-blocking filter and Adaptive Loop Filters may be applied (166) on the reconstructed video block before it is put in the reference picture store (164) and used to code future video blocks. To form the output video bit-stream 120, coding mode (inter or intra), prediction mode information, motion information, and quantized residual coefficients are all sent to the entropy coding unit (108) to be further compressed and packed to form the bit-stream.

[0036]   FIG. 2 is a functional block diagram of a block-based video decoder. The video bit-stream 202 is first unpacked and entropy decoded at entropy decoding unit 208. The coding mode and prediction information are sent to either the spatial prediction unit 260 (if intra coded) or the temporal prediction unit 262 (if inter coded) to form the prediction block. The residual transform coefficients are sent to inverse quantization unit 210 and inverse transform unit 212 to reconstruct the residual block. The prediction block and the residual block are then added together at 226. The reconstructed block may further go through in-loop filtering before it is stored in reference picture store 264. The reconstructed video is then sent out to drive a display device, as well as used to predict future video blocks.

Intra Coding Tools in HEVC.

[0037] HEVC incorporates several coding tools that improve the efficiency of intra prediction over preceding video coding standards. Those intra coding tools can more accurately model different directional texture as well as smooth regions with gradually changing sample values. Moreover, additional efforts were also made during the design to avoid the introduction of artificial edges with potential blocking artifacts by doing adaptive smoothing of reference samples and the boundary samples of the generated intra prediction for DC and directly horizontal and vertical directions.

[0038] The coding structure utilized for intra coding of HEVC closely follows the overall structure of the codec. Specifically, one picture is split into a number of CUs which represent the basic quad-tree splitting regions that are used to conduct intra coding. Inside one CU, multiple non-overlapping prediction units (PUs) can be defined, each of which specifies a region with individual intra prediction mode. In HEVC, one intra-coded CU may have two types of PU partitions, namely, PART_2Nx2N and PART_NxN. PART_2Nx2N indicates that all the samples in the CU are predicted by the same intra prediction mode and PART_NxN indicates that the CU is split into four equal size PUs each of which is predicted by its own intra prediction mode. A given PU can be further split into quad-tree of transform units (TUs), each of which is the basic operational unit of applying residual coding with a transform with the size of the TU. Additionally, in the HEVC specification, the prediction operations of all intra prediction modes in HEVC are defined on the TU basis to minimize the distance between predicted samples and reference samples used in prediction for improved efficiency of intra prediction.

[0039] Due to the number of possible intra prediction modes, HEVC considers five Most Probable Modes (MPMs) when coding luma intra prediction modes. Among the five MPMs, the first two are derived by detecting the intra prediction modes of the above and left neighboring blocks and the remaining three MPMs are defined as Planar, DC and vertical direction (26). Additionally, in cases where both above and left blocks have the same angular mode, two neighboring angular directions closest to that direction will also be included in the MPM list.

[0040] *Intra prediction modes.* For all intra prediction modes in HEVC, the same set of previously decoded boundary samples from above and to the left of the current TU (as shown in FIG. 3) are used for prediction. Directional prediction with 33 different angular orientations are defined for various TU sizes ranging from 4x4 to 32x32. In addition to angular directions, planar prediction (which assumes a gradual changing surface with horizontal and vertical slope derived

from boundaries) and DC prediction (which assumes a flat surface) can also be used. Table 1 specifies the mode index and intra prediction modes associated with each mode index. FIG. 4 illustrates the existing intra prediction modes as defined in HEVC. For chroma, the horizontal, vertical, planar and DC modes can be explicitly signaled in bitstream; additionally, chroma prediction mode can be indicated to be the same as luma prediction mode through DM mode.

Table 1. Specification of intra prediction modes.

Index	Intra Prediction Mode
0	Planar
1	DC
2…34	Angular(N), N = 2, 3, …,34

[0041]   *Reference sample smoothing.* In HEVC, reference samples used for intra prediction are sometimes filtered by a 3-tap smoothing filter [1, 2, 1]/4, depending on TU size and the directionality of intra prediction. For 32x32 TUs, all angular modes except for direct horizontal and vertical use filtered reference samples. For 16x16 TUs, the intra modes that do not use filtered reference samples are extended to four modes (9, 11, 25 and 27) closest to horizontal and vertical. For 8x8 and smaller TUs, only diagonal modes (2, 18 and 34) use filtered reference for prediction. Additionally, reference smoothing is also applied to planar mode when the TU size is 8x8 or larger.

[0042]   *Intra coding tools in JEM.* In general, the basic structure of HEVC intra coding remains unchanged in JEM with several modules associated with intra predictions being further extended and/or improved, including extended intra prediction directions, 4-tap intra interpolation, extended boundary smoothing, position-dependent intra prediction combination, mode-dependent non-separable secondary transform and adaptive reference filtering. In the following, several intra coding tools in JEM are briefly reviewed.

[0043]   *Extended intra prediction directions.* To capture finer edge directions presented in natural videos, in JEM, the number of the angular intra prediction modes is extended from 33 in HEVC to 65. FIG. 5 illustrates the intra prediction modes as defined in JEM where the new directional modes over HEVC are highlighted as dotted lines and the Planar and DC modes remain the same. Like HEVC, the extended directional intra prediction modes apply for all TU sizes and both luma and chroma intra predictions.

[0044]   To accommodate the increased number of directional intra modes, an improved intra mode coding method uses six MPMs. To derive the set of six MPMs, the intra modes of the left and above neighboring blocks are checked. Instead of using the intra modes from top and left

neighboring blocks directly as in HEVC, the most frequently used intra mode along the top neighboring row and along the left neighboring column are computed, and then used as the left and above neighboring modes, respectively. Next, to derive the actual 6 MPMs, the same rule used in HEVC is firstly applied to obtain 3 MPMs; then, the remaining 3 MPMs are defined as the 3 intra modes closest to (in terms of prediction direction) the angular modes which have already been included in the list of MPMs. Table 2 specifies the MPM derivation method currently used in JEM, where L and A specify the most frequently used intra prediction mode of the left and above neighboring blocks, respectively, and Max and Min specify the maximum and the minimum of the intra prediction modes of the left and the above neighboring blocks, respectively.

Table 2. Derivation of 6 MPMs in JEM.

Conditions			MPM0	MPM1	MPM2	MPM3	MPM4	MPM5
L=A	L≠Planar and L≠DC		L	Planar	L+1	L-1	L+2	DC
	Otherwise		Planar	DC	26 (Ver)	10 (Hor)	2	18
L≠A	L≠Planar and R≠Planar	L=DC or A=DC	L	A	Planar	Max-1	Max+1	Max+2
		otherwise	L	A	Planar	DC	Max+1	Min-1
	otherwise	L+A<2	L	A	26 (Ver)	10 (Hor)	2	18
		otherwise	L	A	DC	Max-1	Max+1	Max+2

[0045]   *Reference sample adaptive filtering.* In HEVC intra prediction, a low-pass filter [1, 2, 1]/4 is used to smooth reference samples for some intra prediction modes. The number of the intra prediction modes to which smooth filtering is applied depends on the block size and prediction mode. Specifically, for 32×32 blocks, all angular modes except horizontal and vertical use a filtered reference; for 16 × 16 blocks, the modes not using a filtered reference are extended to the four modes (9, 11, 25, 27) closest to horizontal and vertical; for 8×8 blocks, only the diagonal modes (2, 18, 34) use a filtered reference. Smoothing is also applied where the planar mode is used, for block sizes 8×8 and larger.

[0046]   Applying reference sample smoothing selectively based on the block size and directionality of the prediction may operate to reduce contouring artifacts caused by edges in the reference sample arrays. In JEM, in order to improve intra prediction efficiency, two sets of low-pass filters as below are applied by reference sample adaptive filtering (RASF) to smooth reference samples for intra prediction.

- 3-tap LPF with the coefficients of [1, 2, 1]/4, (default reference smoothing filter of HEVC),
- 5-tap LPF with the coefficients of [2, 3, 6, 3, 2]/16.

[0047] In order to maximize the gain that RSAF could achieve, the selection of two reference smoothing filters are performed on TU-level.

[0048] Data hiding is a technique mainly used for watermarking. In HEVC, this method is applied to hide the signs of the most significant coefficients of TU's coefficients groups (CGs) into the parity of absolute values of CG's coefficients, which is referred to as sign bits hiding (SBH). To reduce the overhead of RSAF flag signaling, data hiding is used to hide the flag that indicates whether the reference filtering is enabled for a given TU into the transform coefficients of that TU. In contrast to the SBH, which uses all the coefficient in a given TU, RSAF only uses those transform coefficients located at an odd positions within the TU to hide a value of the filtering flag.

[0049] *Position dependent intra prediction combination.* In HEVC, intra prediction is computed from either a nonfiltered or a filtered set of reference border pixels, depending on the predictor mode and block size. The position dependent intra prediction combination (PDPC) extends this method by employing weighted combinations of filtered references and prediction samples calculated from unfiltered references, where the weights depend on predictor mode and pixel position. Based on the notation used to define the intra predictors as shown in FIG. 6, r and s represent the unfiltered and filtered reference samples, respectively. $q[x, y]$ is the directional prediction derived from the filtered reference s as defined in HEVC but disabling boundary filters for DC, and direct horizontal and vertical intra directions. The values x and y are the horizontal and vertical distance from the block boundary. Using "$\gg$" to represent right bit-shift operation, the new prediction $p[x, y]$, which combines weighted values of unfiltered reference samples $r[x, y]$ with the prediction signal $q[x, y]$, is derived as

$$p[x,y] = \left\{ \left(c_1^{(v)} \gg \lfloor y/d \rfloor \right) r[x, -1] - \left(c_2^{(v)} \gg \lfloor y/d \rfloor \right) r[-1, -1] + \left(c_1^{(h)} \gg \lfloor x/d \rfloor \right) r[-1, y] - \left(c_2^{(h)} \gg \lfloor x/d \rfloor \right) r[-1, -1] + b[x, y] q[x, y] + 64 \right\} \gg 7 \tag{1}$$

where $c_1^v, c_2^v, c_1^h, c_2^h$ are predefined prediction parameters, $d = 1$ for block sizes up to 16×16, and $d = 2$ for larger blocks, and

$$b[x, y] = 128 - \left(c_1^{(v)} \gg \lfloor y/d \rfloor \right) + \left(c_2^{(v)} \gg \lfloor y/d \rfloor \right) - \left(c_1^{(v)} \gg \lfloor y/d \rfloor \right) + \left(c_2^{(h)} \gg \lfloor y/d \rfloor \right) \tag{2}$$

is a normalization factor.

[0050] The filtered reference is derived by using binomial lowpass filters, since they can be easily adjusted for integer computations, and approximate Gaussian filters. The prediction

parameters $(c_1^v, c_2^v, c_1^h, c_2^h)$ and lowpass filter parameters are defined per prediction direction (neighboring prediction directions are grouped into 35 pairs) and block size. For each of 35 prediction direction pairs for a given block, 3 sets of prediction parameters and lowpass filters are predefined. The selected one from these 3 sets is signaled on the CU level.

[0051]   *Mode-dependent non-separable secondary transform.* Due to the fact that the correlation between spatial neighboring samples is much less than that in temporal domain, there are usually strong correlations within the residual samples generated from intra prediction. In order to further improve intra coding efficiency, mode-dependent non-separable secondary transform (NSST) is applied in the current JEM. NSST applies non-separable transforms to each 4x4 transform coefficient block of intra prediction. Specifically, a CU-level index, valued from 0 to 3, is explicitly signaled to identify which pre-defined secondary transform is applied. If the index is zero, secondary transform is not used; otherwise, one of the three pre-defined 4x4 non-separable transforms is used. For each 4x4 coefficient block of one TU, the 4x4 non-separable transform identified by the CU-level index is applied. Additionally, the CU-level NSST flag is signaled only when there is at least one non-zero transform coefficient in the current CU.

[0052]   To apply the non-separable transform, one 4x4 input block X

$$X = \begin{bmatrix} X_{00} & X_{01} & X_{02} & X_{03} \\ X_{10} & X_{11} & X_{12} & X_{13} \\ X_{20} & X_{21} & X_{22} & X_{23} \\ X_{30} & X_{31} & X_{32} & X_{33} \end{bmatrix} \tag{3}$$

is first rearranged into a vector $\vec{X}$:

$$\vec{X} =$$

$$[X_{00} \quad X_{01} \quad X_{02} \quad X_{03} \quad X_{10} \quad X_{11} \quad X_{12} \quad X_{13} \quad X_{20} \quad X_{21} \quad X_{22} \quad X_{23} \quad X_{30} \quad X_{31} \quad X_{32} \quad X_{33}]^T \tag{4}$$

[0053]   The non-separable transform is calculated as $\vec{F} = T \cdot \vec{X}$, where $\vec{F}$ indicates the transform coefficient vector, and T is a 16x16 transform matrix. The 16x1 coefficient vector $\vec{F}$ is subsequently re-organized as 4x4 block using the scanning order for that block (horizontal, vertical or diagonal).

[0054]   *Cross component prediction for chroma intra prediction.* Coding performance can be improved by exploring the cross component correlation. In JEM the cross-component Linear

Model (LM) prediction mode is used for chroma intra prediction. In LM prediction mode, the chroma samples are predicted based on reconstructed luma samples of the same block by using a linear model as follows:

$$pred_C(i,j) = \alpha \cdot rec_L(i,j) + \beta \qquad (5)$$

where $pred_C(i,j)$ represents the prediction of chroma samples in a block and $rec_L(i,j)$ represents the reconstructed luma samples of the same block in the same resolution as chroma block, which may be down sampled for 4:2:0 chroma format content. Parameters α and β are derived by minimizing regression error between the neighboring reconstructed luma and chroma samples around the current block.

[0055]   Additionally, the LM prediction mode also exploits the correlation between two chroma components, i.e. Cr component is predicted from Cb component. Instead of using the reconstructed sample signal, the cross component prediction is applied in residual domain. This is implemented by adding a weighted reconstructed Cb residual to the original Cr intra prediction to form the final Cr prediction:

$$pred_{Cr}^*(i,j) = pred_{Cr}(i,j) + \gamma \cdot resi_{Cb}'(i,j) \qquad (6)$$

The scaling factor γ is derived in the same way as the derivation of parameters α and β described above. The only difference is an addition of a regression cost relative to a default γ value in the error function so that derived scaling factor is biased towards the default value (- 0.5).

Intra Prediction Coding Overhead.

[0056]   As discussed above, high coding efficiency can be achieved for the intra coding in HEVC and JEM by increasing the number of directions supported for the intra prediction. Additionally, CUs can be further split into multiple non-overlapped PUs, each of which can have a different intra prediction mode. This PU-based representation of intra prediction directions provides good coding efficiency especially for regions with different directional edges. However, in natural content, the edges of one object may be gradually changing along a certain range of directions, which does not always perfectly align with the shape of CU/PU partitions that are restricted to be squares. Therefore, a large amount of signaling overhead (e.g., the bits to indicate the depth of CU partition, the PU partition type and intra prediction directions) is expended on splitting the region into small CU/PUs, each with different intra prediction directions, in order for the predicted texture direction to be aligned with the actual edge direction of the object. To illustrate the abovementioned problem, FIGs. 7A-7D show a comparison of several partition

methods for intra coding. In FIGs. 7A-7D, it is assumed that the current region is a 16x16 block which contains an arc-shape object of FIG. 7A (as enclosed in the patterned stripe) with boundary samples along a gradually changing edge from 0° (direct horizontal) to 90° (direct vertical). In FIG. 7B, the whole region is coded as one CU (as enclosed in dashed lines) by one single intra prediction direction; however, samples located on the edges in different directions within the block cannot be precisely predicted. FIG. 7C divides the whole region into four CUs (as enclosed in dashed lines) and each CU is predicted by its own direction; as a result, the number of precisely predicted samples is increased. FIG. 7D shows the partition structure where both CU (as enclosed in dashed lines) and PU (as enclosed in dotted lines) partitions are applied. This method further increases the intra prediction accuracy by assigning different intra prediction modes to each CU/PU. However, as mentioned earlier, the penalty of this method is a large number of bits need to be signaled to indicate partition structure and intra prediction directions in bit-stream.

Control-Point Based Intra Prediction.

[0057]   Systems and methods are disclosed herein for intra coding based on a control-point representation of prediction directions. The proposed systems and methods may improve the efficiency of intra coding by increasing the precision of intra prediction directions without substantially increasing the overhead of signaling block partitions and intra prediction modes. In an exemplary method, the intra prediction direction of samples inside a video block is derived from a set of controlling intra prediction directions for the current block. Additionally, a flag may be signaled to indicate whether the control-point based intra prediction is enabled or not. In case control-point based intra prediction is enabled, the intra modes of the control-points for the current block may be transmitted to the decoder, and the intra modes of the samples inside the block are derived from the control-point intra directions. In cases where the texture of the samples in the current block changes direction, control-point based intra prediction can provide a good estimate of the intra prediction directions for the samples inside the block. Systems and methods of control-point based intra prediction can make the use of relatively large intra coding units (CUs) and prediction units (PUs), therefore reducing the overhead used for signaling CU/PU partitions and intra direction information.

