

METHOD AND DEVICE FOR PROCESSING IMAGE INFORMATION FOR IMAGE/VIDEO CODING

BACKGROUND OF THE DISCLOSURE

5 Field of the disclosure

[1] The present technology relates to a method and an apparatus for processing image information in coding image/video.

Related Art

[2] Recently, the demand for high resolution, high quality image/video such as 4K, 8K or
10 more Ultra High Definition (UHD) image/video is increasing in various fields. As the image/video resolution or quality becomes higher, relatively more amount of information or bits are transmitted than for conventional image/video data. Therefore, if image/video data are transmitted via a medium such as an existing wired/wireless broadband line or stored in a legacy storage medium, costs for transmission and storage are readily increased.

15 **[3]** Moreover, interests and demand are growing for virtual reality (VR) and artificial reality (AR) contents, and immersive media such as hologram; and broadcasting of images/videos exhibiting image/video characteristics different from those of an actual image/video, such as game images/videos, are also growing.

[4] Therefore, a highly efficient image/video compression technique is required to
20 effectively compress and transmit, store, or play high resolution, high quality images/videos showing various characteristics as described above.

SUMMARY

[5] A technical subject of the present document is to provide a method and an apparatus
25 for enhancing image/video coding efficiency.

[6] Another technical subject of the present document is to provide a method and an apparatus for efficiently performing inter prediction and/or intra prediction in image/video coding.

[7] Still another technical subject of the present document is to provide a method and an apparatus for preventing unnecessary signaling in coding image/video.

[8] Still another technical subject of the present document is to provide a method and an apparatus for skipping unnecessary signaling for inter prediction and/or intra prediction in transferring image/video information.

[9] According to an embodiment of the present document, a video decoding method performed by a video decoding apparatus may include: obtaining image information from a bitstream, the image information including a picture header related to a current picture, and the current picture including a plurality of slices; parsing, from the picture header, at least one of a first flag representing whether information necessary for an inter prediction operation for a decoding process is present in the picture header or a second flag representing whether information necessary for an intra prediction operation for the decoding process is present in the picture header; generating prediction samples by performing at least one of an intra prediction or an inter prediction for the slices in the current picture based on at least one of the first flag or the second flag; and generating reconstructed samples based on the prediction samples.

[10] According to another embodiment of the present document, a video encoding method performed by a video encoding apparatus may include: determining a prediction mode of a current block in a current picture, the current picture including a plurality of slices; generating prediction samples for the current block based on the prediction mode; generating at least one of first information representing whether information necessary for an inter prediction operation for a decoding process is present in a picture header related to the current picture or

second information representing whether information necessary for an intra prediction operation for the decoding process is present in the picture header based on the prediction mode; and encoding image information including at least one of the first information or the second information, wherein the first information and the second information are included in the picture header of the image information.

[11] According to still another embodiment of the present document, a computer-readable digital storage medium including information causing a decoding apparatus to perform a video decoding method, wherein the decoding method may include: obtaining image information from a bitstream, the image information including a picture header related to a current picture, and the current picture including a plurality of slices; parsing, from the picture header, at least one of a first flag representing whether information necessary for an inter prediction operation for a decoding process is present in the picture header or a second flag representing whether information necessary for an intra prediction operation for the decoding process is present in the picture header; generating prediction samples by performing at least one of an intra prediction or an inter prediction for the slices in the current picture based on at least one of the first flag or the second flag; and generating reconstructed samples based on the prediction samples.

[12] According to an embodiment of the present document, the overall image/video compression efficiency can be enhanced.

[13] According to an embodiment of the present document, the inter prediction and/or the intra prediction can be efficiently performed during image/video coding.

[14] According to an embodiment of the present document, unnecessary signaling can be skipped during image/video coding.

[15] According to an embodiment of the present document, signaling of an unnecessary syntax element for the inter prediction or the intra prediction can be prevented during

image/video transmission.

BRIEF DESCRIPTION OF THE DRAWINGS

[16] FIG. 1 schematically shows an example of a video/image coding system to which
5 embodiments of the present disclosure may be applied.

[17] FIG. 2 is a diagram schematically illustrating a configuration of a video/image
encoding apparatus to which embodiments of the present disclosure may be applied.

[18] FIG. 3 is a diagram schematically illustrating a configuration of a video/image
decoding apparatus to which embodiments of the present disclosure may be applied.

10 [19] FIG. 4 shows an example of a video/image encoding method based on intra prediction.

[20] FIG. 5 schematically illustrates an intra predictor in an encoding apparatus.

[21] FIG. 6 illustrates an example of a video/image decoding method based on intra
prediction.

[22] FIG. 7 schematically illustrates an inter predictor in a decoding apparatus.

15 [23] FIG. 8 illustrates an example of a video/image encoding method based on inter
prediction.

[24] FIG. 9 schematically illustrates an inter predictor in an encoding apparatus.

[25] FIG. 10 illustrates an example of a video/image decoding method based on inter
prediction.

20 [26] FIG. 11 schematically illustrates an inter predictor in a decoding apparatus.

[27] FIGS. 12 and 13 schematically illustrate a video/image encoding method and an
example of related components according to an embodiment of the present document.

[28] FIGS. 14 and 15 schematically illustrate a video/image decoding method and an
example of related components according to an embodiment of the present document.

25 [29] FIG. 16 illustrates an example of a content streaming system to which embodiments

disclosed in the present document are applicable.

DESCRIPTION OF EXEMPLARY EMBODIMENTS

[30] The disclosure of the present disclosure may be modified in various forms, and specific
5 embodiments thereof will be described and illustrated in the drawings. The terms used in the
present disclosure are used to merely describe specific embodiments, but are not intended to
limit the disclosed method in the present disclosure. An expression of a singular number
includes an expression of 'at least one', so long as it is clearly read differently. The terms such
as “include” and “have” are intended to indicate that features, numbers, steps, operations,
10 elements, components, or combinations thereof used in the document exist and it should be
thus understood that the possibility of existence or addition of one or more different features,
numbers, steps, operations, elements, components, or combinations thereof is not excluded.

[31] In addition, each configuration of the drawings described in this document is an
independent illustration for explaining functions as features that are different from each other,
15 and does not mean that each configuration is implemented by mutually different hardware or
different software. For example, two or more of the configurations may be combined to form
one configuration, and one configuration may also be divided into multiple configurations.
Without departing from the gist of the disclosed method of the present disclosure, embodiments
in which configurations are combined and/or separated are included in the scope of the
20 disclosure of the present document.