[0058]   As discussed above, for PU-based intra direction signaling, the adaptation of intra directions is restricted by the square shapes of PU partitions. Since the local object edge(s) may gradually change directions within a block, in order to align the intra prediction direction with the local object edge(s), a large amount of overhead is usually needed in order to split the current block

into small PUs. To address this issue, a new control-point based representation method of intra prediction directions is described herein. Exemplary embodiments can improve the intra coding efficiency by increasing the precision of intra prediction directions without significantly increasing the overhead of signaling block partitions and intra prediction directions. Specifically, in exemplary embodiments, intra prediction directions at one or more control points are signaled for a given video block. Then, by exploiting the spatial correlation between adjacent samples (or blocks), the intra prediction direction of each sample (or a sub-block of samples) within the current block is derived from the intra prediction directions of the control points.

[0059] Given that the intra directions of samples in the current block are directly derived from those at the control points, the proposed systems and methods have relatively small signaling overhead of intra directions, as only one or more intra directions per block need to be signaled. Additionally, as the proposed representation method allows derivation of intra prediction directions at finer granularity level (e.g., sample level and various sub-block levels), there is less of a need to split the current region into small CU/PUs. This could lead to an increased usage of large intra CU/PUs, therefore reduce the overhead used for signaling CU/PU partitions. Both of those two aspects can result in improved coding performance of video blocks that are coded using control-point intra prediction.

Hybrid Video Coding with a Control-Point Based Intra Prediction Mode.

[0060] As described above, exemplary control-point based intra prediction systems and methods may reduce the signaling overhead of block partitions (e.g., CU, PU and TU partitions) and intra prediction directions. If the control-point based intra prediction model can provide a good estimate of local intra directions within a region, there is no need to split the region into small CU/PU partitions in order to align block boundaries with the variation of local edges; also, the intra prediction directions inside the region can be derived from those directions at control-points, instead of being explicitly signaled. Therefore, compression gain is achieved. However, the proposed control-point based intra prediction model may not be always accurate enough and may impair the intra prediction efficiency due to large prediction errors. Therefore, to further improve the coding performance, the proposed method is proposed to be adaptively selected at certain block level. In some embodiments, a flag is signaled in the bitstream to transmit the encoder decision to the decoder, indicating that the current block is predicted using either a control-point intra prediction mode or the existing intra prediction mode. For example, the encoder may use Rate-Distortion (R-D) optimization decision in a way similar to the Lagrangian optimization method

used in HM and JEM to determine whether to use the control-point intra prediction to predict a current block.

[0061] FIG. 8 is a functional block diagram of an exemplary hybrid video encoder configured to implement a control-point based intra prediction mode. For intra-coded blocks, the encoder estimates the intra prediction mode by selecting from conventional intra prediction modes and the proposed method, for example, using R-D optimization. Similar to the existing video encoder (as shown in FIG. 1), the intra prediction module generates the intra prediction signal by predicting the samples of the block from its adjacent reconstructed samples belonging to neighboring block. The original signal is subtracted by the intra prediction signal to remove spatial redundancy and produce the corresponding residual signal. The transform and quantization are applied to the residual signal which are then entropy-coded and output to bit-stream. To obtain the reconstructed signal, the reconstructed residual signal is inverse quantized followed by inverse transform. The reconstructed residual is added to the prediction signal. Further, in-loop filtering processes, e.g., de-blocking, SAO and ALF, are applied to the reconstructed video signal for output. As mentioned above, to improve the coding gain, a control point intra mode prediction flag (e.g. *control_point_intra_prediction_mode_flag*) may be signaled at a given block level (e.g., CU, PU or TU level) to indicate whether the proposed control-point based intra prediction is applied to the current block. If it is applied to the current block, a control-point based intra prediction module 802 performs intra prediction using techniques described herein.

[0062] FIG. 9 is a functional block diagram of a hybrid video decoder configured to implement a control-point based intra prediction mode. The decoder of FIG. 9 may be configured to receive the bit-stream produced by the encoder in FIG. 8. At the decoder, the bit-stream is parsed by the entropy decoder. The residual coefficients are inverse quantized and inverse transformed to obtain the reconstructed residual. The coding mode and prediction information are used to obtain the prediction signal using either spatial prediction or temporal prediction. For intra-coded blocks, if the control point intra mode prediction flag (e.g. *control_point_intra_prediction_mode_flag*) is true, the prediction signal is generated by the proposed control-point based intra prediction method using a control-point based intra prediction module 902; otherwise, the prediction signal is generated by a conventional intra prediction process. The prediction signal and the reconstructed residual are added together to get the reconstructed video. The reconstructed video may additionally go through loop filtering before being stored in the reference picture store to be displayed and/or to be used to decode future video signal.

[0063] Several features of various control-point based intra prediction encoding and decoding methods proposed herein are described in greater detail below.

Direction Representation of a Control-Point Based Intra Prediction Mode.

[0064] In an exemplary coding process, consider a current CU, which is an $N \times N$ block, and set the coordinate system as shown in FIG. 10. The top-left, top-right and bottom-left corners, $(0,0)$, $(N-1,0)$ and $(0, N-1)$, serve as three control points, denoted as CP_0, CP_1 and CP_2, respectively. The target sample, which is located at position (x, y), specifies the current sample for which it is desired to estimate the intra prediction direction. Let $IM(CP_i)$ and $IM(x, y)$ be the intra prediction direction of the i-th controlling point $CP_i, i = 0,1,2$, and the intra prediction direction of the target sample to be estimated. Then the proposed control-point based intra direction derivation method in the present embodiment is described as

$$IM(\text{x, y}) = \sum_{i=0}^{K-1} \phi_i(x, y) \, IM(CP_i) \qquad (7)$$

where $\phi_i(x, y)$ is an interpolation kernel function which specifies the contribution of the i-th controlling intra direction to the resulting intra direction of the target sample, and K is the number of control points used in the direction interpolation process (K is equal to 3 in the example shown by FIG. 10). Different interpolation kernel functions, e.g., triangular interpolation, bilinear interpolation, Gaussian, spline interpolation and so forth, may be applied in exemplary embodiments. It should be noted that the term "interpolation" as used herein further encompasses extrapolation in that interpolation is not limited herein to points that fall strictly within geometric bounds of the control points.

[0065] In an exemplary embodiment in which triangular interpolation is used, the interpolation kernel functions associated with the three controlling points in (7) may be given as

$$\phi_0(x, y) = (N - 1 - x - y)/(N - 1)$$

$$\phi_1(x, y) = x/(N - 1) \qquad (8)$$

$$\phi_2(x, y) = y/(N - 1)$$

[0066] If the target point is located in the triangle formed by CP_1, CP_0 and CP_2 (the samples within the shadow region in FIG. 10), the triangular interpolation result will be within the range restricted by control points. If the target point is outside the triangle (samples in the blank region

in FIG. 10), the triangular interpolation result may be clipped into a valid range because it is extrapolated and possibly outside the range restricted by control points.

[0067]   In another example, bilinear interpolation is used. In this case, the interpolation kernel functions of three controlling points may be given as

$$\phi_0(x, y) = (N - 1 - x)(N - 1 - y)/(N - 1)^2$$

$$\phi_1(x, y) = \big(x(N - 1 - y) + x(N - 1)\big)/(2(N - 1)^2) \tag{9}$$

$$\phi_2(x, y) = \big(y(N - 1 - x) + y(N - 1)\big)/(2(N - 1)^2)$$

As shown in Equations (7)-(9), the interpolation processes of the proposed control-point based intra direction derivation method may involve division by $(N - 1)$ or $(N - 1)^2$. Given that video block size N is usually a power of 2, division by $(N - 1)$ or $(N - 1)^2$ usually cannot be implemented as bit-shift operations and thus requires undesirable computational complexity. In some embodiments, to perform division more efficiently, fixed point approximation may be used to calculate the result. Without loss of generality, consider division by $(N - 1)$ as an example. This division may be approximated by two integer values M and B, as indicated as

$$\frac{1}{N - 1} \approx \frac{M}{2^B}$$

Thus, in some embodiments, the division may be implemented as a multiplication by M followed by a bit-shift of B to the right.

[0068]   Alternatively, in some embodiments, to promote efficient division, the control-point locations may be slightly overlapped between neighboring blocks. Specifically, using the coordinate system as defined in FIG. 10, the location of the top-left control-point (CP_0,) is unchanged, at $(0, 0)$. However, the locations of the top-right, bottom-left control-points (CP_1 and CP_2) are moved to cross the block boundary by one additional column and one additional row respectively, such that the new locations CP_1 and CP_2 will become $(N, 0)$ and $(0, N)$ which belong, respectively, to the neighboring blocks to the right of and below the current block. Correspondingly, the value of $(N - 1)$ in equations (7)-(9) becomes N after this adjustment. Thus, where N is a power of 2, the divisions in equations (7)-(9) may be implemented using bit shifts.

[0069]   In yet another embodiment, the control-points CP_1 and CP_2 are obtained from the top-right and bottom-left corners of the current block, namely from the locations of $(N - 1, 0)$ and $(0, N - 1)$ respectively, while the control-point CP_0 is obtained from the top-left neighboring sample of the current block, namely the location $(-1, -1)$. This way, the denominators in

equations (7)-(9) also become N and the divisions can be implemented using bit shifts where N is a power of 2.

Selection of Control Points.

[0070]   In video coding, the coding of the intra direction can take a significant amount of bits and therefore impact the overall coding performance, especially at low bit-rate. The resulting overhead of signaling control intra directions could neutralize the advantages of the proposed control-point representation. In different embodiments, different number of control-points with various locations can be selected, which may result in different coding performance of video content. Using more control points is advantageous in the sense that it can yield a more accurate estimation of varying intra prediction directions inside one video block. However, a larger number of control points also leads to an increased overhead given that more controlling intra directions need to be sent. On the other hand, a smaller number of control-points can reduce signaling overhead while not always being able to provide an accurate intra prediction estimation. Therefore, exemplary embodiments allow a different number of control-points to be selected for different video content, with the number of control points being large enough for accurate estimation of intra prediction directions inside the block while leading to an appropriate signaling overhead. In one embodiment, the top-left and top-right corners of the current video block are used as the control-points for the proposed intra prediction derivation method (as shown in FIG. 11A). In a case where bilinear interpolation is applied, the interpolation kernel functions of control-points (as shown in Equation (9)) can be represented as

$$\phi_0(x, y) = (N - 1 - x)/(N - 1)$$
$$\phi_1(x, y) = x/(N - 1) \tag{10}$$

[0071]   Given that two control-points are obtained horizontally, this control-point selection is more efficient for estimating intra directions of video blocks that have dominant edges in the horizontal direction. In another embodiment, the top-left and bottom-left corners of the current video block are used as the control-points for the proposed intra prediction derivation method (as shown in FIG. 11B). Assuming the bilinear interpolation is applied, the interpolation kernel functions of control-points (as shown in Eq. (9)) can be represented as

$$\phi_0(x, y) = (N - 1 - y)/(N - 1)$$
$$\phi_2(x, y) = y/(N - 1) \tag{11}$$

[0072]   Because the two control-points are obtained vertically, this control-point selection is efficient for estimating intra directions of video blocks that have dominant edges in a vertical direction.

[0073]   In embodiments in which the top-left and bottom-right corners of the current video block are used as control-points (as shown in FIG. 11C), the interpolation kernel function may be represented as the following:

$$\phi_0(x, y) = (N - 1 - (x + y)/2)/(N - 1)$$

$$\phi_3(x, y) = ((x + y)/2)/(N - 1)$$

(12)

[0074]   If top-right and bottom-left corners of the current video block are used as control-points (as shown in FIG. 11D), the interpolation kernel function may be represented as the following:

$$\phi_1(x, y) = ((N - 1 + x - y)/2)/(N - 1)$$

$$\phi_2(x, y) = ((N - 1 - x + y)/2)/(N - 1)$$

(13)

[0075]   The proposed control-point based intra prediction derivation methods may use various numbers of control points for the coding based on the content's characteristics, such as three control points or two control points. In some embodiments, the control-point configuration is signaled at the picture level or the slice level. For example, the identification of a control point within the block can be defined as in Table 3, where block_width and block_height are the width and height of the block that control-point is applied to. An exemplary control-point table is defined in Table 4, which specifies the number of control-points and the position of each control point. At the picture level (e.g. in the picture parameter set) or sequence level (e.g. in the sequence parameter set), multiple control-point tables defined in Table 4 may be signaled, and they may be differentiated by the *"control_points_model_id"*. The *"control_points_model_id"* may also be signaled at the slice level or coding block level (e.g. CTU, CU).

Table 3. Index of control point.

CP_Id	Position
0	top-left: (0, 0)
1	top-right: (block_width-1, 0)
2	bottom-left: (0, block_height-1)
3	bottom-right: (block_with-1, block_height-1)
4	customized: (x, y) to be signaled explicitly

Table 4. Signaling of control-points

control_points_table ( ) {	Descriptor
control_points_model_id	ue(v)
num_control_points	ue(v)
for (int i=0; i< num_control_points; i++) {	
CP_Id	ue(v)
if (cp_id == 4) {	
pos_x	u(8)
pos_y	u(8)
}	
}	
}	

control_points_model_id : indicates the ID of control_point model signaled in the table.

num_control_points: specifies the number of control points for this model.

CP_Id: specifies the ID defined in Table 3 of the control point.

pos_x: specifies horizontal position of control point with ID being equal to 4. The actual position is calculated as (pos_x*block_width + 128)/256. block_width is the width of the block.

pos_y: specifies vertical position of control point with ID being equal to 4. The actual position within the block is calculated as (pos_x*block_height + 128)/256. block_height is the height of the block.

[0076]   It should be noted that with respect to Equations (10)-(13), techniques as described above for efficient division by $(N - 1)$ may also be applied.

[0077]   As shown in FIGs. 11A and 11B, selecting different control-points may be beneficial to video regions presenting different directional characteristics. Taking this aspect into consideration, region-based control-point adaptation may be applied. Specifically, in an exemplary method, the horizontal control-point selection (as shown in FIG. 11A) is enabled for regions that have dominant horizontal edge information, and the vertical control-point selection (as shown in FIG. 11B) is enabled for regions that have dominant vertical edge information. For non-flat regions whose edges are neither horizontal nor vertical, the three control-points (as shown in FIG. 10) may be selected. The dominant edge direction in one region can be determined from the reconstructed neighboring samples of the current video block. In one embodiment, the Sobel operator is used to calculate the horizontal and vertical derivatives of neighboring reconstructed samples.

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (14)$$

[0078] If the ratio between G_y and G_x is larger than a predefined threshold (T_1), then the region may be classified as horizontal edge region; if the ratio between G_y and G_x is smaller than another predefined threshold $(T_2, T_2 \ll T_1)$, then region may be classified as vertical edge region; otherwise, the edge of the region may be regarded as neither horizontal nor vertical.

[0079] In another embodiment, the likely direction of the texture inside the current block may be derived from the set of MPMs. For example, if majority modes in the MPM set indicate a horizontal direction, then the two horizontal control points in FIG. 11A may be used. If majority modes in the MPM set indicate vertical direction, then the two vertical control points in FIG. 11B may be used. If there is not strong indication of horizontal or vertical from the MPM modes, then the three control points in FIG. 10 may be used.

Granularity of Deriving Intra Prediction Direction Based on Control Points.

[0080] In the embodiments of FIG. 10 and FIGs. 11A-11D, the intra prediction directions are derived based on sample unit (such that each sample inside the video block may have its own intra direction); however, intra direction derivation methods in accordance with the present disclosure can be conducted at various levels, such as CU-, PU- and TU-level, and each adaptation level may provide different intra direction adaptation capability and encoding/decoding complexity. For example, if the intra prediction direction is derived at the TU level, the encoder/decoder derives the intra prediction mode for each TU and the samples within one TU will use the same derived intra direction. If the intra prediction direction is derived at the PU, the encoder or decoder may only need to derive intra directions for each PU, and all the TUs inside that the PU will use the same derived intra direction for their own intra prediction. In a case where three control-points are used, FIGs. 12A-12B show the cases of deriving intra prediction directions at the PU-level (FIG. 12A) and at the TU-level (FIG. 12B) for the proposed control-point based intra prediction method, where solid and dashed lines depict the boundaries of PUs and TUs respectively. In FIG. 12A the target block is prediction unit PU₃, and in FIG. 12B the target block is transform unit TU₀. As an exemplary embodiment, in FIGs. 12A-12B, the center position of the target block is used as the representative coordinate to derive the intra direction for the block. In other embodiments, other positions may be selected as the representative coordinate.

[0081] The level at which control-point based intra prediction is enabled/disabled (e.g. the level at which the flag *control_point_intra_prediction_mode_flag* is signaled, or the level at which control-points are defined) is in many embodiments higher than the level of intra direction derivation. For example, the control-point based intra prediction may be enabled at the CU level

while the corresponding intra prediction derivation may be performed at PU-level, TU level or sample level. If the intra direction derivation is carried over at the same level of defining control-points (e.g., both at CU-level), then the proposed method degenerates to the normal intra mode because only one single control-point is needed, which is actually the signaled intra prediction mode of the block.

Predicting and Signaling Control-Point Intra Directions.

[0082]   The signaling overhead of the control-point's intra directions affects the overall performance of the proposed control-point based intra prediction mode. To reduce the signaling overhead, in one embodiment of the disclosure, it is proposed to directly reuse the existing intra mode signaling methods in HM and JEM for signaling control-point intra directions. Specifically, for each control point, a set of MPM candidates are generated from DC, planar and intra modes of its left and top neighbors. In case the control intra direction is identical to one of the MPM candidates, only the index corresponding to the candidate is sent; otherwise, the control intra direction is signaled using a fixed length code. In another embodiment, the MPM candidate list may be constructed by only including angular directions when control-point based intra prediction mode is enabled because non-angular intra directions (DC and planar) are not used in control-point intra prediction.

[0083]   Differential coding is widely employed to reduce the signaling overhead of various syntax elements in modern video codec. For example, in HM and JEM, differential coding is used for efficient motion vector (MV) coding where one MV is predicted from one of its spatial and temporal neighbors and only the residue needs to be transmitted. Given the strong correlation with intra prediction directions of neighboring video blocks, for one video block, the control-point's intra prediction directions may be very similar. Therefore, in one embodiment of the disclosure, it is proposed to use differential coding for signaling control-point intra directions. More specifically, one of the control-points (e.g., CP_0 in FIGs. 10 and 11) is firstly signaled based on the existing MPM-based intra mode prediction method. Then, the intra direction of this control-point is used to predict the intra directions of the other control-points (e.g., CP_1 and CP_2 in FIGs. 10 and 11). Correspondingly, only the difference between the modes is signaled.

High-Accuracy Intra Prediction for Control-Point Based Intra Prediction Mode.

[0084]   In JEM, the number of the angular intra modes is 65. The number of 65 is selected in order to provide an increased intra prediction granularity without overwhelmingly increasing the signaling overhead of intra prediction mode which may neutralize the coding gain obtained from

the increased intra prediction accuracy. In embodiments disclosed herein, the intra directions of samples inside the current block are derived from control-points, and no overhead needs to be signaled for those derived intra directions inside the current block (e.g., using Equation (7)) in the bit-stream. Therefore, when a control-point based intra prediction mode is enabled, the number of directional intra prediction modes of those samples inside the block can be increased with no signaling cost. In order to more accurately capture the edge directions in natural videos, in one embodiment of the disclosure, it is proposed to increase the number of angular intra prediction modes with a finer granularity direction indication when the control-point based intra prediction mode is enabled.

[0085]   In HEVC, bilinear interpolation is applied to obtain intra predictions when the target samples are predicted from one reference sample at a fractional position. Specifically, each predicted sample P(x, y) at location (x, y) is obtained by projecting its location onto a reference row of pixels based on the selected prediction direction and interpolating a value for the predicted sample at 1/32 sample accuracy, where bilinear interpolation is performed utilizing the two closest reference samples, R_i and R_{i+1}, at integer positions.

$$p(x, y) = \left(\left(32 - w_y \right) \cdot R_i + w_y \cdot R_{i+1} + 16 \right) \gg 5 \qquad (15)$$

where w_y is the weighting parameter between two integer reference samples R_i and R_{i+1} which specifies the segment covering the projected fractional location and $\gg$ denotes right bit-shift operation. The reference sample index i and w_y are calculated based on the projection displacement d corresponding to the selected intra prediction mode (which describes the tangent of the intra prediction direction in the precision of 1/32 sample) as

$$c_y = (y \cdot d) \gg 5$$

$$w_y = (y \cdot d) \& 31 \qquad (16)$$

$$i = x + c_y$$

where "&" denotes bitwise AND operation. Table 5 specifies exemplary values of the projection displacement d for different intra prediction directions.

Table 5. Mapping between intra prediction direction and projection displacement d for the 33 angular intra prediction directions in HEVC.

predModeIntra	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
d	-	32	26	21	17	13	9	5	2	0	−2	−5	−9	−13	−17	−21	−26
predModeIntra	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
d	−32	−26	−21	−17	−13	−9	−5	−2	0	2	5	9	13	17	21	26	32

[0086] As discussed above, when a control-point intra prediction mode is selected, the number of supported angular intra prediction directions may be increased. Correspondingly, the existing 1/32 sample accuracy may be insufficient to represent all possible directional intra predictions in the finer granularity. Therefore, in one embodiment of the disclosure, it is proposed to increase the accuracy of the intra interpolation filtering when the proposed control-point based intra prediction is enabled for the current video block. Denoting 2^B as the number of defined fractional samples for the finer granularity intra prediction, the bilinear filtering process (of Equation (15)) and the reference sample index derivation (of Equation (16)) may be expressed as

$$p(x, y) = \left(\left(2^B - w_y \right) \cdot R_i + w_y \cdot R_{i+1} + 2^{B-1} \right) \gg B \tag{17}$$

$$c_y = (y \cdot d) \gg B \tag{18}$$

$$w_y = (y \cdot d) \& 2^B$$

$$i = x + c_y$$

[0087] Additionally, the dynamic range of the projection displacement d may also be enlarged in order to represent the tangent values of the increased intra prediction granularity. In an embodiment in which the number of intra prediction directions is 131 (including planar, DC and 129 angular intra directions) with the accuracy of intra interpolation being 1/64 sample, Table 6 shows exemplary projection displacement values d associated with the 129 angular intra prediction directions.

Table 6. Mapping between intra prediction direction and projection displacement d for 129 angular intra prediction directions.

predModeIntra	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
d	-	64	61	58	55	52	49	46	44	42	40	38	36	34	32	30	28
predModeIntra	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
d	26	24	22	20	18	16	14	12	10	8	6	5	4	3	2	1	0
predModeIntra	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51
d	-1	-2	-3	-4	-5	-6	-8	-10	-12	-14	-16	-18	-20	-22	-24	-26	-28
predModeIntra	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68
d	-30	-32	-34	-36	-38	-40	-42	-44	-46	-49	-52	-55	-58	-61	-64	-61	-58
predModeIntra	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85
d	-55	-52	-49	-46	-44	-42	-40	-38	-36	-34	-32	-30	-28	-26	-24	-22	-20
predModeIntra	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102
d	-18	-16	-14	-12	-10	-8	-6	-5	-4	-3	-2	-1	0	1	2	3	4
predModeIntra	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
d	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36
predModeIntra	120	121	122	123	124	125	126	127	128	129	130						
d	38	40	42	44	46	49	52	55	58	61	64						

[0088] In another embodiment, a control-point based intra prediction method with increased angular direction granularity may be used in combination with other interpolation methods. Denoting $f_{w_y}(k)$ as the coefficients of the applied interpolation filter, $k = -(L/2 - 1), \ldots, 0, 1, \ldots L/2$, where L is the interpolation filter length, the value of the predicted sample P(x, y) at location (x, y) may be calculated as

$$p(x,y) = \left(\sum_{k=-(L/2-1)}^{L/2} f_{w_y}(k)\, R_{i+k} + 2^{H-1} \right) \gg H \tag{19}$$

$$c_y = (y \cdot d) \gg B \tag{20}$$

$$w_y = (y \cdot d) \& 2^B$$

$$i = x + c_y$$

where H is the number of bits for representing filter coefficients. As shown in Equation (19), the filter coefficients $f_{w_y}(k)$ is dependent on the corresponding phase (i.e., w_y, $w_y = 0, 1, \ldots, 2^B - 1$) of the interpolated fractional sample in-between the corresponding integer samples.

[0089] When one intra block is coded in control-point based intra prediction mode, the derived intra prediction direction may be used as one of MPM candidates to predict the intra prediction modes of neighboring intra blocks. Correspondingly, the high-accuracy intra prediction directions of the control-point coded blocks may be rounded to the intra prediction accuracy of normal intra mode in order to ensure the predicted intra prediction direction and the MPM candidate having the same precision. Therefore, in one embodiment of the disclosure, when finer granularity intra

prediction is enabled (e.g. 129 angular directions) for control-point coded blocks, it is proposed to round the high-accuracy intra prediction mode of the current block to the nearest intra prediction mode in the coarse granularity (e.g. 65 angular directions in JEM and 33 in HM) before it is stored and used as MPM candidate for predicting the intra prediction mode of its neighboring intra blocks. In a case where two neighboring coarse-accuracy intra directions have equal distance to the high-accuracy intra direction, a fixed rule (e.g., always rounding to the coarse-accuracy intra direction with smaller intra direction index, or always rounding to the coarse-accuracy intra direction which is closer to 45 degrees) may be applied.

[0090]   In some embodiments, signaled intra prediction directions at the control points of the current block are used to predict the intra directions of adjacent blocks for which the intra directions are explicitly signaled. Therefore, to reduce the overhead of signaling intra mode, it is proposed to use the intra direction of the top-right control-point (CP_1 in in FIG. 10) as one candidate to generate the MPM list of the neighboring block to the right of the current block; similarly, it is also proposed to use the intra direction of the bottom-left control-point (CP_2 of FIG. 10) as one candidate to generate the MPM list of the neighboring block below the current block. Additionally, in one embodiment of the disclosure, it is proposed to add the intra direction of the control-point as one additional MPM candidate and increase the size of the MPM list by one. In another embodiment, to not increase the signaling overhead of MPM index, it is proposed to use the intra direction of the control-point to replace one of the existing MPM candidates (e.g., the last candidate in the MPM list) such that the size of the MPM list is kept unchanged.

Decoder-Side Control-Point Derivation.

[0091]   In exemplary methods described above, the control-point's intra directions are signaled from encoder to decoder for the control-point based intra prediction mode, which could consume a non-negligible number of bits and therefore reduce the overall coding performance of the proposed control-point based intra prediction mode. In some embodiments, a decoder-side intra mode derivation (DIMD) approach may be used to reduce the overhead of signaling intra prediction parameters by estimating the intra prediction mode of one video block based on a group of reconstructed neighboring samples (referred to as template) of the current block at the decoder side. To reduce the signaling overhead of controlling intra directions, it is proposed in one embodiment of the disclosure to apply the DIMD approach to the proposed control-point based intra prediction method. In this case, the intra directions of the control points may be derived at the decoder side using DIMD instead of explicitly signaled. Specifically, when the control-point

based intra prediction is enabled, instead of indicating control-point's intra directions in bit-stream, the encoder/decoder derives the control-point's intra directions from the reconstructed samples adjacent to the current block. FIG. 13 illustrates the proposed decoder-side control-point derivation method. Given the locations of the selected control-points (CP_0, CP_1 and CP_2), template samples (as indicated by different patterned regions for three control-points) refer to a set of already reconstructed samples that are used to derive the intra prediction direction of the corresponding control-point. Based on the template samples of each control-point, a decoder-side control intra direction derivation method may include a step of calculating, for each intra prediction direction, the cost between the reconstructed template samples and the prediction samples of the template. The template cost may be measured as the distortion between the two. The intra direction that yields the minimum cost may then be selected as the control intra direction.

[0092]   In different embodiments, different cost measurements may be applied in the decoder-side intra direction derivation method, such as sum of absolute difference (SAD), sum of square difference (SSD) and sum of absolute transformed difference (SATD). The cost may also be measured as the sum of cost measurement of small sub-block unit (e.g. 2x2 or 4x4) contained in the template.

[0093]   The selection of template samples may be restricted based on the encoding/decoding order of blocks. In FIG. 13, the closest reconstructed samples around each control-point are used as a template. In practice, different template sizes may be used. A large number of template samples can lead to a more reliable estimation of intra direction in the presence of coding noise. However, large template size also increases the distance between the template samples and the control-point position for which the mode is estimated. Therefore, as the distance increases, the accuracy of such estimation may be reduced (given the reduced correlation between the template and the control-point). On the other hand, while a smaller template may provide more accurate estimation, it may be more sensitive to coding noise. Therefore, it is desirable to select an optimal template size for different block sizes, which should be large enough to be robust against noise, while not exceeding the size limit to accurate capture the direction of local edge. In one embodiment, a template size of 2 (i.e., $L = 2$ in FIG. 13) is used for 4x4 and 8x8 blocks and a template size of 4 (i.e., $L = 4$ in FIG. 13) is used for 16x16 and larger video blocks.

[0094]   In other embodiments of the disclosure, other intra mode derivation methods, including other decoder-side intra mode derivation techniques may be employed in an exemplary control-point based intra prediction method.

Control-Point Based Intra Prediction for Chroma Coding.

[0095] Since the human vision system is much more sensitive to variations in brightness than color, a video coding system usually devotes more bits to the luma component than chroma components, e.g., by adjusting the quantization parameter (QP) delta value between luma component and chroma components. Therefore, chroma components contain much less detail information (e.g., edges and textures) than the luma component. Based on this observation, one may signal fewer control points for chroma intra prediction than for luma intra prediction. On the other hand, given the strong correlation between the luma plane and chroma planes, it is highly probable that chroma predication uses the same intra prediction direction as the luma prediction. Therefore, in one embodiment of the disclosure, when the control-point intra prediction mode is enabled for one video block, the encoder/decoder reuses the same control-points of the luma component to derive the intra prediction directions of chroma components, such that the same intra prediction direction will be used for both luma and chroma components. In another embodiment of the disclosure, two different set of control-points are transmitted to decoder for luma and chroma components separately when the control-point based intra prediction mode is enabled.

Intra Prediction Methods Using Control-Point Based Intra Mode Representation.

[0096] In exemplary methods described above, control-point based intra prediction is generally performed in such a way that the intra directions of the control-points are firstly used to derive the intra direction for each sample (or sub-block) within the current video block according to Equation (7); then, the existing intra prediction process (as defined in HEVC or JEM) is applied to generate the prediction signal of the current block using the derived intra directions. Reusing the notations in Equation (7) and assuming $IP(intraDir)$ is the function of performing intra prediction using intra prediction direction $intraDir$, the control-point based intra prediction process as mentioned above can be described as

$$IP\big(IM(x, y)\big)= IP\left(\sum_{i=0}^{K-1} \big(\phi_i(x,y)IM(CP_i)\big)\right) \qquad (21)$$

[0097] In some embodiments, intra prediction operation $IP(\cdot)$ can be closely approximated as a linear function, and Equation (21) can be expressed in approximate form as follows

$$IP\big(IM(x, y)\big)= IP\left(\sum_{i=0}^{K-1}\big(\phi_i(x,y)IM(CP_i)\big)\right) \approx \sum_{i=0}^{K-1} IP\left(\phi_i(x,y)IM(CP_i)\right)$$

$$\approx \sum_{i=0}^{K-1}\phi_i(x,y)IP\big(IM(CP_i)\big) \tag{22}$$

[0098] In this equation, $IP\big(IM(CP_i)\big)$ corresponds to applying intra prediction to one sample using the intra direction of the control-point CP_i. Therefore, based on Equation (22), another intra prediction method may be employed using the proposed control-point based intra prediction method. In this method, given K control-points $CP_i, i = 0, 1, K-1$, the encoder/decoder firstly generates K different intra prediction signals for the current block, each using the intra prediction direction of one control-point, $IM(CP_i)$, respectively; then, the final intra prediction is obtained by employing weighted combination of these K control-point based intra prediction signals $(IM(CP_i))$ where the weight may be calculated according to the interpolation kernel function as specified in Equation (7). In a case in which three control points are used for the current block $(K = 3)$, FIG. 14 shows one example to illustrate the weighted combination of intra prediction signals using the intra directions of control-points. In Equation (22), the weight factor is dependent on the location of the target sample (x, y). In another embodiment, other weight factors (e.g., not position-dependent or not using the kernel functions) may be applied in Equation (22) to combine intra prediction signals generated with the direction from different control-points.

[0099] As compared to methods using Equation (7), a control-point based intra prediction method using Equation (22) allows the intra prediction directions at the control points to be non-angular direction modes, such as, DC and planar. Therefore, when MPM-based intra mode prediction is applied to predict the controlling intra directions, the MPM candidate list may be generated by including both angular and non-angular intra directions.

[0100] FIG. 15 is a schematic diagram illustrating an example of a coded bitstream structure. A coded bitstream 1000 consists of a number of NAL (Network Abstraction layer) units 1001. A NAL unit may contain coded sample data such as coded slice 1006, or high level syntax metadata such as parameter set data, slice header data 1005 or supplemental enhancement information data 1007 (which may be referred to as an SEI message). Parameter sets are high level syntax structures containing essential syntax elements that may apply to multiple bitstream layers (e.g. video parameter set 1002 (VPS)), or may apply to a coded video sequence within one layer (e.g. sequence parameter set 1003 (SPS)), or may apply to a number of coded pictures within one coded video sequence (e.g. picture parameter set 1004 (PPS)). The parameter sets can be either sent together

with the coded pictures of the video bit stream, or sent through other means (including out-of-band transmission using reliable channels, hard coding, etc.). Slice header 1005 is also a high level syntax structure that may contain some picture-related information that is relatively small or relevant only for certain slice or picture types. SEI messages 1007 carry the information that may not be needed by the decoding process but can be used for various other purposes such as picture output timing or display as well as loss detection and concealment.

[0101]   FIG. 16 is a schematic diagram illustrating an example of a communication system. The communication system 1300 may comprise an encoder 1302, a communication network 1304, and a decoder 1306. The encoder 1302 may be in communication with the network 1304 via a connection 1308, which may be a wireline connection or a wireless connection. The encoder 1302 may be similar to the block-based video encoder of FIG. 1. The encoder 1302 may include a single layer codec (*e.g.*, FIG. 1) or a multilayer codec. For example, the encoder 1302 may be a multi-layer (*e.g.*, two-layer) scalable coding system with picture-level ILP support. The decoder 1306 may be in communication with the network 1304 via a connection 1310, which may be a wireline connection or a wireless connection. The decoder 1306 may be similar to the block-based video decoder of FIG. 2. The decoder 1306 may include a single layer codec (*e.g.*, FIG. 2) or a multilayer codec. For example, the decoder 1306 may be a multi-layer (*e.g.*, two-layer) scalable decoding system with picture-level ILP support.