[32] This document relates to video/image coding. For example, a method/embodiment
disclosed in this document may be applied to a method disclosed in a versatile video coding
(VVC) standard. In addition, the method/embodiment disclosed in this document may be
applied to a method disclosed in an essential video coding (EVC) standard, AOMedia Video 1
25 (AV1) standard, 2nd generation of audio video coding standard (AVS2), or a next-generation

video/image coding standard (e.g., H.267, H.268, etc.).

[33] Various embodiments related to video/image coding are presented in this document, and the embodiments may be combined with each other unless otherwise stated.

[34] In this document, a video may refer to a series of images over time. A picture generally refers to the unit representing one image at a particular time frame, and a slice/tile refers to the unit constituting a part of the picture in terms of coding. A slice/tile may include one or more coding tree units (CTUs). One picture may consist of one or more slices/tiles. One picture may consist of one or more tile groups. One tile group may include one or more tiles. A brick may represent a rectangular region of CTU rows within a tile in a picture). A tile may be partitioned into multiple bricks, each of which consisting of one or more CTU rows within the tile. A tile that is not partitioned into multiple bricks may be also referred to as a brick. A brick scan is a specific sequential ordering of CTUs partitioning a picture in which the CTUs are ordered consecutively in CTU raster scan in a brick, bricks within a tile are ordered consecutively in a raster scan of the bricks of the tile, and tiles in a picture are ordered consecutively in a raster scan of the tiles of the picture. A tile is a rectangular region of CTUs within a particular tile column and a particular tile row in a picture. The tile column is a rectangular region of CTUs having a height equal to the height of the picture and a width specified by syntax elements in the picture parameter set. The tile row is a rectangular region of CTUs having a height specified by syntax elements in the picture parameter set and a width equal to the width of the picture). A tile scan is a specific sequential ordering of CTUs partitioning a picture in which the CTUs are ordered consecutively in CTU raster scan in a tile whereas tiles in a picture are ordered consecutively in a raster scan of the tiles of the picture. A slice includes an integer number of bricks of a picture that may be exclusively contained in a single NAL unit. A slice may be composed of either a number of complete tiles or only a consecutive sequence of complete bricks of one tile. In this document, tile group and slice may be used interchangeably. For

example, in this document, a tile group/tile group header may be referred to as a slice/slice header.

[35] A pixel or a pel may mean a smallest unit constituting one picture (or image). Also, 'sample' may be used as a term corresponding to a pixel. A sample may generally represent a pixel or a value of a pixel, and may represent only a pixel/pixel value of a luma component or only a pixel/pixel value of a chroma component.

[36] A unit may represent a basic unit of image processing. The unit may include at least one of a specific region of the picture and information related to the region. One unit may include one luma block and two chroma (ex. Cb, cr) blocks. The unit may be used interchangeably with terms such as block or area in some cases. In a general case, an $M \times N$ block may include samples (or sample arrays) or a set (or array) of transform coefficients of M columns and N rows. Alternatively, the sample may mean a pixel value in the spatial domain, and when such a pixel value is transformed to the frequency domain, it may mean a transform coefficient in the frequency domain.

[37] In this document, the term "/" and "," should be interpreted to indicate "and/or." For instance, the expression "A/B" may mean "A and/or B." Further, "A, B" may mean "A and/or B." Further, "A/B/C" may mean "at least one of A, B, and/or C." Also, "A/B/C" may mean "at least one of A, B, and/or C."

[38] "Further, in the document, the term "or" should be interpreted to indicate "and/or." For instance, the expression "A or B" may comprise 1) only A, 2) only B, and/or 3) both A and B. In other words, the term "or" in this document should be interpreted to indicate "additionally or alternatively."

[39] Further, the parentheses used in the present specification may mean "for example". Specifically, in the case that "prediction (intra prediction)" is expressed, it may be indicated that "intra prediction" is proposed as an example of "prediction". In other words, the term

“prediction” in the present specification is not limited to “intra prediction”, and it may be indicated that “intra prediction” is proposed as an example of “prediction”. Further, even in the case that “prediction (i.e., intra prediction)” is expressed, it may be indicated that “intra prediction” is proposed as an example of “prediction”.

5 **[40]** Hereinafter, embodiments of the present disclosure will be described in detail with reference to the accompanying drawings. In addition, like reference numerals are used to indicate like elements throughout the drawings, and the same descriptions on the like elements may be omitted.

10 **[41]** FIG. 1 illustrates an example of a video/image coding system to which the embodiments of the present disclosure may be applied.

[42] Referring to FIG. 1, a video/image coding system may include a first device (a source device) and a second device (a reception device). The source device may transmit encoded video/image information or data to the reception device through a digital storage medium or network in the form of a file or streaming.

15 **[43]** The source device may include a video source, an encoding apparatus, and a transmitter. The receiving device may include a receiver, a decoding apparatus, and a renderer. The encoding apparatus may be called a video/image encoding apparatus, and the decoding apparatus may be called a video/image decoding apparatus. The transmitter may be included in the encoding apparatus. The receiver may be included in the decoding apparatus. The
20 renderer may include a display, and the display may be configured as a separate device or an external component.

[44] The video source may acquire video/image through a process of capturing, synthesizing, or generating the video/image. The video source may include a video/image capture device and/or a video/image generating device. The video/image capture device may
25 include, for example, one or more cameras, video/image archives including previously

captured video/images, and the like. The video/image generating device may include, for example, computers, tablets and smartphones, and may (electronically) generate video/images. For example, a virtual video/image may be generated through a computer or the like. In this case, the video/image capturing process may be replaced by a process of generating related
5 data.

[45] The encoding apparatus may encode input video/image. The encoding apparatus may perform a series of procedures such as prediction, transform, and quantization for compaction and coding efficiency. The encoded data (encoded video/image information) may be output in the form of a bitstream.

10 **[46]** The transmitter may transmit the encoded image/image information or data output in the form of a bitstream to the receiver of the receiving device through a digital storage medium or a network in the form of a file or streaming. The digital storage medium may include various storage mediums such as USB, SD, CD, DVD, Blu-ray, HDD, SSD, and the like. The transmitter may include an element for generating a media file through a predetermined file
15 format and may include an element for transmission through a broadcast/communication network. The receiver may receive/extract the bitstream and transmit the received bitstream to the decoding apparatus.

[47] The decoding apparatus may decode the video/image by performing a series of procedures such as dequantization, inverse transform, and prediction corresponding to the
20 operation of the encoding apparatus.

[48] The renderer may render the decoded video/image. The rendered video/image may be displayed through the display.