[0102]   The encoder 1302 and/or the decoder 1306 may be incorporated into a wide variety of wired communication devices and/or wireless transmit/receive units (WTRUs), such as, but not limited to, digital televisions, wireless broadcast systems, a network element/terminal, servers, such as content or web servers (*e.g.*, such as a Hypertext Transfer Protocol (HTTP) server), personal digital assistants (PDAs), laptop or desktop computers, tablet computers, digital cameras, digital recording devices, video gaming devices, video game consoles, cellular or satellite radio telephones, digital media players, and/or the like.

[0103]   The communications network 1304 may be a suitable type of communication network. For example, the communications network 1304 may be a multiple access system that provides content, such as voice, data, video, messaging, broadcast, *etc.*, to multiple wireless users. The communications network 1304 may enable multiple wireless users to access such content through the sharing of system resources, including wireless bandwidth. For example, the communications network 1304 may employ one or more channel access methods, such as code division multiple access (CDMA), time division multiple access (TDMA), frequency division multiple access (FDMA), orthogonal FDMA (OFDMA), single-carrier FDMA (SC-FDMA), and/or the like. The

communication network 1304 may include multiple connected communication networks. The communication network 1304 may include the Internet and/or one or more private commercial networks such as cellular networks, WiFi hotspots, Internet Service Provider (ISP) networks, and/or the like.

[0104]   FIG. 17 is a system diagram of an example WTRU in which an encoder or decoder as described herein may be implemented. As shown the example WTRU 1202 may include a processor 1218, a transceiver 1220, a transmit/receive element 1222, a speaker/microphone 1224, a keypad or keyboard 1226, a display/touchpad 1228, non-removable memory 1230, removable memory 1232, a power source 1234, a global positioning system (GPS) chipset 1236, and/or other peripherals 1238. It will be appreciated that the WTRU 1202 may include any sub-combination of the foregoing elements while remaining consistent with an embodiment. Further, a terminal in which an encoder (*e.g.*, encoder 100) and/or a decoder (*e.g.*, decoder 200) is incorporated may include some or all of the elements depicted in and described herein with reference to the WTRU 1202 of FIG. 17.

[0105]   The processor 1218 may be a general purpose processor, a special purpose processor, a conventional processor, a digital signal processor (DSP), a graphics processing unit (GPU), a plurality of microprocessors, one or more microprocessors in association with a DSP core, a controller, a microcontroller, Application Specific Integrated Circuits (ASICs), Field Programmable Gate Array (FPGAs) circuits, any other type of integrated circuit (IC), a state machine, and the like. The processor 1218 may perform signal coding, data processing, power control, input/output processing, and/or any other functionality that enables the WTRU 1500 to operate in a wired and/or wireless environment. The processor 1218 may be coupled to the transceiver 1220, which may be coupled to the transmit/receive element 1222. While FIG. 17 depicts the processor 1218 and the transceiver 1220 as separate components, it will be appreciated that the processor 1218 and the transceiver 1220 may be integrated together in an electronic package and/or chip.

[0106]   The transmit/receive element 1222 may be configured to transmit signals to, and/or receive signals from, another terminal over an air interface 1215. For example, in one or more embodiments, the transmit/receive element 1222 may be an antenna configured to transmit and/or receive RF signals. In one or more embodiments, the transmit/receive element 1222 may be an emitter/detector configured to transmit and/or receive IR, UV, or visible light signals, for example. In one or more embodiments, the transmit/receive element 1222 may be configured to transmit

and/or receive both RF and light signals. It will be appreciated that the transmit/receive element 1222 may be configured to transmit and/or receive any combination of wireless signals.

[0107] In addition, although the transmit/receive element 1222 is depicted in FIG. 17 as a single element, the WTRU 1202 may include any number of transmit/receive elements 1222. More specifically, the WTRU 1202 may employ MIMO technology. Thus, in one embodiment, the WTRU 1202 may include two or more transmit/receive elements 1222 (*e.g.*, multiple antennas) for transmitting and receiving wireless signals over the air interface 1215.

[0108] The transceiver 1220 may be configured to modulate the signals that are to be transmitted by the transmit/receive element 1222 and/or to demodulate the signals that are received by the transmit/receive element 1222. As noted above, the WTRU 1202 may have multi-mode capabilities. Thus, the transceiver 1220 may include multiple transceivers for enabling the WTRU 1500 to communicate via multiple RATs, such as UTRA and IEEE 802.11, for example.

[0109] The processor 1218 of the WTRU 1202 may be coupled to, and may receive user input data from, the speaker/microphone 1224, the keypad 1226, and/or the display/touchpad 1228 (*e.g.*, a liquid crystal display (LCD) display unit or organic light-emitting diode (OLED) display unit). The processor 1218 may also output user data to the speaker/microphone 1224, the keypad 1226, and/or the display/touchpad 1228. In addition, the processor 1218 may access information from, and store data in, any type of suitable memory, such as the non-removable memory 1230 and/or the removable memory 1232. The non-removable memory 1230 may include random-access memory (RAM), read-only memory (ROM), a hard disk, or any other type of memory storage device. The removable memory 1232 may include a subscriber identity module (SIM) card, a memory stick, a secure digital (SD) memory card, and the like. In one or more embodiments, the processor 1218 may access information from, and store data in, memory that is not physically located on the WTRU 1202, such as on a server or a home computer (not shown).

[0110] The processor 1218 may receive power from the power source 1234, and may be configured to distribute and/or control the power to the other components in the WTRU 1202. The power source 1234 may be any suitable device for powering the WTRU 1202. For example, the power source 1234 may include one or more dry cell batteries (*e.g.*, nickel-cadmium (NiCd), nickel-zinc (NiZn), nickel metal hydride (NiMH), lithium-ion (Li-ion), *etc.*), solar cells, fuel cells, and the like.

[0111] The processor 1218 may be coupled to the GPS chipset 1236, which may be configured to provide location information (*e.g.*, longitude and latitude) regarding the current location of the

WTRU 1202. In addition to, or in lieu of, the information from the GPS chipset 1236, the WTRU 1202 may receive location information over the air interface 1215 from a terminal (*e.g.*, a base station) and/or determine its location based on the timing of the signals being received from two or more nearby base stations. It will be appreciated that the WTRU 1202 may acquire location information by way of any suitable location-determination method while remaining consistent with an embodiment.

[0112]   The processor 1218 may further be coupled to other peripherals 1238, which may include one or more software and/or hardware modules that provide additional features, functionality and/or wired or wireless connectivity. For example, the peripherals 1238 may include an accelerometer, orientation sensors, motion sensors, a proximity sensor, an e-compass, a satellite transceiver, a digital camera and/or video recorder (*e.g.*, for photographs and/or video), a universal serial bus (USB) port, a vibration device, a television transceiver, a hands free headset, a Bluetooth® module, a frequency modulated (FM) radio unit, and software modules such as a digital music player, a media player, a video game player module, an Internet browser, and the like.

[0113]   By way of example, the WTRU 1202 may be configured to transmit and/or receive wireless signals and may include user equipment (UE), a mobile station, a fixed or mobile subscriber unit, a pager, a cellular telephone, a personal digital assistant (PDA), a smartphone, a laptop, a netbook, a tablet computer, a personal computer, a wireless sensor, consumer electronics, or any other terminal capable of receiving and processing compressed video communications.

[0114]   The WTRU 1202 and/or a communication network (*e.g.*, communication network 804) may implement a radio technology such as Universal Mobile Telecommunications System (UMTS) Terrestrial Radio Access (UTRA), which may establish the air interface 1215 using wideband CDMA (WCDMA). WCDMA may include communication protocols such as High-Speed Packet Access (HSPA) and/or Evolved HSPA (HSPA+). HSPA may include High-Speed Downlink Packet Access (HSDPA) and/or High-Speed Uplink Packet Access (HSUPA). The WTRU 1202 and/or a communication network (*e.g.*, communication network 804) may implement a radio technology such as Evolved UMTS Terrestrial Radio Access (E-UTRA), which may establish the air interface 1515 using Long Term Evolution (LTE) and/or LTE-Advanced (LTE-A).

[0115]   The WTRU 1202 and/or a communication network (*e.g.*, communication network 804) may implement radio technologies such as IEEE 802.16 (*e.g.*, Worldwide Interoperability for

Microwave Access (WiMAX)), CDMA2000, CDMA2000 1X, CDMA2000 EV-DO, Interim Standard 2000 (IS-2000), Interim Standard 95 (IS-95), Interim Standard 856 (IS-856), Global System for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE), GSM EDGE (GERAN), and the like. The WTRU 1500 and/or a communication network (*e.g.*, communication network 804) may implement a radio technology such as IEEE 802.11, IEEE 802.15, or the like.

[0116]   Note that various hardware elements of one or more of the described embodiments are referred to as "modules" that carry out (i.e., perform, execute, and the like) various functions that are described herein in connection with the respective modules. As used herein, a module includes hardware (e.g., one or more processors, one or more microprocessors, one or more microcontrollers, one or more microchips, one or more application-specific integrated circuits (ASICs), one or more field programmable gate arrays (FPGAs), one or more memory devices) deemed suitable by those of skill in the relevant art for a given implementation. Each described module may also include instructions executable for carrying out the one or more functions described as being carried out by the respective module, and it is noted that those instructions could take the form of or include hardware (i.e., hardwired) instructions, firmware instructions, software instructions, and/or the like, and may be stored in any suitable non-transitory computer-readable medium or media, such as commonly referred to as RAM, ROM, etc.

[0117]   Although features and elements are described above in particular combinations, one of ordinary skill in the art will appreciate that each feature or element can be used alone or in any combination with the other features and elements. In addition, the methods described herein may be implemented in a computer program, software, or firmware incorporated in a computer-readable medium for execution by a computer or processor. Examples of computer-readable media include electronic signals (transmitted over wired or wireless connections) and computer-readable storage media. Examples of computer-readable storage media include, but are not limited to, a read only memory (ROM), a random access memory (RAM), a register, cache memory, semiconductor memory devices, magnetic media such as internal hard disks and removable disks, magneto-optical media, and optical media such as CD-ROM disks, and digital versatile disks (DVDs). A processor in association with software may be used to implement a radio frequency transceiver for use in a WTRU, UE, terminal, base station, RNC, or any host computer.

CLAIMS

1. A method of coding a video in a bitstream, wherein the video comprises a plurality of frames, each frame comprising blocks of samples, the method comprising:

for at least a current block, identifying at least a first intra direction for a first control point and a second intra direction for a second control point;

for at least one current sample in the current block, deriving an intra prediction direction for the sample by interpolating a direction from at least the first intra direction and the second intra direction; and

predicting the sample with intra prediction using the derived intra prediction direction.

2. The method of claim 1, wherein the derived intra direction is interpolated from the first intra direction, the second intra direction, and a third intra direction for a third control point.

3. The method of any of claims 1-2, wherein positions of at least the first control point and the second control point are signaled in the bitstream.

4. The method of any of claims 1-3, wherein at least the first intra prediction direction is signaled in the bitstream.

5. The method of any of claims 1-4, wherein at least the first and second control points are selected from the group consisting of: a first control point at the top-left corner of the block, a second control point at the top-right corner of the block, a third control point at the bottom-left corner of the block, a fourth control point at the bottom-right corner of the block, a custom control point signaled in the bitstream.

6. The method of any of claims 1-5, wherein the first control point is at a top-left corner of the current block and the second control point is at a top-right corner or a bottom-left corner of the current block.

7. The method of claim 6, wherein the position of the second control point is determined based on intra modes in a list of most probable modes for the current block, the list being determined based on intra modes used to code neighboring blocks and wherein:

the second control point is at a top-right corner of the current block if more of the most probable modes are horizontal than vertical; and

the second control point is at a bottom-left corner of the current block if more of the most probable modes are vertical than horizontal.

8. The method of any of claims 1-7, wherein the current sample is in a current sub-block of the current block, further comprising predicting all samples of the current sub-block using the derived intra prediction direction.

9. The method of claim 8, wherein the current block includes a plurality of sub-blocks of samples, the method further comprising:

deriving respective intra prediction directions for each of the sub-blocks by interpolating respective directions from at least the first intra direction and the second intra direction; and

predicting samples in each of the sub-blocks by intra prediction using the respective derived intra prediction direction.

10. The method of any of claims 1-9, wherein the first and second intra directions are selected from a list of most probable modes, the list being determined based on intra modes used to code neighboring blocks.

11. The method of any of claims 1-10, wherein the intra direction of at least one of the control points for the current block is included in a list of most probable modes for intra coding of a neighboring block.

12. The method of any of claims 1-11, wherein at least the second intra direction is coded in the bitstream using differential coding with respect to the first intra direction.

13. The method of any of claims 1-3 or 5-12 further comprising deriving at least the first intra direction from samples in a template region adjacent to the control point.

14. The method of any of claims 1-13, wherein the interpolating is performed using triangular interpolation.

15. The method of any of claims 1-13, wherein the interpolating is performed using bilinear interpolation.

16. The method of any of claims 1-15, wherein the bitstream includes a flag indicating that control-point intra coding is used for the current block.

17. The method of any of claims 1-16, performed by a video encoder, the method further comprising determining a prediction residual for the sample and encoding the prediction residual in the bitstream.

18. The method of any of claims 1-16, performed by a video decoder, the method further comprising decoding a prediction residual for the sample and reconstructing the sample using the prediction and the residual.

19. An encoder for encoding a video into a bitstream, wherein the video comprises a plurality of frames, each frame comprising blocks of samples, the encoder being operative to perform functions comprising:

      for at least a current block, identifying at least a first intra direction for a first control point and a second intra direction for a second control point;

      for at least one current sample in the current block, deriving an intra prediction direction for the sample by interpolating a direction from at least the first intra direction and the second intra direction; and

   predicting the sample with intra prediction using the derived intra prediction direction.

20. A decoder for decoding a video from a bitstream, wherein the video comprises a plurality of frames, each frame comprising blocks of samples, the encoder being operative to perform functions comprising:

      for at least a current block, identifying at least a first intra direction for a first control point and a second intra direction for a second control point;

      for at least one current sample in the current block, deriving an intra prediction direction for the sample by interpolating a direction from at least the first intra direction and the second intra direction; and

   predicting the sample with intra prediction using the derived intra prediction direction.

FIG. 1

FIG. 2

$R_{0,0}$	$R_{1,0}$	$R_{2,0}$	...	$R_{N,0}$	$R_{N+1,0}$	...		$R_{2N,0}$

| $R_{0,1}$ |
| $R_{0,2}$ |
| ⋮ |
| $R_{0,N}$ |

Current TU

| $R_{0,N+1}$ |
| ⋮ |
| |
| $R_{0,2N}$ |

FIG. 3

FIG. 4

0: Planar
1: DC

FIG. 5

Filtered reference s[x,-1]

Intermediate prediction q[x,y]

Filtered reference s[-1,y]

Unfiltered reference r[x,-1]

Unfiltered reference r[-1,y]