[49] In the present specification, technical features individually explained in one drawing may be individually implemented, or may be simultaneously implemented.

25 **[50]** FIG. 2 is a diagram schematically illustrating the configuration of a video/image

encoding apparatus to which the embodiments of the present disclosure may be applied. Hereinafter, what is referred to as the video encoding apparatus may include an image encoding apparatus.

[51] Referring to FIG. 2, the encoding apparatus 200 may include and be configured with an image partitioner 210, a predictor 220, a residual processor 230, an entropy encoder 240, an adder 250, a filter 260, and a memory 270. The predictor 220 may include an inter predictor 221 and an intra predictor 222. The residual processor 230 may include a transformer 232, a quantizer 233, a dequantizer 234, and an inverse transformer 235. The residual processor 230 may further include a subtractor 231. The adder 250 may be called a reconstructor or reconstructed block generator. The image partitioner 210, the predictor 220, the residual processor 230, the entropy encoder 240, the adder 250, and the filter 260, which have been described above, may be configured by one or more hardware components (e.g., encoder chipsets or processors) according to an embodiment. In addition, the memory 270 may include a decoded picture buffer (DPB), and may also be configured by a digital storage medium. The hardware component may further include the memory 270 as an internal/external component.

[52] The image partitioner 210 may split an input image (or, picture, frame) input to the encoding apparatus 200 into one or more processing units. As an example, the processing unit may be called a coding unit (CU). In this case, the coding unit may be recursively split according to a Quad-tree binary-tree ternary-tree (QTBTBT) structure from a coding tree unit (CTU) or the largest coding unit (LCU). For example, one coding unit may be split into a plurality of coding units of a deeper depth based on a quad-tree structure, a binary-tree structure, and/or a ternary-tree structure. In this case, for example, the quad-tree structure is first applied and the binary-tree structure and/or the ternary-tree structure may be later applied. Alternatively, the binary-tree structure may also be first applied. A coding procedure according to the present disclosure may be performed based on a final coding unit which is not split any

more. In this case, based on coding efficiency according to image characteristics or the like, the maximum coding unit may be directly used as the final coding unit, or as necessary, the coding unit may be recursively split into coding units of a deeper depth, such that a coding unit having an optimal size may be used as the final coding unit. Here, the coding procedure may include a procedure such as prediction, transform, and reconstruction to be described later. As another example, the processing unit may further include a prediction unit (PU) or a transform unit (TU). In this case, each of the prediction unit and the transform unit may be split or partitioned from the aforementioned final coding unit. The prediction unit may be a unit of sample prediction, and the transform unit may be a unit for inducing a transform coefficient and/or a unit for inducing a residual signal from the transform coefficient.

[53] The unit may be interchangeably used with the term such as a block or an area in some cases. Generally, an $M \times N$ block may represent samples composed of M columns and N rows or a group of transform coefficients. The sample may generally represent a pixel or a value of the pixel, and may also represent only the pixel/pixel value of a luma component, and also represent only the pixel/pixel value of a chroma component. The sample may be used as the term corresponding to a pixel or a pel configuring one picture (or image).

[54] The encoding apparatus 200 may subtract the prediction signal (predicted block, prediction sample array) output from the inter predictor 221 or the intra predictor 222 from the input image signal (original block, original sample array) to generate a residual signal (residual block, residual sample array), and the generated residual signal is transmitted to the transformer 232. In this case, as illustrated, a unit for subtracting the prediction signal (prediction block, prediction sample array) from an input image signal (original block, original sample array) in the encoder 200 may be referred to as a subtractor 231. The predictor 220 may perform prediction on a processing target block (hereinafter, referred to as a current block) and generate a predicted block including prediction samples for the current block. The predictor 220 may

determine whether intra prediction or inter prediction is applied in units of a current block or CU. The predictor 220 may generate various information on prediction, such as prediction mode information, and transmit the generated information to the entropy encoder 240, as is described below in the description of each prediction mode. The information on prediction may
5 be encoded by the entropy encoder 240 and output in the form of a bitstream.

[55] The intra predictor 222 may predict a current block with reference to samples within a current picture. The referenced samples may be located neighboring to the current block, or may also be located away from the current block according to the prediction mode. The prediction modes in the intra prediction may include a plurality of non-directional modes and
10 a plurality of directional modes. The non-directional mode may include, for example, a DC mode or a planar mode. The directional mode may include, for example, 33 directional prediction modes or 65 directional prediction modes according to the fine degree of the prediction direction. However, this is illustrative and the directional prediction modes which are more or less than the above number may be used according to the setting. The intra predictor
15 222 may also determine the prediction mode applied to the current block using the prediction mode applied to the neighboring block.

[56] The inter predictor 221 may induce a predicted block of the current block based on a reference block (reference sample array) specified by a motion vector on a reference picture. At this time, in order to decrease the amount of motion information transmitted in the inter
20 prediction mode, the motion information may be predicted in units of a block, a sub-block, or a sample based on the correlation of the motion information between the neighboring block and the current block. The motion information may include a motion vector and a reference picture index. The motion information may further include inter prediction direction (L0 prediction, L1 prediction, Bi prediction, or the like) information. In the case of the inter
25 prediction, the neighboring block may include a spatial neighboring block existing within the

current picture and a temporal neighboring block existing in the reference picture. The reference picture including the reference block and the reference picture including the temporal neighboring block may also be the same as each other, and may also be different from each other. The temporal neighboring block may be called the name such as a collocated reference block, a collocated CU (colCU), or the like, and the reference picture including the temporal neighboring block may also be called a collocated picture (colPic). For example, the inter predictor 221 may configure a motion information candidate list based on the neighboring blocks, and generate information indicating what candidate is used to derive the motion vector and/or the reference picture index of the current block. The inter prediction may be performed based on various prediction modes, and for example, in the case of a skip mode and a merge mode, the inter predictor 221 may use the motion information of the neighboring block as the motion information of the current block. In the case of the skip mode, the residual signal may not be transmitted unlike the merge mode. A motion vector prediction (MVP) mode may indicate the motion vector of the current block by using the motion vector of the neighboring block as a motion vector predictor, and signaling a motion vector difference.

[57] The predictor 220 may generate a prediction signal based on various prediction methods to be described below. For example, the predictor 220 may apply intra prediction or inter prediction for prediction of one block and may simultaneously apply intra prediction and inter prediction. This may be called combined inter and intra prediction (CIIP). In addition, the predictor may be based on an intra block copy (IBC) prediction mode or based on a palette mode for prediction of a block. The IBC prediction mode or the palette mode may be used for image/video coding of content such as games, for example, screen content coding (SCC). IBC basically performs prediction within the current picture, but may be performed similarly to inter prediction in that a reference block is derived within the current picture. That is, IBC may use at least one of the inter prediction techniques described in this document. The palette mode

may be viewed as an example of intra coding or intra prediction. When the palette mode is applied, a sample value in the picture may be signaled based on information on the palette table and the palette index.

[58] The prediction signal generated by the predictor (including the inter predictor 221 and/or the intra predictor 222) may be used to generate a reconstructed signal or may be used to generate a residual signal.

[59] The transformer 232 may generate transform coefficients by applying a transform technique to the residual signal. For example, the transform technique may include at least one of a discrete cosine transform (DCT), a discrete sine transform (DST), a graph-based transform (GBT), or a conditionally non-linear transform (CNT). Here, GBT refers to transformation obtained from a graph when expressing relationship information between pixels in the graph. CNT refers to transformation obtained based on a prediction signal generated using all previously reconstructed pixels. Also, the transformation process may be applied to a block of pixels having the same size as a square or may be applied to a block of a variable size that is not a square.

[60] The quantizer 233 quantizes the transform coefficients and transmits the same to the entropy encoder 240, and the entropy encoder 240 encodes the quantized signal (information on the quantized transform coefficients) and outputs the encoded signal as a bitstream. Information on the quantized transform coefficients may be referred to as residual information.

The quantizer 233 may rearrange the quantized transform coefficients in the block form into a one-dimensional vector form based on a coefficient scan order and may generate information on the transform coefficients based on the quantized transform coefficients in the one-dimensional vector form.

[61] The entropy encoder 240 may perform various encoding methods such as, for example, exponential Golomb, context-adaptive variable length coding (CAVLC), and context-adaptive

binary arithmetic coding (CABAC). The entropy encoder 240 may encode information necessary for video/image reconstruction (e.g., values of syntax elements, etc.) other than the quantized transform coefficients together or separately. Encoded information (e.g., encoded video/image information) may be transmitted or stored in units of a network abstraction layer (NAL) unit in the form of a bitstream. The video/image information may further include information on various parameter sets, such as an adaptation parameter set (APS), a picture parameter set (PPS), a sequence parameter set (SPS), or a video parameter set (VPS). Also, the video/image information may further include general constraint information. In this document, information and/or syntax elements transmitted/signaled from the encoding apparatus to the decoding apparatus may be included in video/image information. The video/image information may be encoded through the encoding procedure described above and included in the bitstream. The bitstream may be transmitted through a network or may be stored in a digital storage medium. Here, the network may include a broadcasting network and/or a communication network, and the digital storage medium may include various storage media such as USB, SD, CD, DVD, Blu-ray, HDD, and SSD. A transmitting unit (not shown) and/or a storing unit (not shown) for transmitting or storing a signal output from the entropy encoder 240 may be configured as internal/external elements of the encoding apparatus 200, or the transmitting unit may be included in the entropy encoder 240.

[62] The quantized transform coefficients output from the quantizer 233 may be used to generate a prediction signal. For example, the residual signal (residual block or residual samples) may be reconstructed by applying dequantization and inverse transform to the quantized transform coefficients through the dequantizer 234 and the inverse transform unit 235. The adder 250 may add the reconstructed residual signal to the prediction signal output from the inter predictor 221 or the intra predictor 222 to generate a reconstructed signal (reconstructed picture, reconstructed block, reconstructed sample array). When there is no

residual for the processing target block, such as when the skip mode is applied, the predicted block may be used as a reconstructed block. The adder 250 may be referred to as a restoration unit or a restoration block generator. The generated reconstructed signal may be used for intra prediction of a next processing target block in the current picture, or may be used for inter
5 prediction of the next picture after being filtered as described below.

[63] Meanwhile, luma mapping with chroma scaling (LMCS) may be applied during a picture encoding and/or reconstruction process.

[64] The filter 260 may improve subjective/objective image quality by applying filtering to the reconstructed signal. For example, the filter 260 may generate a modified reconstructed
10 picture by applying various filtering methods to the reconstructed picture, and store the modified reconstructed picture in the memory 270, specifically, in a DPB of the memory 270. The various filtering methods may include, for example, deblocking filtering, a sample adaptive offset, an adaptive loop filter, a bilateral filter, and the like. The filter 260 may generate various kinds of information related to the filtering, and transfer the generated
15 information to the entropy encoder 240 as described later in the description of each filtering method. The information related to the filtering may be encoded by the entropy encoder 240 and output in the form of a bitstream.

[65] The modified reconstructed picture transmitted to the memory 270 may be used as a reference picture in the inter predictor 221. When the inter prediction is applied through the
20 encoding apparatus, prediction mismatch between the encoding apparatus 200 and the decoding apparatus may be avoided and encoding efficiency may be improved.

[66] The DPB of the memory 270 may store the modified reconstructed picture for use as the reference picture in the inter predictor 221. The memory 270 may store motion information of a block from which the motion information in the current picture is derived (or encoded)
25 and/or motion information of blocks in the picture, having already been reconstructed. The

stored motion information may be transferred to the inter predictor 221 to be utilized as motion information of the spatial neighboring block or motion information of the temporal neighboring block. The memory 270 may store reconstructed samples of reconstructed blocks in the current picture, and may transfer the reconstructed samples to the intra predictor 222.

5 **[67]** FIG. 3 is a diagram for schematically explaining the configuration of a video/image decoding apparatus to which the embodiments of the present disclosure may be applied.

[68] Referring to FIG. 3, the decoding apparatus 300 may include and configured with an entropy decoder 310, a residual processor 320, a predictor 330, an adder 340, a filter 350, and a memory 360. The predictor 330 may include an inter predictor 331 and an intra predictor 332.

10 The residual processor 320 may include a dequantizer 321 and an inverse transformer 322. The entropy decoder 310, the residual processor 320, the predictor 330, the adder 340, and the filter 350, which have been described above, may be configured by one or more hardware components (e.g., decoder chipsets or processors) according to an embodiment. Further, the memory 360 may include a decoded picture buffer (DPB), and may be configured by a digital
15 storage medium. The hardware component may further include the memory 360 as an internal/external component.