Prediction p[x,y]
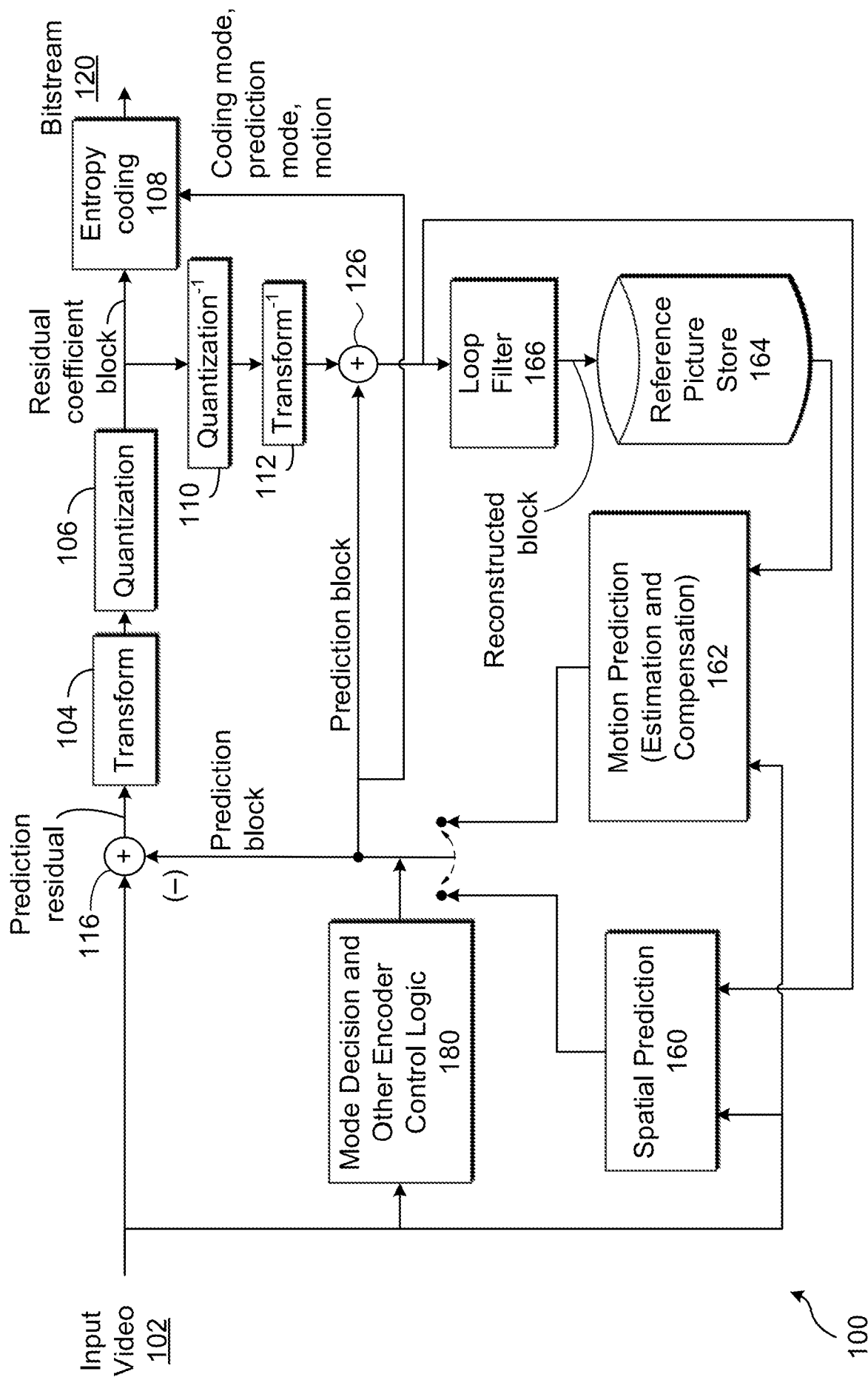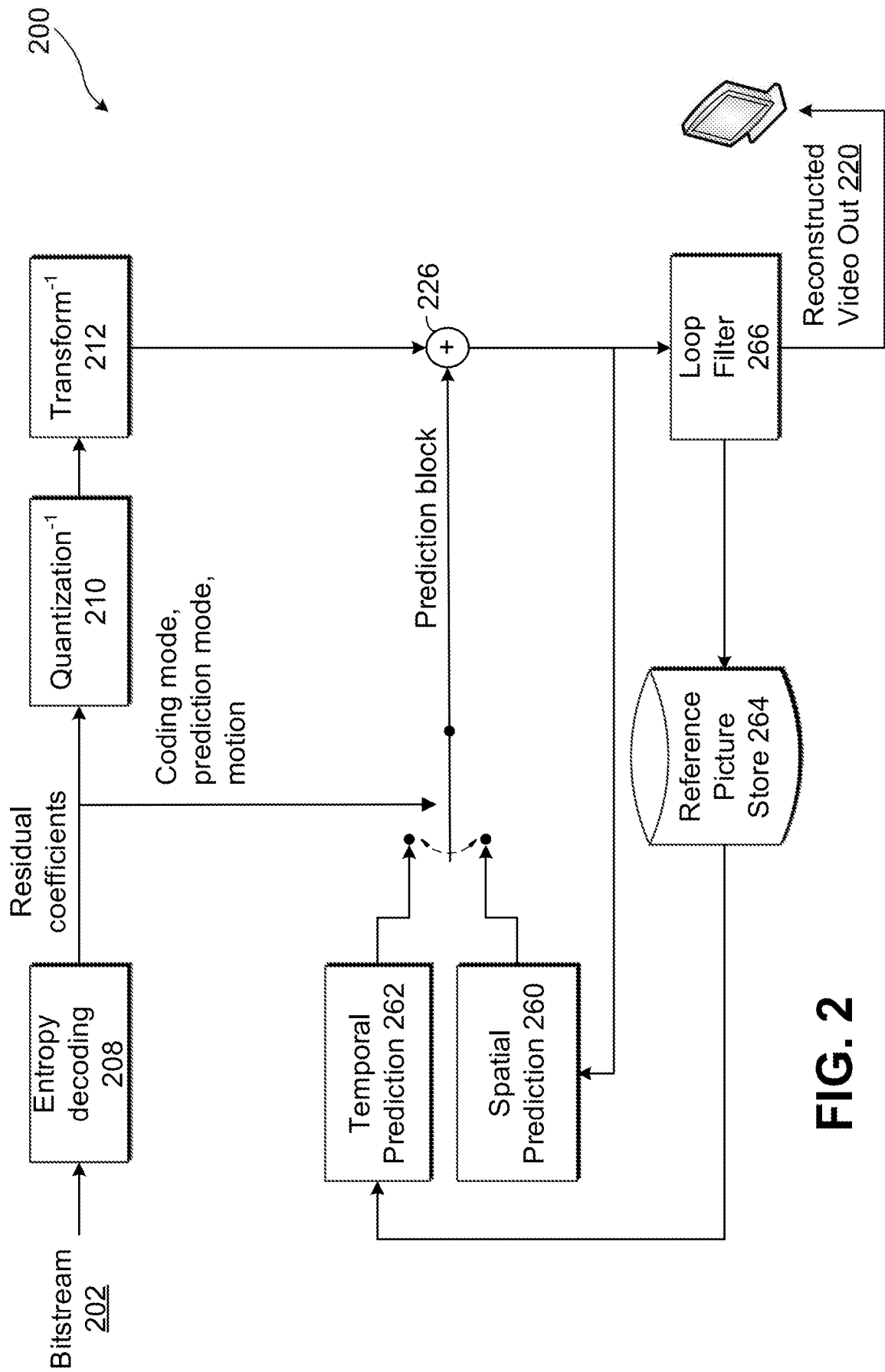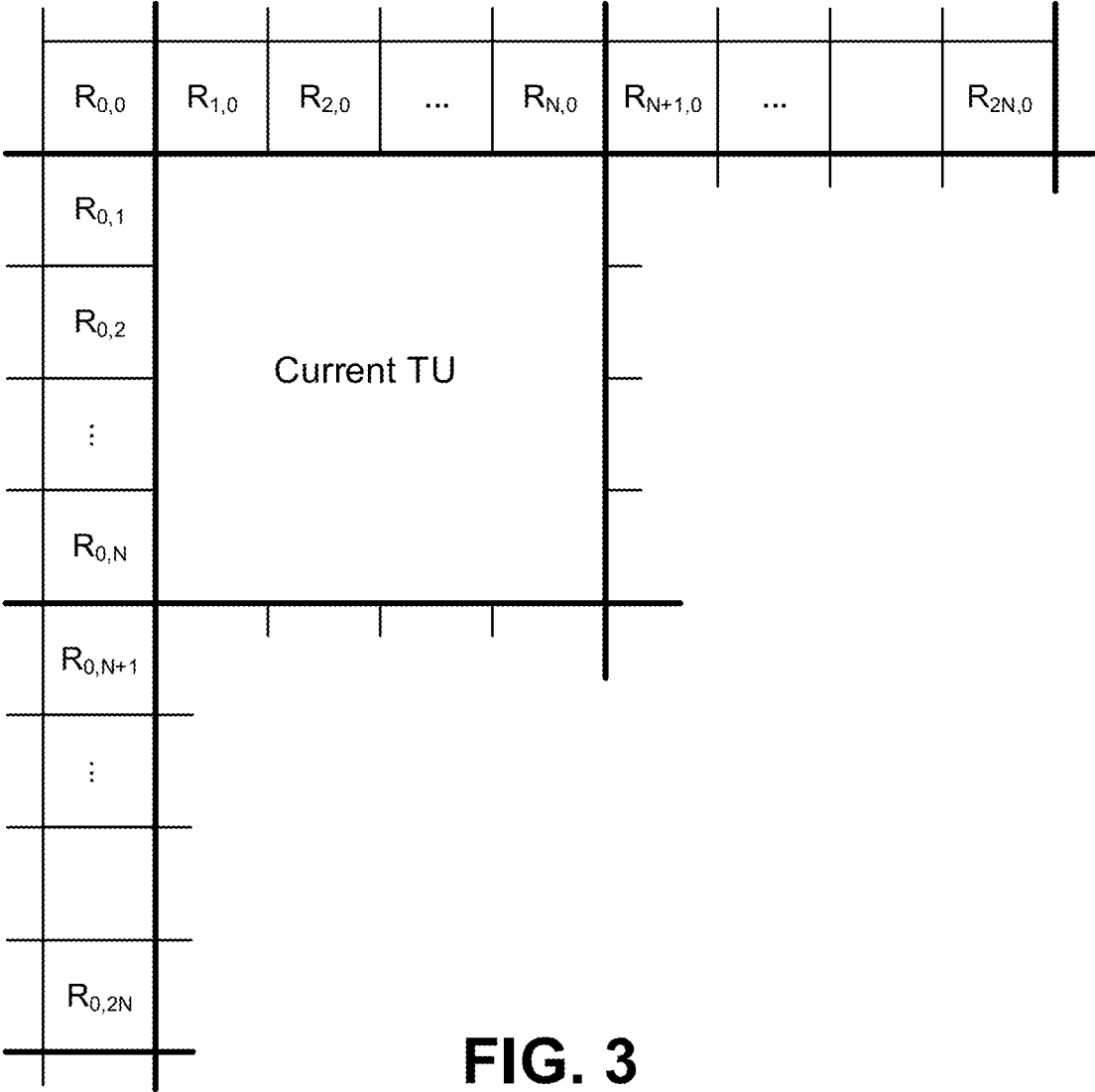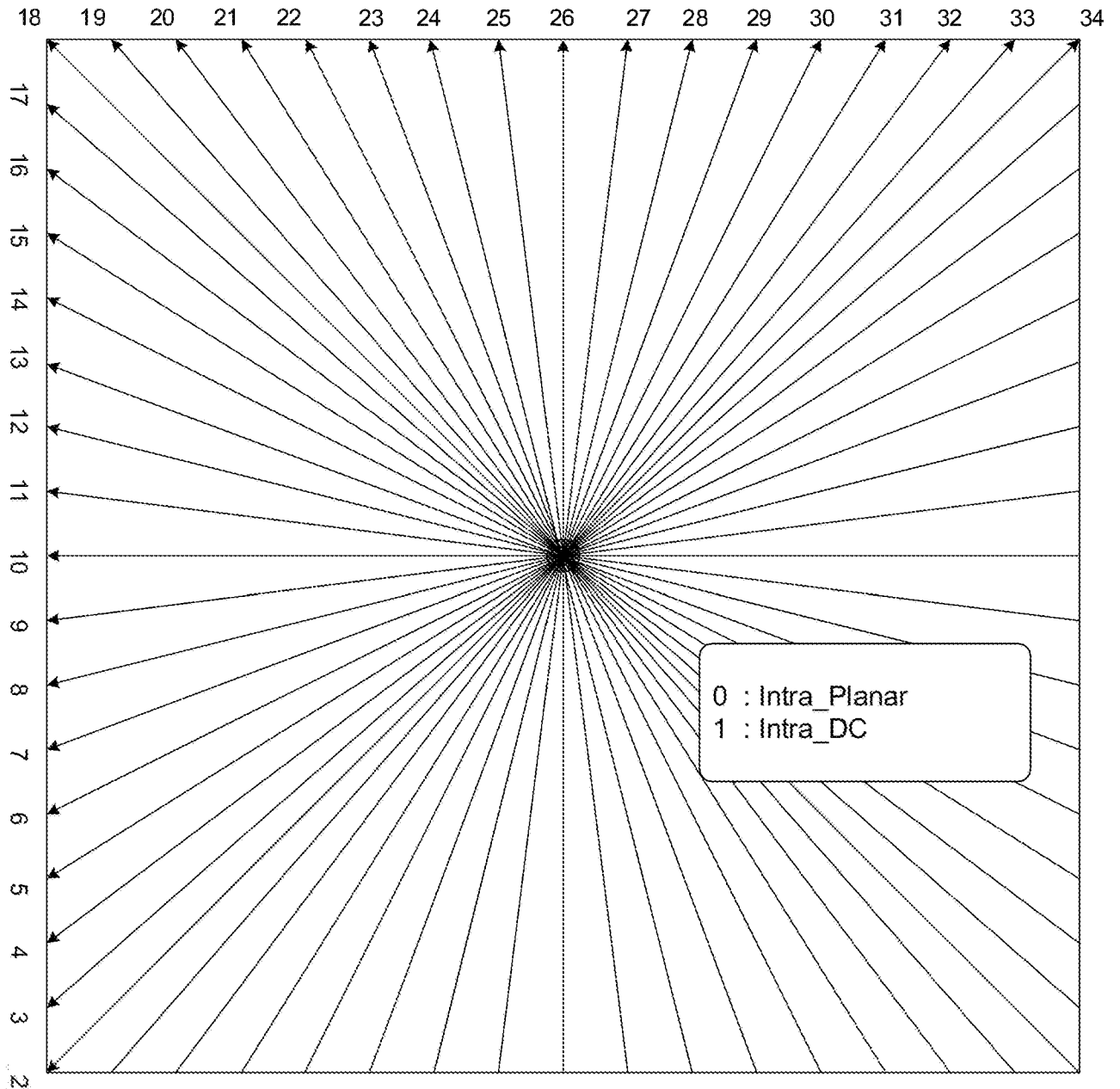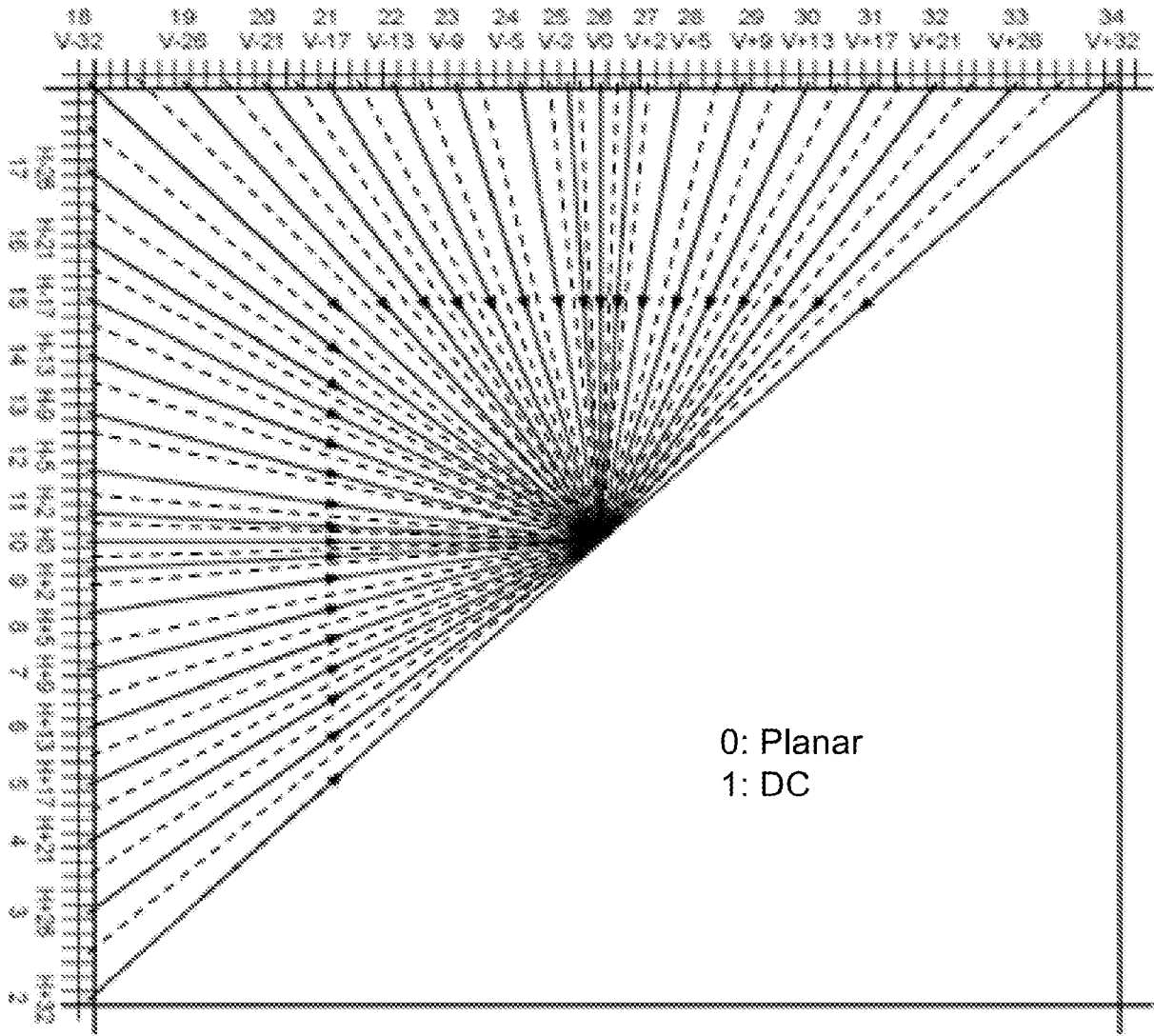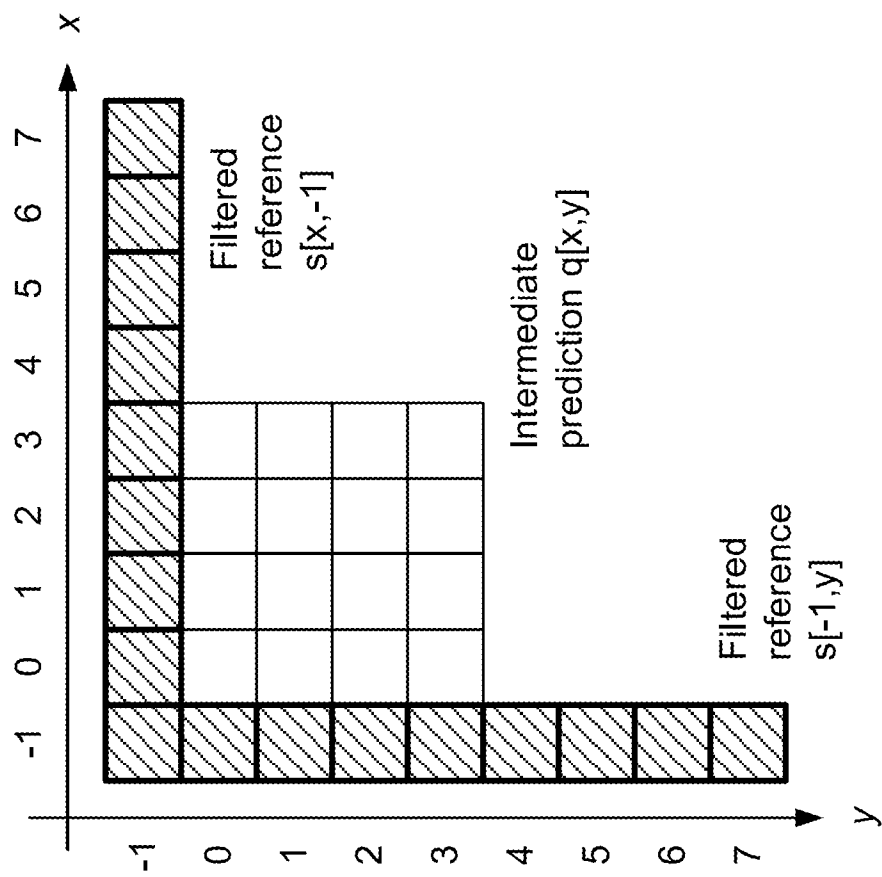
FIG. 6
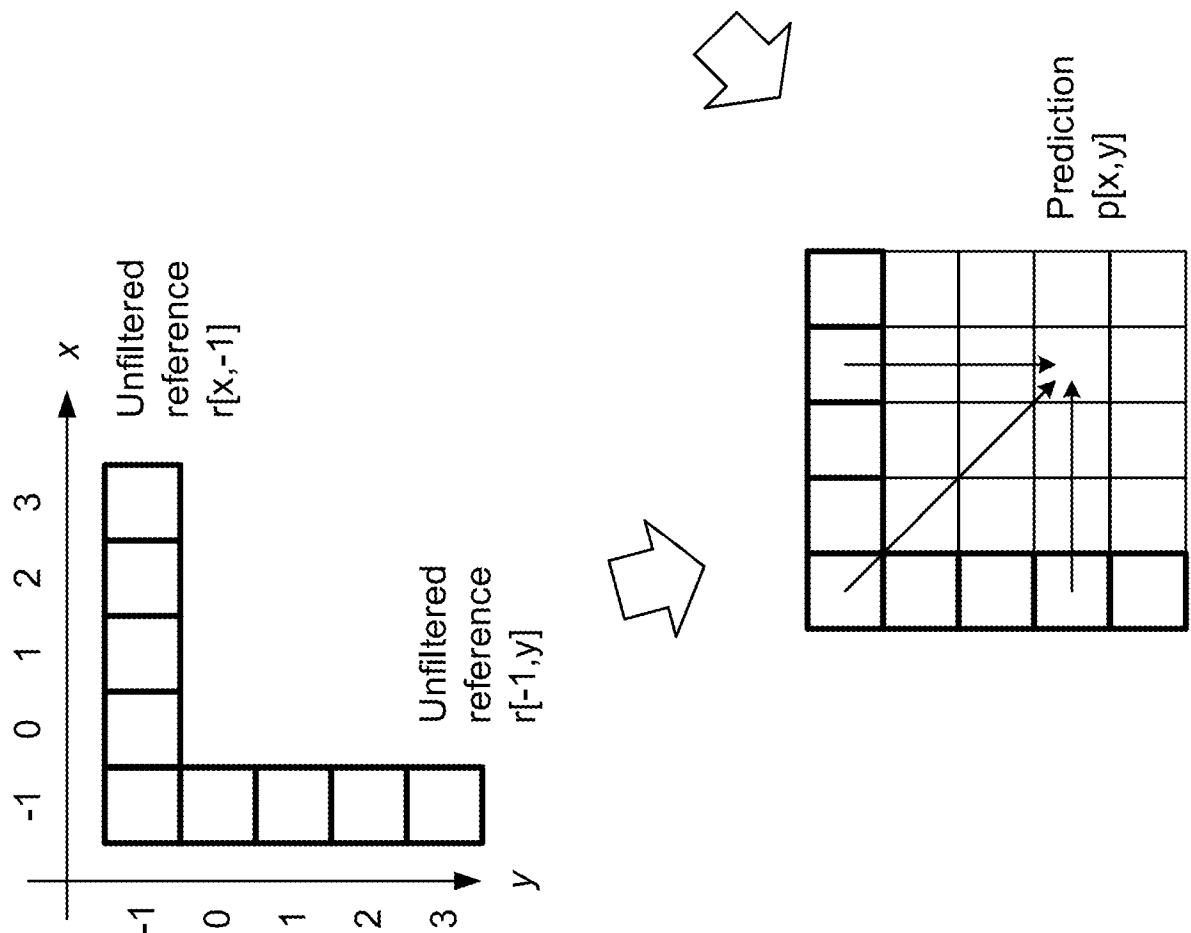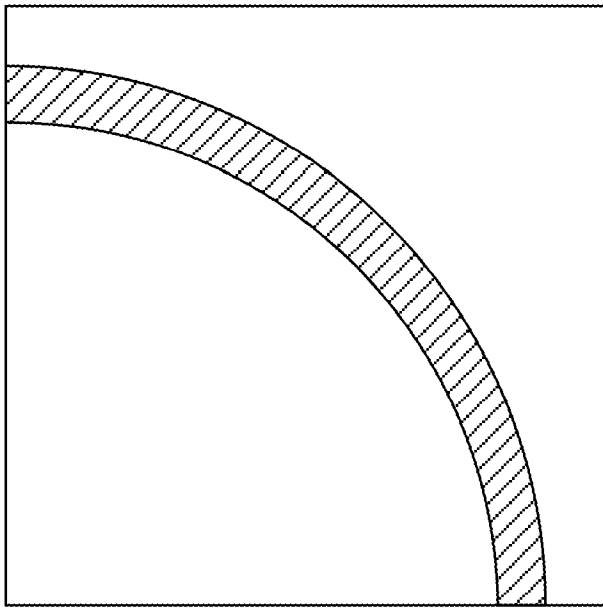
FIG. 7A