[69] When the bitstream including the video/image information is input, the decoding apparatus 300 may reconstruct the image in response to a process in which the video/image information is processed in the encoding apparatus illustrated in FIG. 2. For example, the
20 decoding apparatus 300 may derive the units/blocks based on block split-related information acquired from the bitstream. The decoding apparatus 300 may perform decoding using the processing unit applied to the encoding apparatus. Therefore, the processing unit for the decoding may be, for example, a coding unit, and the coding unit may be split according to the quad-tree structure, the binary-tree structure, and/or the ternary-tree structure from the coding
25 tree unit or the maximum coding unit. One or more transform units may be derived from the

coding unit. In addition, the reconstructed image signal decoded and output through the decoding apparatus 300 may be reproduced through a reproducing apparatus.

[70] The decoding apparatus 300 may receive a signal output from the encoding apparatus of Figure 2 in the form of a bitstream, and the received signal may be decoded through the entropy decoder 310. For example, the entropy decoder 310 may parse the bitstream to derive information (e.g., video/image information) necessary for image reconstruction (or picture reconstruction). The video/image information may further include information on various parameter sets such as an adaptation parameter set (APS), a picture parameter set (PPS), a sequence parameter set (SPS), or a video parameter set (VPS). In addition, the video/image information may further include general constraint information. The decoding apparatus may further decode picture based on the information on the parameter set and/or the general constraint information. Signaled/received information and/or syntax elements described later in this document may be decoded may decode the decoding procedure and obtained from the bitstream. For example, the entropy decoder 310 decodes the information in the bitstream based on a coding method such as exponential Golomb coding, context-adaptive variable length coding (CAVLC), or context-adaptive arithmetic coding (CABAC), and output syntax elements required for image reconstruction and quantized values of transform coefficients for residual. More specifically, the CABAC entropy decoding method may receive a bin corresponding to each syntax element in the bitstream, determine a context model by using a decoding target syntax element information, decoding information of a decoding target block or information of a symbol/bin decoded in a previous stage, and perform an arithmetic decoding on the bin by predicting a probability of occurrence of a bin according to the determined context model, and generate a symbol corresponding to the value of each syntax element. In this case, the CABAC entropy decoding method may update the context model by using the information of the decoded symbol/bin for a context model of a next symbol/bin after determining the

context model. The information related to the prediction among the information decoded by the entropy decoder 310 may be provided to the predictor (inter predictor 332 and intra predictor 331), and residual values on which the entropy decoding has been performed in the entropy decoder 310, that is, the quantized transform coefficients and related parameter information, may be input to the residual processor 320.

[71] The residual processor 320 may derive a residual signal (residual block, residual samples, and residual sample array). Also, information on filtering among the information decoded by the entropy decoder 310 may be provided to the filter 350. Meanwhile, a receiving unit (not shown) for receiving a signal output from the encoding apparatus may be further configured as an internal/external element of the decoding apparatus 300, or the receiving unit may be a component of the entropy decoder 310. Meanwhile, the decoding apparatus according to this document may be called a video/image/picture decoding apparatus, and the decoding apparatus may be divided into an information decoder (video/image/picture information decoder) and a sample decoder (video/image/picture sample decoder). The information decoder may include the entropy decoder 310, and the sample decoder may include at least one of the dequantizer 321, the inverse transformer 322, the adder 340, the filter 350, the memory 360, an inter predictor 332, and an intra predictor 331.

[72] The dequantizer 321 may dequantize the quantized transform coefficients to output the transform coefficients. The dequantizer 321 may rearrange the quantized transform coefficients in a two-dimensional block form. In this case, the rearrangement may be performed based on a coefficient scan order performed by the encoding apparatus. The dequantizer 321 may perform dequantization for the quantized transform coefficients using a quantization parameter (e.g., quantization step size information), and acquire the transform coefficients.

[73] The inverse transformer 322 inversely transforms the transform coefficients to acquire the residual signal (residual block, residual sample array).

[74] The predictor 330 may perform the prediction of the current block, and generate a predicted block including the prediction samples of the current block. The predictor may determine whether the intra prediction is applied or the inter prediction is applied to the current block based on the information on prediction output from the entropy decoder 310, and
5 determine a specific intra/inter prediction mode.

[75] The predictor 330 may generate a prediction signal based on various prediction methods to be described later. For example, the predictor may apply intra prediction or inter prediction for prediction of one block, and may simultaneously apply intra prediction and inter prediction. This may be called combined inter and intra prediction (CIIP). In addition, the
10 predictor may be based on an intra block copy (IBC) prediction mode or based on a palette mode for prediction of a block. The IBC prediction mode or the palette mode may be used for image/video coding of content such as games, for example, screen content coding (SCC). IBC may basically perform prediction within the current picture, but may be performed similarly to inter prediction in that a reference block is derived within the current picture. That is, IBC may
15 use at least one of the inter prediction techniques described in this document. The palette mode may be considered as an example of intra coding or intra prediction. When the palette mode is applied, information on the palette table and the palette index may be included in the video/image information and signaled.

[76] The intra predictor 331 may predict the current block by referring to the samples in the
20 current picture. The referred samples may be located in the neighborhood of the current block, or may be located apart from the current block according to the prediction mode. In intra prediction, prediction modes may include a plurality of non-directional modes and a plurality of directional modes. The intra predictor 331 may determine the prediction mode to be applied to the current block by using the prediction mode applied to the neighboring block.

[77] The inter predictor 332 may derive a predicted block for the current block based on a
25

reference block (reference sample array) specified by a motion vector on a reference picture. In this case, in order to reduce the amount of motion information being transmitted in the inter prediction mode, motion information may be predicted in the unit of blocks, subblocks, or samples based on correlation of motion information between the neighboring block and the current block. The motion information may include a motion vector and a reference picture index. The motion information may further include information on inter prediction direction (L0 prediction, L1 prediction, Bi prediction, and the like). In case of inter prediction, the neighboring block may include a spatial neighboring block existing in the current picture and a temporal neighboring block existing in the reference picture. For example, the inter predictor 332 may construct a motion information candidate list based on neighboring blocks, and derive a motion vector of the current block and/or a reference picture index based on the received candidate selection information. Inter prediction may be performed based on various prediction modes, and the information on the prediction may include information indicating a mode of inter prediction for the current block.

[78] The adder 340 may generate a reconstructed signal (reconstructed picture, reconstructed block, or reconstructed sample array) by adding the obtained residual signal to the prediction signal (predicted block or predicted sample array) output from the predictor (including inter predictor 332 and/or intra predictor 331). If there is no residual for the processing target block, such as a case that a skip mode is applied, the predicted block may be used as the reconstructed block.

[79] The adder 340 may be called a reconstructor or a reconstructed block generator. The generated reconstructed signal may be used for the intra prediction of a next block to be processed in the current picture, and as described later, may also be output through filtering or may also be used for the inter prediction of a next picture.

[80] Meanwhile, a luma mapping with chroma scaling (LMCS) may also be applied in the

picture decoding process.

[81] The filter 350 may improve subjective/objective image quality by applying filtering to the reconstructed signal. For example, the filter 350 may generate a modified reconstructed picture by applying various filtering methods to the reconstructed picture, and store the modified reconstructed picture in the memory 360, specifically, in a DPB of the memory 360. The various filtering methods may include, for example, deblocking filtering, a sample adaptive offset, an adaptive loop filter, a bilateral filter, and the like.

[82] The (modified) reconstructed picture stored in the DPB of the memory 360 may be used as a reference picture in the inter predictor 332. The memory 360 may store the motion information of the block from which the motion information in the current picture is derived (or decoded) and/or the motion information of the blocks in the picture having already been reconstructed. The stored motion information may be transferred to the inter predictor 332 so as to be utilized as the motion information of the spatial neighboring block or the motion information of the temporal neighboring block. The memory 360 may store reconstructed samples of reconstructed blocks in the current picture, and transfer the reconstructed samples to the intra predictor 331.

[83] In the present document, the embodiments described in the filter 260, the inter predictor 221, and the intra predictor 222 of the encoding apparatus 200 may be applied equally or to correspond to the filter 350, the inter predictor 332, and the intra predictor 331.

[84] The video/image coding method according to the present document may be performed based on the following partitioning structure. Specifically, procedures of prediction, residual processing ((inverse) transform and (de)quantization), syntax element coding, and filtering to be described later may be performed based on CTU and CU (and/or TU and PU) derived based on the partitioning structure. A block partitioning procedure may be performed by the image partitioner 210 of the above-described encoding apparatus, and partitioning related information

may be processed (encoded) by the entropy encoder 240, and may be transferred to the decoding apparatus in the form of a bitstream. The entropy decoder 310 of the decoding apparatus may derive the block partitioning structure of the current picture based on the partitioning related information obtained from the bitstream, and based on this, may perform a series of procedures for image decoding (e.g., prediction, residual processing, block/picture reconstruction, and in-loop filtering). The CU size and the TU size may be equal to each other, or a plurality of TUs may be present in the CU area. Meanwhile, the CU size may generally represent a luma component (sample) coding block (CB) size. The TU size may generally represent a luma component (sample) transform block (TB) size. A chroma component (sample) CB or TB size may be derived based on the luma component (sample) CB or TB size in accordance with a component ratio according to a color format (chroma format, e.g., 4:4:4, 4:2:2, 4:2:0, and the like) of the picture/image. The TU size may be derived based on maxTbSize. For example, if the CU size is larger than the maxTbSize, a plurality of TUs (TBs) of the maxTbSize may be derived, and the transform / inverse transform may be performed in the unit of the TU (TB). Further, for example, in case that intra prediction is applied, the intra prediction mode/type may be derived in the unit of CU (or CB), and the derivation of a neighboring reference sample and the generation of a prediction sample may be performed in the unit of TU (or TB). In this case, one or a plurality of TUs (or TBs) may be present in one CU (or CB) area, and in this case, the plurality of TUs (or TBs) may share the same intra prediction mode/type.

[85] Further, in coding the video/image according to the present document, the image processing unit may have a hierarchical structure. One picture may be divided into one or more tiles, bricks, slices, and/or tile groups. One slice may include one or more bricks. One brick may include one or more CTU rows in a tile. The slice may include an integer number of bricks of the picture. One tile group may include one or more tiles. One tile is a rectangular region of

CTUs within a particular tile column and a particular tile row in a picture. The tile group may include an integer number of tiles in accordance with the tile raster scan in the picture. A slice header may carry information/parameters that can be applied to the corresponding slice (blocks in the slice). In case that the encoding/decoding apparatus has a multi-core processor, the encoding/decoding procedures for the tile, slice, brick, and/or tile group may be processed in parallel. In the present document, the slice or the tile group may be interchangeably used. That is, the tile group header may be called a slice header. Here, the slice may have one of slice types including an intra (I) slice, a predictive (P) slice, and a bi-predictive (B) slice. For prediction of blocks in I slice, the inter prediction may not be used, but only the intra prediction may be used. Even in this case, the original sample value may be coded and signaled without the prediction. For blocks in the P slice, the intra prediction or the inter prediction may be used, and in case of using the inter prediction, only uni-prediction may be used. Meanwhile, for blocks in the B slice, the intra prediction or the inter prediction may be used, and in case of using the inter prediction, maximally up to bi-prediction may be used.

[86] In accordance with the characteristics (e.g., resolution) of the video image, or in consideration of the coding efficiency or parallel processing, the encoder may determine the tile / tile group, brick, slice, maximum and minimum coding unit sizes, and corresponding information or information capable of inducing the same may be included in the bitstream.

[87] The decoder may obtain information representing whether the tile / tile group, brick, slice, or CTU in the tile of the current picture has been partitioned into a plurality of coding units. By obtaining (transmitting) such information only under a specific condition, the efficiency can be enhanced.

[88] Pictures may be divided into sequences of coding tree units (CTUs). The CTU may correspond to a coding tree block (CTB). Further, the CTU may include a coding tree block of luma samples, and two coding tree blocks of corresponding chroma samples. In other words,

for the picture including three sample arrays, the CTU may include an NxN block of the luma samples and two corresponding blocks of the chroma samples.

[89] The maximum supported size of the CTU for coding and prediction may differ from the maximum supported size of the CTU for transform. For example, the maximum supported size of the luma block in the CTU may be 128x128.

[90] The CTU may be divided into CUs based on a quad-tree (QT) structure. The quad-tree structure may be called a quaternary tree structure. This is to reflect various local characteristics. Meanwhile, in the present document, the CTU may be divided based on splitting of the multi-type tree structure including not only a quad-tree but also a binary-tree (BT) and a ternary-tree (TT). Hereinafter, a QTBT structure may include a splitting structure based on the quad-tree and the binary-tree, and a QTBT TT may include a splitting structure based on the quad-tree, the binary-tree, and the ternary-tree. Further, the QTBT structure may include a splitting structure based on the quad-tree, the binary-tree, and the ternary-tree. In the coding tree structure, the CU may have a square or rectangular shape. The CTU may be first split into the quad-tree structure. Thereafter, leaf nodes of the quad-tree structure may be further split by the multi-type tree structure. For example, schematically four splitting types may be included in the multi-type tree structure.