FIG. 7B

FIG. 7C

FIG. 7D

FIG. 8

FIG. 9

FIG. 10

$CP_0(0,0)$ $CP_1(N-1,0)$

(x,y)

x $N-1-x$

FIG. 11A

$CP_0(0,0)$

y

(x,y)

$N-1-y$

FIG. 11B

$CP_2(0,N-1)$

FIG. 11C

FIG. 11D

CP₀(0,0)   CP₁(N-1,0)

PU₀	PU₁
PU₂	PU₃

y

(x,y)

x

N-1-x

N-1-y

CP₂(0,N-1)

FIG. 12A

CP₀(0,0)   CP₁(N-1,0)

y

(x,y)

x   N-1-x

TU₀   TU₁

TU₂   TU₃

N-1-y

CP₂(0,N-1)

FIG. 12B

FIG. 13

FIG. 14

Bitstream 1000

VPS	SPS	PPS	Slice header	Slice data	SEI

1002 1003 1004 1005 1006   1007

NAL Unit | NAL Unit | NAL Unit | NAL Unit | NAL Unit | NAL Unit

1001

FIG. 15

1300

COMMUNICATION NETWORK
1304

1308   1310

| ENCODER 1302 | | DECODER 1306 |
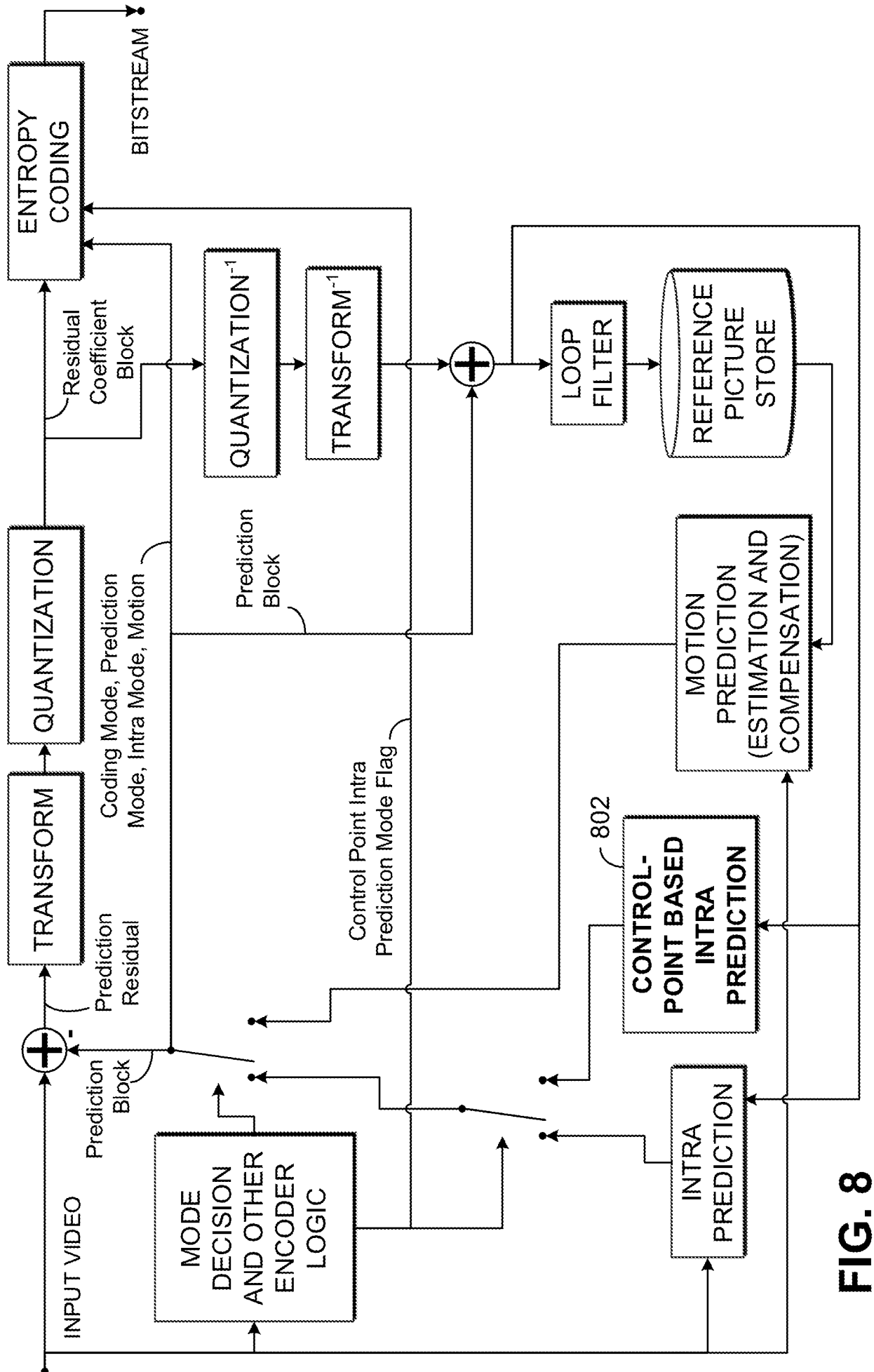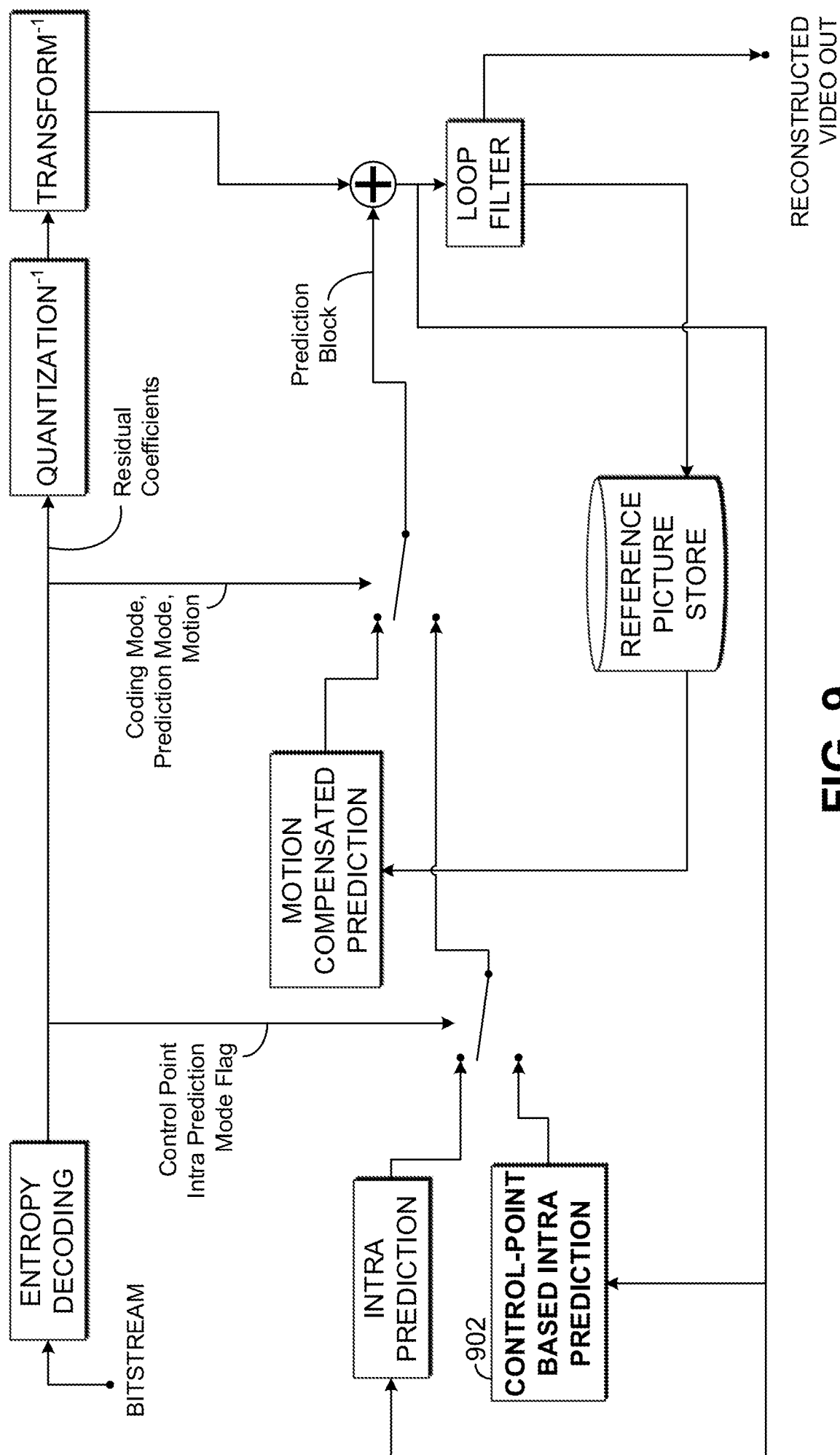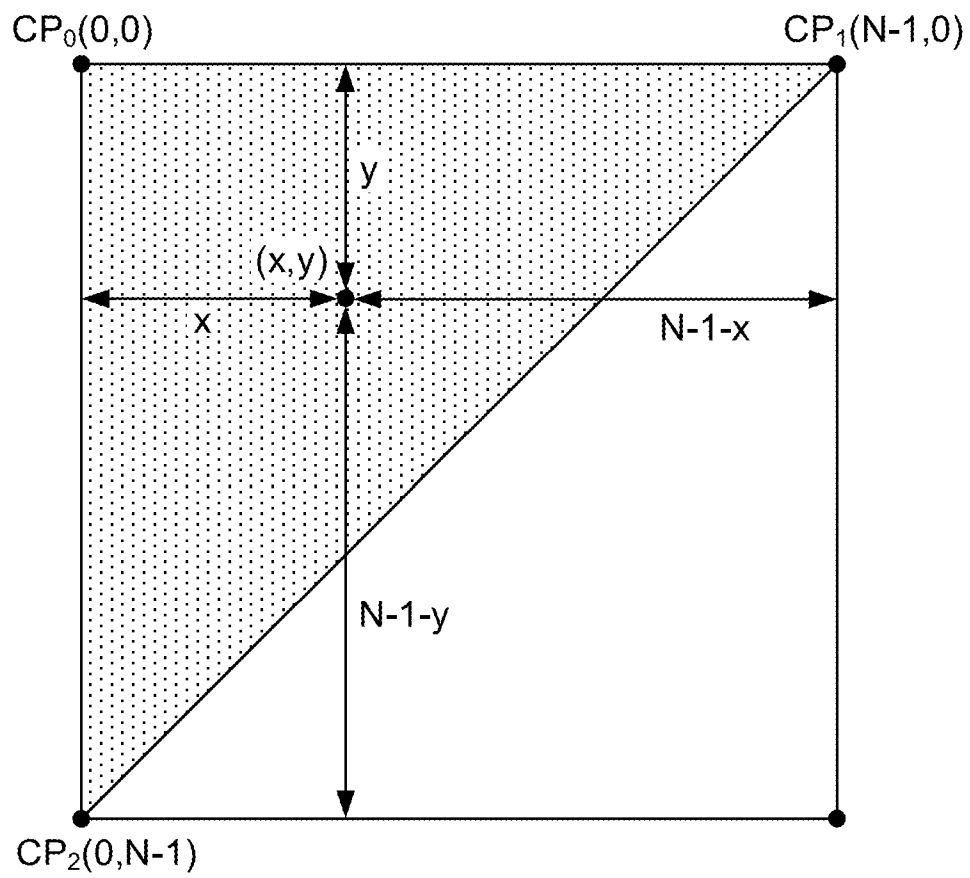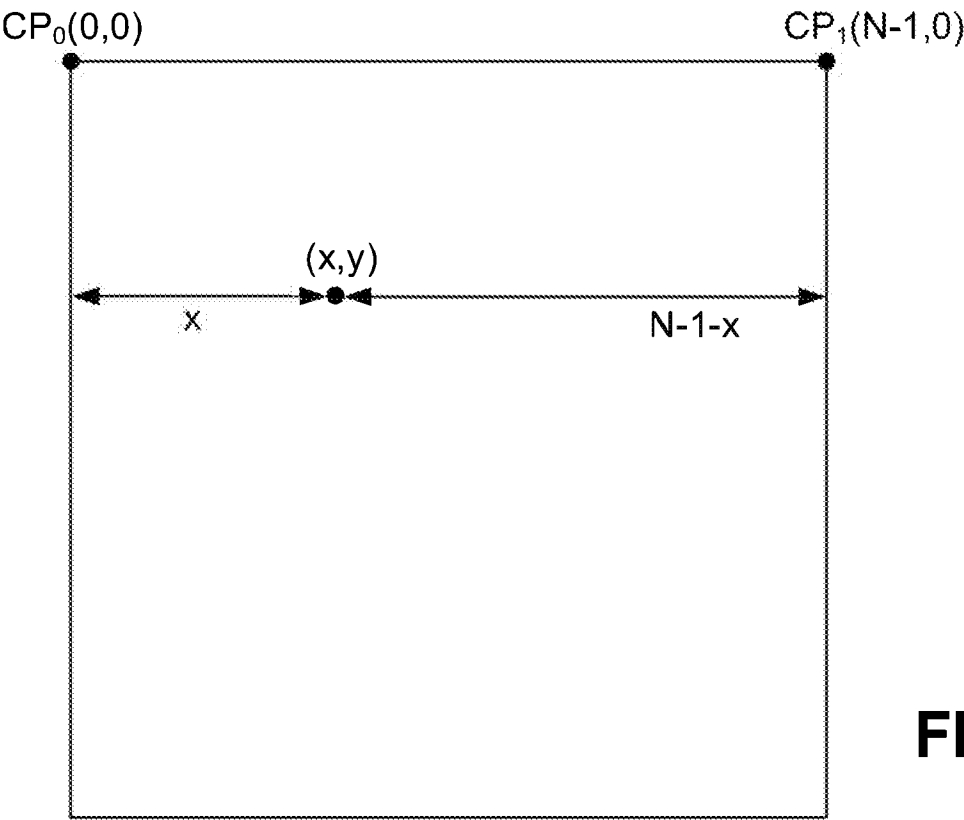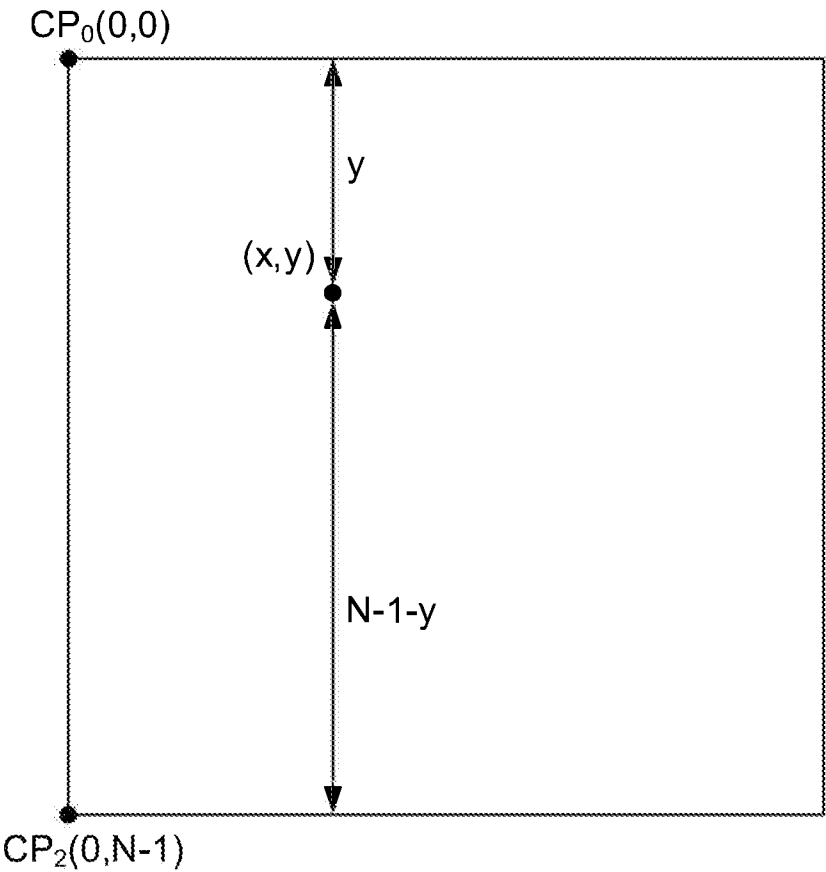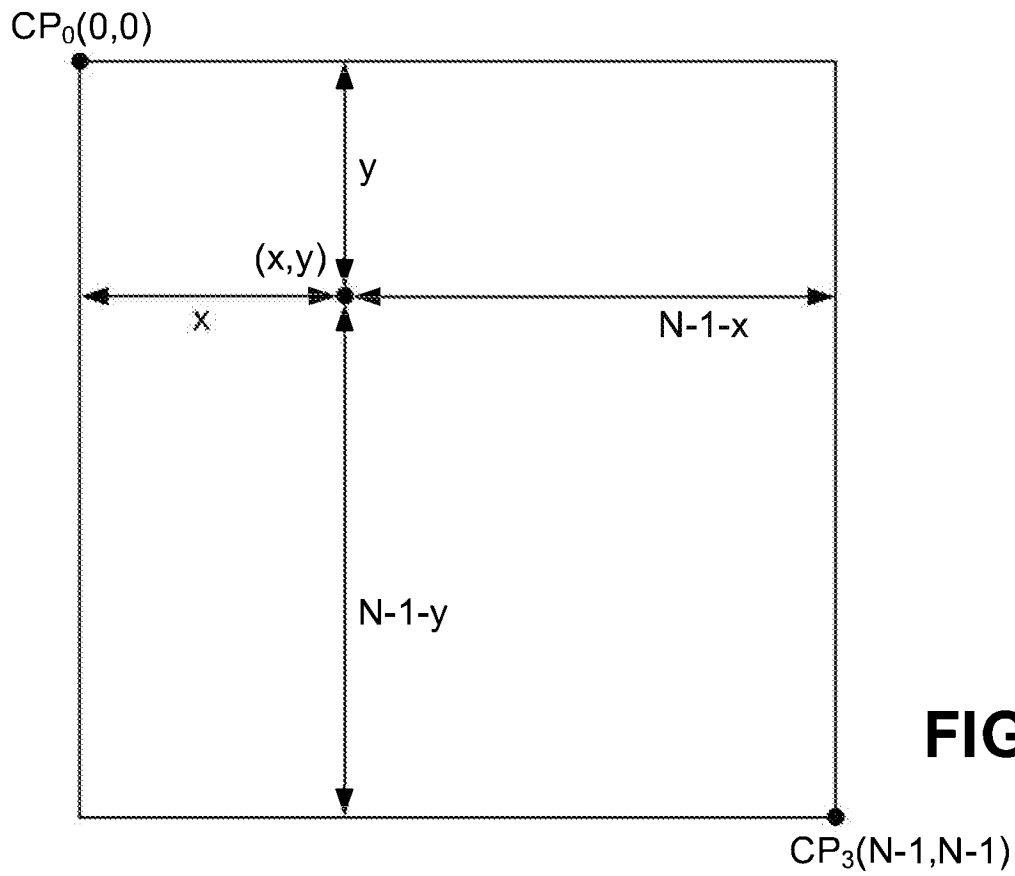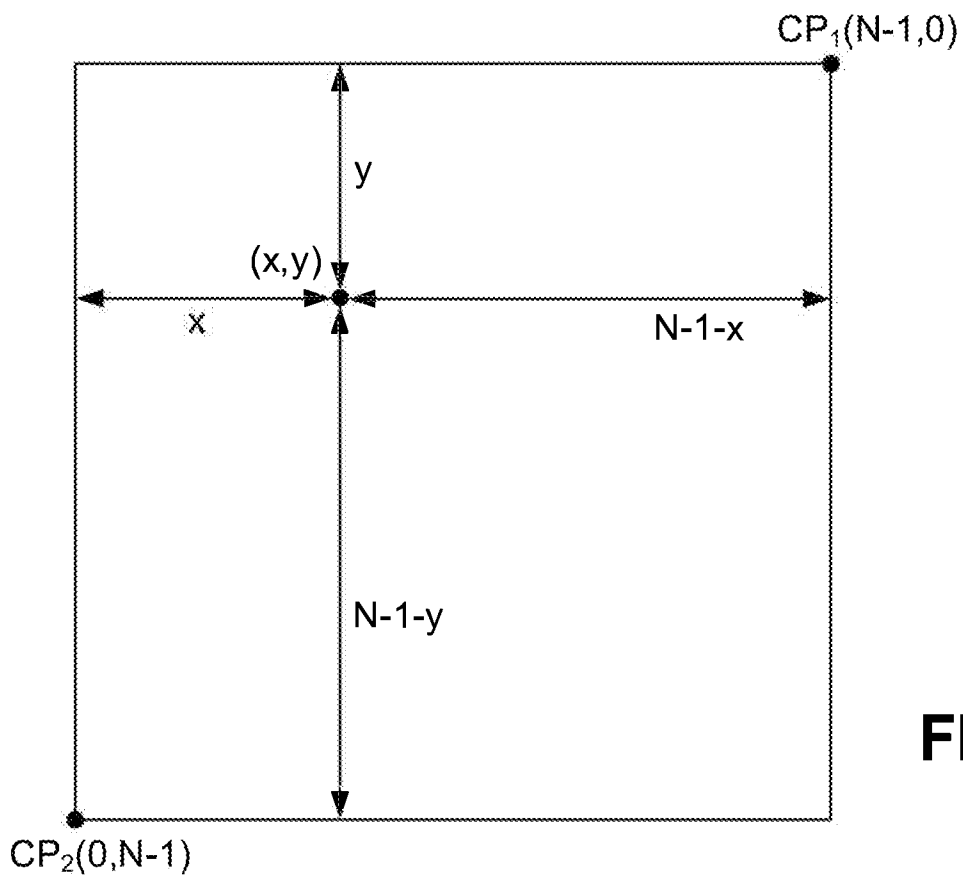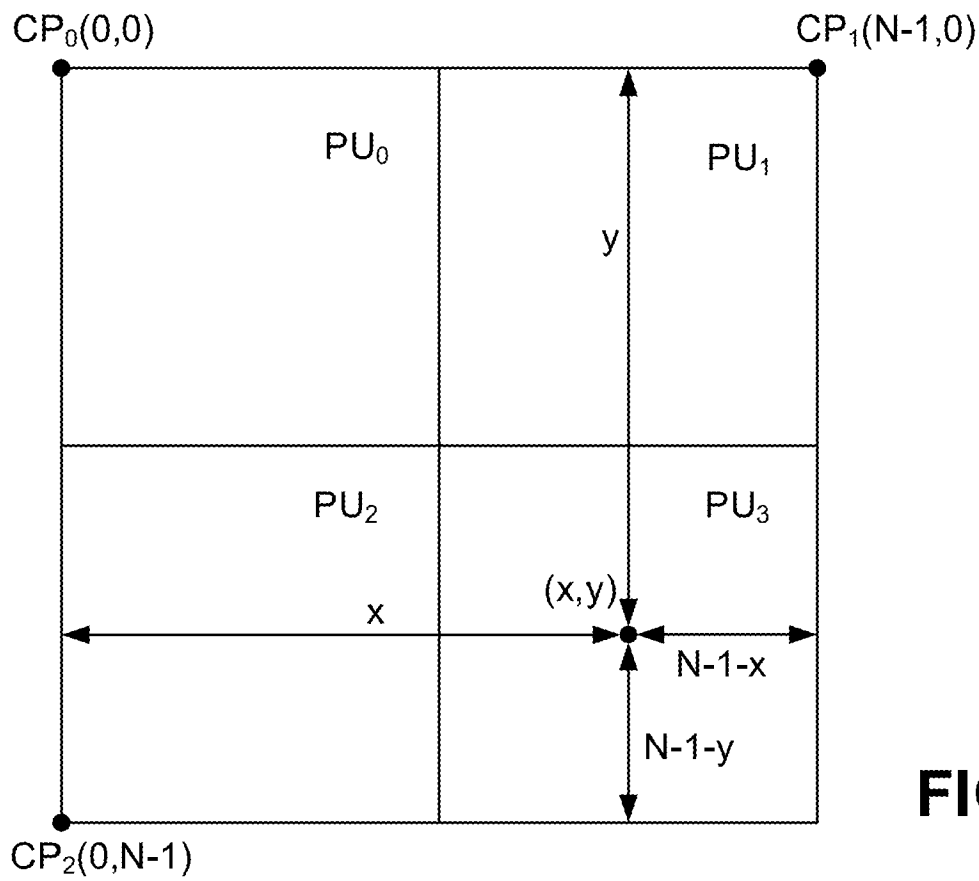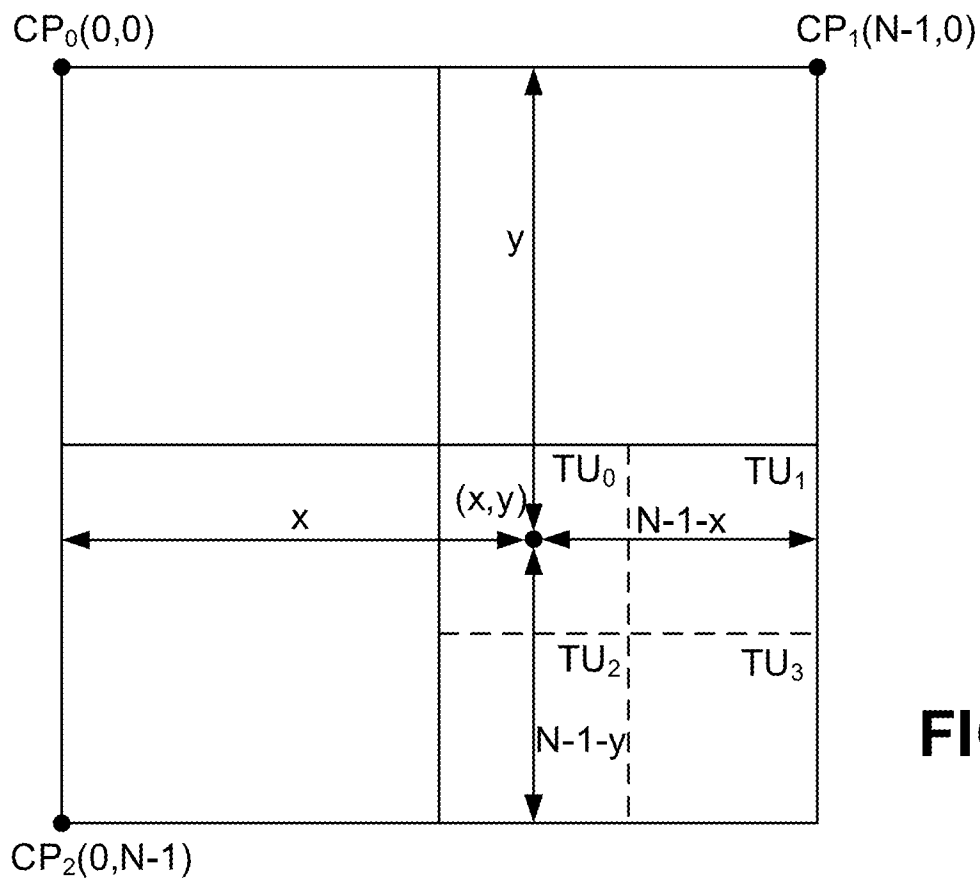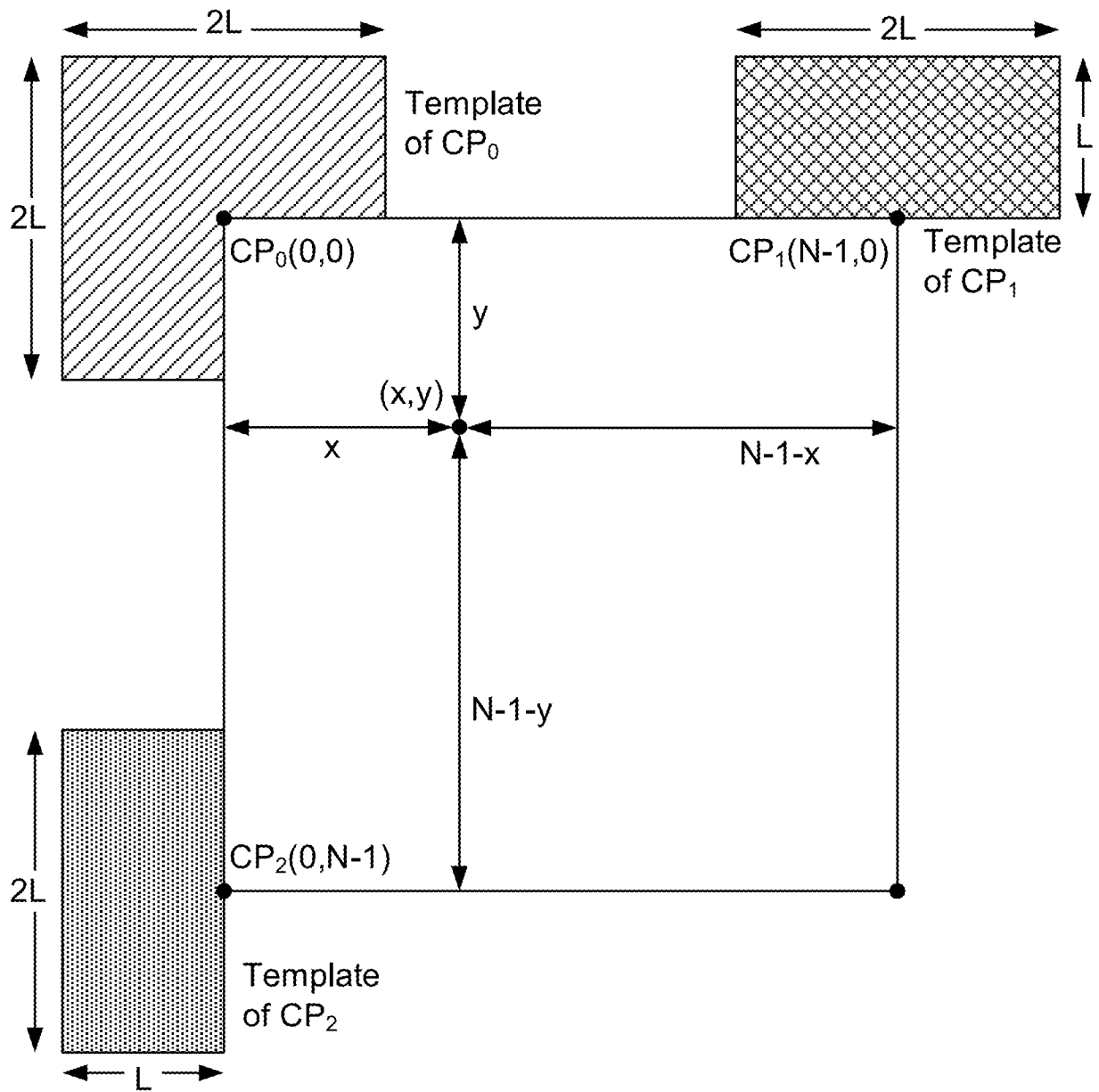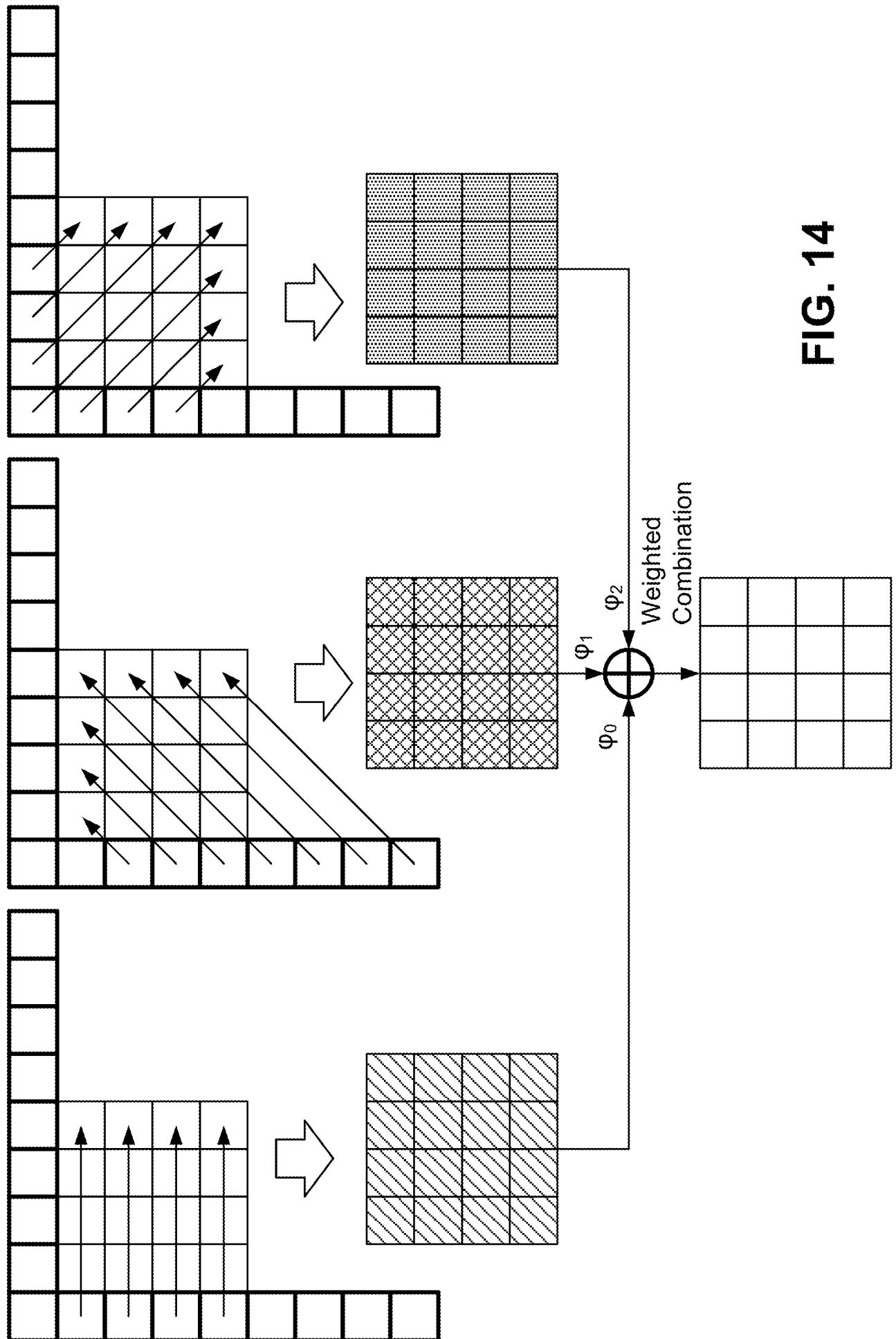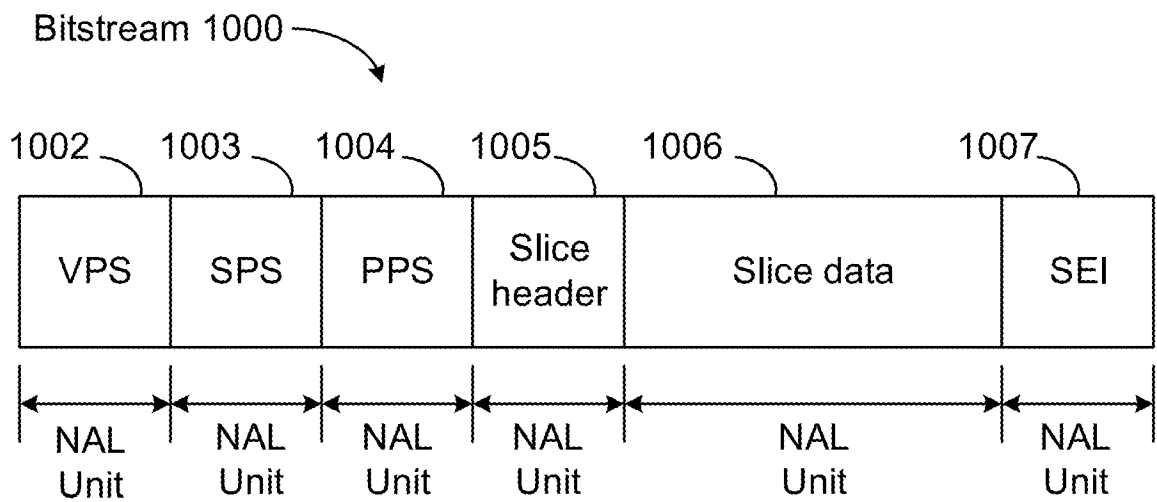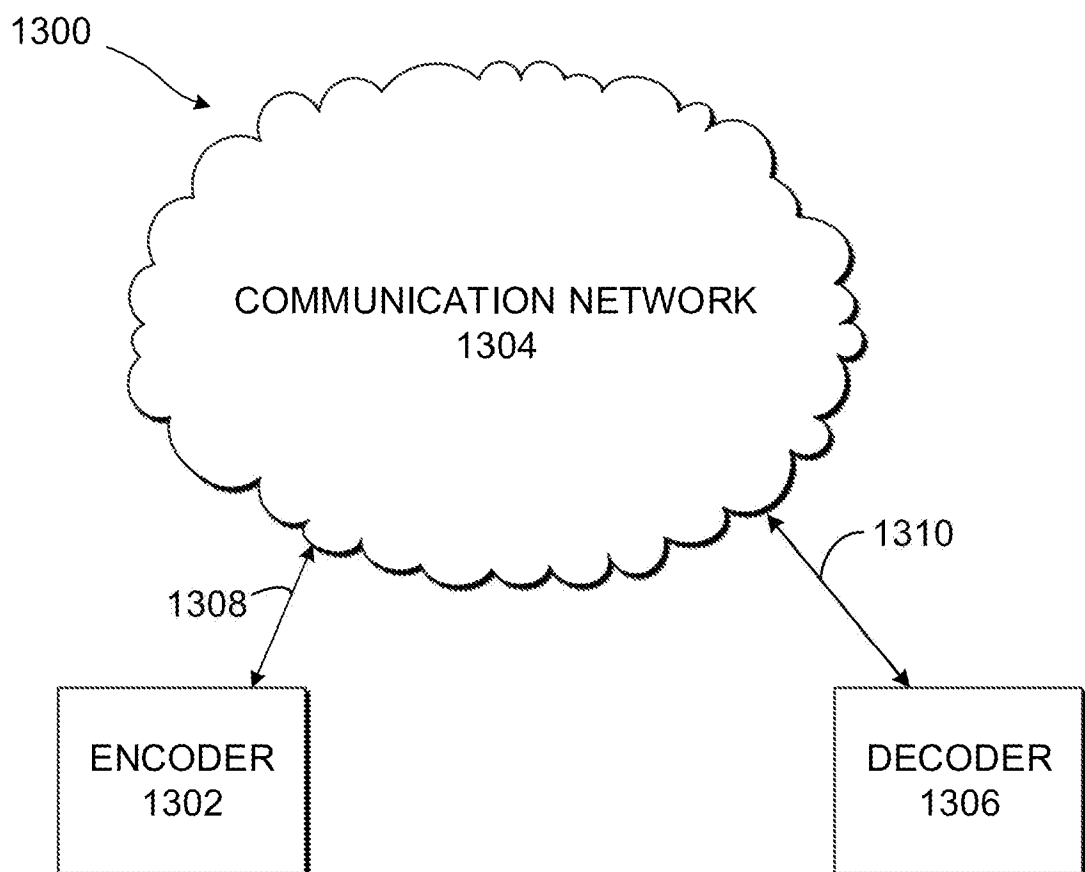
FIG. 16
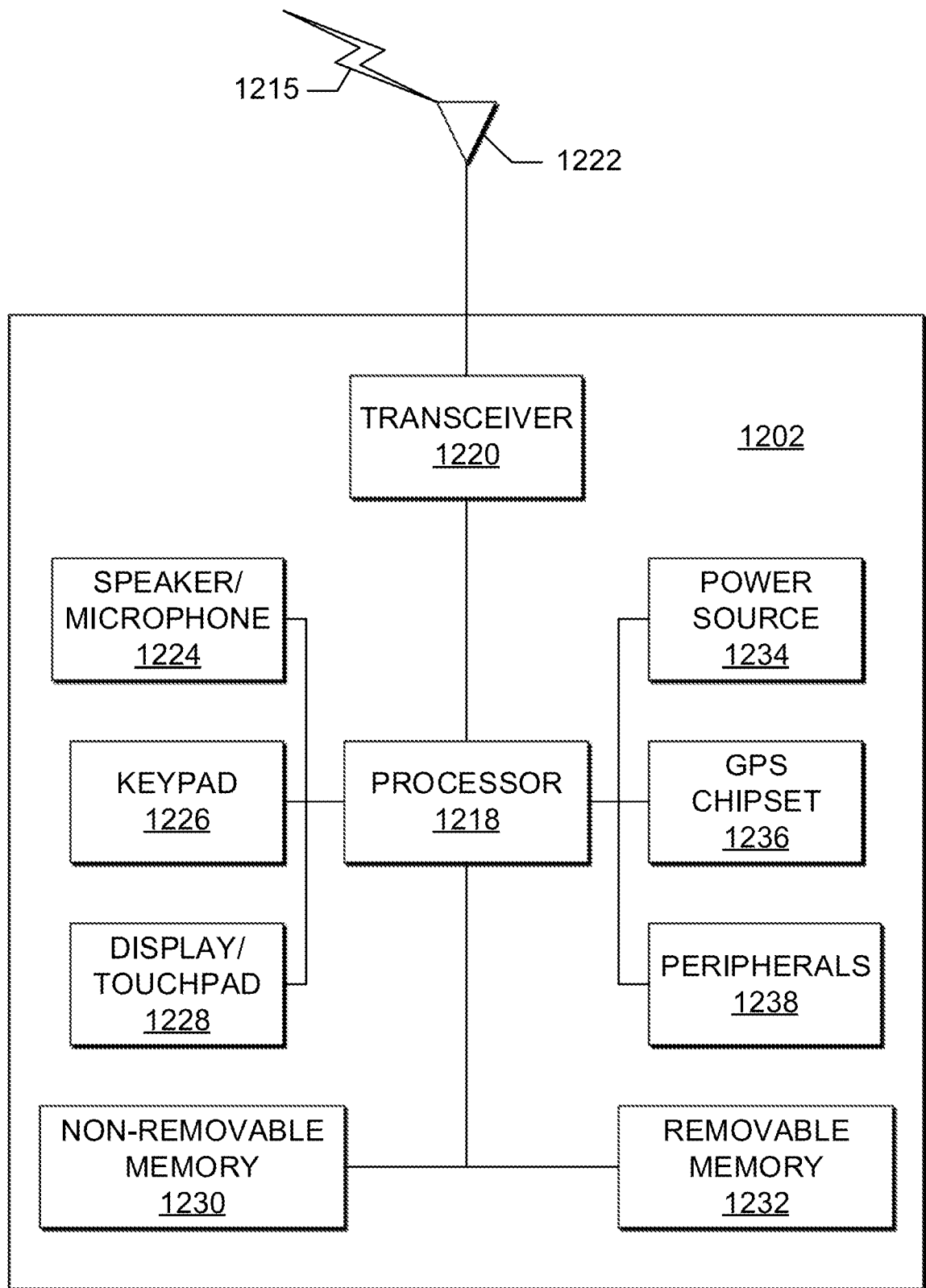
FIG. 17

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04N19/176   H04N19/147   H04N19/11   H04N19/463   H04N19/182
H04N19/59

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, COMPENDEX, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2011/083439 A1 (NOKIA CORP [FI]; LAINEMA JANI [FI]; UGUR KEMAL [FI]) 14 July 2011 (2011-07-14) page 13, line 4 - page 16, line 13; figures 5a, 5b, 9a,9b, 12a, 12b -----	1-20
A	JANI LAINEMA ET AL: "Intra picture coding with planar representations", 28TH PICTURE CODING SYMPOSIUM, 1 December 2010 (2010-12-01), pages 198-201, XP55100091, DOI: 10.1109/PCS.2010.5702460 ISBN: 978-1-42-447134-8 the whole document ----- -/--	1-20

[X] Further documents are listed in the continuation of Box C.   [X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 June 2017	04/07/2017

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Colesanti, Carlo

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	RAPPORTEUR Q6/16: "Future video coding: Joint Exploration Model 1 (JEM1) for future video coding investigation", ITU-T SG16 MEETING; 12-10-2015 - 23-10-2015; GENEVA,, no. T13-SG16-151012-TD-WP3-0213, 21 October 2015 (2015-10-21), XP030100762, Sectionb 2.2. with sub-sections -----	1-20

1

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 2011083439	A1	14-07-2011	CN	102783151 A	14-11-2012
			EP	2522144 A1	14-11-2012
			KR	20120112773 A	11-10-2012
			TW	201201589 A	01-01-2012
			US	2012121013 A1	17-05-2012
			WO	2011083439 A1	14-07-2011