[91] The four splitting types may include vertical binary splitting (SPLIT_BT_VER), horizontal binary splitting (SPLIT_BT_HOR), vertical ternary splitting (SPLIT_TT_VER), and horizontal ternary splitting (SPLIT_TT_HOR). The leaf nodes of the multi-type tree structure may be called CUs. Such CUs may be used for the prediction and transform procedure. In the present document, the CU, PU, and TU may generally have the same block size. However, if the maximum supported transform length is smaller than the width or height of a color component of the CU, the CU and the TU may have different block sizes.

[92] Meanwhile, as described above, one picture may include a plurality of slices, and one

slice may include a slice header and slice data. In this case, one picture header may be further added for the plurality of slices (set of the slice header and slice data) in one picture. The picture header (picture header syntax) may include information/parameters capable of being commonly applied to the picture. The slice header (slice header syntax) may include
5 information/parameters capable of being commonly applied to the slice. An adaptation parameter set (APS) or a picture parameter set (PPS) may include information/parameters capable of being commonly applied to one or more pictures. A sequence parameter set (SPS) may include information/parameters capable of being commonly applied to one or more sequences. A video parameter set (VPS) may include information/parameters capable of being
10 commonly applied to multiple layers. A decoding parameter set (DPS) may include information/parameters capable of being commonly applied to the overall video. The DPS may include information/parameters related to concatenation of a coded video sequence (CVS).

[93] In the present document, an upper layer syntax may include at least one of the APS syntax, PPS syntax, SPS syntax, VPS syntax, DPS syntax, picture header syntax, and slice
15 header syntax.

[94] Further, for example, information on the division and the configuration of the tile / tile group / brick / slice may be configured by the encoding end through the upper level syntax, and may be transferred to the decoding apparatus in the form of a bitstream.

[95] In the present document, at least one of quantization/dequantization and/or
20 transform/inverse transform may be omitted. When the quantization/dequantization is omitted, the quantized transform coefficient may be referred to as a transform coefficient. When the transform/inverse transform is omitted, the transform coefficient may be called a coefficient or a residual coefficient or may still be called the transform coefficient for uniformity of expression.

[96] In the present document, the quantized transform coefficient and the transform
25

coefficient may be referred to as a transform coefficient and a scaled transform coefficient, respectively. In this case, the residual information may include information on transform coefficient(s), and the information on the transform coefficient(s) may be signaled through residual coding syntax. Transform coefficients may be derived based on the residual
5 information (or information on the transform coefficient(s)), and scaled transform coefficients may be derived through inverse transform (scaling) on the transform coefficients. Residual samples may be derived based on inverse transform (transform) of the scaled transform coefficients. This may be applied/expressed in other parts of the present document as well.

[97] As the above-described contents, the encoding apparatus may perform various
10 encoding methods, for example, such as exponential Golomb, context-adaptive variable length coding (CAVLC), and context-adaptive binary arithmetic coding (CABAC). Further, the decoding apparatus may decode information in the bitstream based on the coding method, such as the exponential Golomb, CAVLC, or CABAC, and may output a value of a syntax element and quantized values of transform coefficients for residual, being necessary for image
15 reconstruction. For example, the above-described coding methods may be performed as in the contents to be described later.

[98] In the present document, the intra prediction may represent the prediction that generates the prediction samples for the current block based on the reference samples in the picture (hereinafter, current picture) to which the current block belongs. In case that the intra
20 prediction is applied to the current block, the neighboring reference samples to be used for the intra prediction of the current block may be derived. The neighboring reference samples of the current block may include a sample adjacent to the left boundary of the current block having the size of $nW \times nH$ and total $2 \times nH$ samples neighboring the bottom-left, a sample adjacent to the top boundary of the current block and total $2 \times nW$ samples neighboring the top-right, and
25 one sample neighboring the top-left of the current block. Further, the neighboring reference

samples of the current block may include a plurality of columns of top neighboring samples and a plurality of rows of left neighboring samples. Further, the neighboring reference samples of the current block may include total nH samples adjacent to the right boundary of the current block having the size of $nW \times nH$, total nW samples adjacent to the bottom boundary of the current block, and one sample neighboring the bottom-right of the current block.

[99] However, some of the neighboring reference samples of the current block may have not yet been decoded or enabled. In this case, the decoding apparatus may configure the neighboring reference samples to be used for the prediction through substitution of enabled samples for non-enabled sample. Further, the neighboring reference samples to be used for the prediction may be configured through interpolation of the enabled samples.

[100] In case that the neighboring reference samples are derived, (i) the prediction sample may be induced based on an average or interpolation of the neighboring reference samples of the current block, and (ii) the prediction sample may be induced based on the reference sample that is present in a specific (prediction) direction for the prediction sample among the neighboring reference samples of the current block. The case of (i) may be called a non-directional mode or a non-angular mode, and the case of (ii) may be called a directional mode or an angular mode. Further, the prediction sample may be generated through interpolation of the first neighboring sample with the second neighboring sample located in an opposite direction to the prediction direction of the intra prediction mode of the current block based on the prediction sample of the current block among the neighboring reference samples. The above-described case may be called a linear interpolation intra prediction (LIP). Further, chroma prediction samples may be generated based on luma samples using a linear model. This case may be called an LM mode. Further, a temporary prediction sample of the current block may be derived based on the filtered neighboring reference samples, and a prediction sample of the current block may be derived by calculating a weighted sum of the temporary prediction

sample and at least one reference sample derived in accordance with the intra prediction mode among the existing neighboring reference samples, that is, non-filtered neighboring reference samples. The above-described case may be called a position dependent intra prediction (PDPC). Further, the prediction sample may be derived using a reference sample located in a prediction direction on a reference sample line having the highest prediction accuracy among neighboring multiple reference sample lines of the current block through selection of the corresponding line, and in this case, intra prediction coding may be performed in a method for indicating (signaling) the used reference sample line to the decoding apparatus. The above-described case may be called multi-reference line (MRL) intra prediction or MRL-based intra prediction. Further, the intra prediction may be performed based on the same intra prediction mode through division of the current block into vertical or horizontal subpartitions, and the neighboring reference samples may be derived and used in the unit of a subpartition. That is, in this case, since the intra prediction mode for the current block is equally applied to the subpartitions, and the neighboring reference samples are derived and used in the unit of the subpartition, the intra prediction performance can be enhanced in some cases. Such a prediction method may be called intra subpartitions (ISP) or ISP-based intra prediction. The above-described intra prediction method may be called the intra prediction type in distinction from the intra prediction mode. The intra prediction type may be called by various terms, such as an intra prediction technique or an additional intra prediction mode. For example, the intra prediction type (or additional intra prediction mode) may include at least one of LIP, PDPC, MRL, or ISP described above. A general intra prediction method excluding a specific intra prediction type, such as the LIP, PDPC, MRL, or ISP, may be called a normal intra prediction type. The normal intra prediction type may be generally applied in case that the specific intra prediction type is not applied, and the prediction may be performed based on the above-described intra prediction mode. Meanwhile, as needed, post-filtering for the derived prediction sample may be

performed.

[101] Specifically, the intra prediction procedure may include steps of intra prediction mode/type determination, neighboring reference sample derivation, and intra prediction mode/type-based prediction sample derivation. Further, as needed, a post-filtering step for the derived prediction sample may be performed.

[102] Meanwhile, in addition to the above-described prediction types, an affine linear weighted intra prediction (ALWIP) may be used. The ALWIP may be called linear weighted intra prediction (LWIP) or matrix weighted intra prediction (MIP) or matrix based intra prediction. In case that the MIP is applied for the current block, i) by using the neighboring reference samples for which an averaging procedure has been performed ii) a matrix-vector-multiplication procedure may be performed, and iii) as needed, the prediction samples for the current block may be derived by further performing a horizontal/vertical interpolation. The intra prediction modes being used for the MIP may be configured differently from the above-described LIP, PDPC, MRL, or ISP intra prediction, or the intra prediction modes being used for the normal intra prediction. The intra prediction mode for the MIP may be called an MIP intra prediction mode, an MIP prediction mode, or an MIP mode. For example, in accordance with the intra prediction mode for the MIP, a matrix and an offset being used for the matrix vector multiplication may be differently configured. Here, the matrix may be called an (MIP) weighted matrix, and the offset may be called an (MIP) offset vector or an (MIP) bias vector.

[103] Schematically for example, the intra prediction based video/image encoding procedure may include the followings.

[104] FIG. 4 illustrates an example of a video/image encoding method based on intra prediction, and FIG. 5 schematically illustrates an intra predictor in an encoding apparatus.

[105] Referring to FIGS. 4 and 5, S400 may be performed by the intra predictor 222 of the encoding apparatus, and S410 to S430 may be performed by the residual processor 230 of the

encoding apparatus. Specifically, S410 may be performed by the subtractor 231 of the encoding apparatus, S420 may be performed by the transformer 232 and the quantizer 233 of the encoding apparatus, and S430 may be performed by the dequantizer 234 and the inverse transformer 235 of the encoding apparatus. In S400, prediction information may be derived by the intra predictor 222, and may be encoded by the entropy encoder 240. Through S410 and S420, residual information may be derived, and may be encoded by the entropy encoder 240. The residual information is information about the residual samples. The residual information may include information about the quantized transform coefficients for the residual samples. As described above, the residual samples may be derived as transform coefficients through the transformer 232 of the encoding apparatus, and the transform coefficients may be derived as the quantized transform coefficients through the quantizer 233. The information about the quantized transform coefficients may be encoded by the entropy encoder 240 through the residual coding procedure.

[106] The encoding apparatus performs intra prediction for the current block (S400). The encoding apparatus may derive an intra prediction mode for the current block, derive neighboring reference samples of the current block, and generate prediction samples in the current block based on the intra prediction mode and the neighboring reference samples. Here, the procedures of determining the intra prediction mode, deriving the neighboring reference samples, and generating the prediction samples may be simultaneously performed, and any one procedure may be performed prior to another procedure. For example, the intra predictor 222 of the encoding apparatus may include a prediction mode/type determiner 222_1, a reference sample deriver 222_2, and a prediction sample deriver 222_3, and the prediction mode/type determiner 222_1 may determine the intra prediction mode/type for the current block, the reference sample deriver 222_2 may derive the neighboring reference samples of the current block, and the prediction sample deriver 222_3 may derive motion samples of the current block.

Meanwhile, in case that a prediction sample filtering procedure to be described later is performed, the intra predictor 222 may further include a prediction sample filter (not illustrated). The encoding apparatus may determine a mode being applied for the current block among a plurality of intra prediction modes. The encoding apparatus may compare RD costs for the intra prediction modes, and may determine the optimum intra prediction mode for the current block.

[107] Meanwhile, the encoding apparatus may perform the prediction sample filtering procedure. The prediction sample filtering may be called post-filtering. Some or all of the prediction samples may be filtered by the prediction sample filtering procedure. In some cases, the prediction sample filtering procedure may be omitted.

[108] The encoding apparatus derives residual samples for the current block based on the prediction samples (S410). The encoding apparatus may compare the prediction samples with the original samples of the current block based on phase, and may derive the residual samples.

[109] The encoding apparatus may derive quantized transform coefficients through transform/quantization of the residual samples (S420), and then may derive the (modified) residual samples by performing the dequantization / inverse transform of the quantized transform coefficients again (S430). The reason why to perform the dequantization / inverse transform again after the transform/quantization is to derive the same residual samples as the residual samples being derived by the decoding apparatus as described above.

[110] The encoding apparatus may generate a reconstructed block including reconstructed samples for the current block based on the prediction samples and the (modified) residual samples (S440). Based on the reconstructed block a reconstructed picture for the current picture may be generated.

[111] As described above, the encoding apparatus may encode image information including prediction information on the intra prediction (e.g., prediction mode information representing

the prediction mode) and the residual information on the intra/residual samples, and may output the encoded image information in the form of the bitstream. The residual information may include a residual coding syntax. The encoding apparatus may derive quantized transform coefficients through transform/quantization of the residual samples. The residual information
5 may include information on the quantized transform coefficients.

[112] Schematically for example, the intra prediction based video/image decoding procedure may include the followings.

[113] FIG. 6 illustrates an example of a video/image decoding method based on intra prediction, and FIG. 7 schematically illustrates an intra predictor in a decoding apparatus.

10 **[114]** The decoding apparatus may perform an operation corresponding to the operation performed by the encoding apparatus.

[115] Referring to FIGS. 6 and 7, S600 to 610 may be performed by the intra predictor 331 of the decoding apparatus, and prediction information of S600 and residual information of S630 may be obtained from a bitstream by the entropy decoder 310 of the decoding apparatus. The
15 residual processor 320 of the decoding apparatus may derive residual samples for the current block based on the residual information. Specifically, the dequantizer 321 of the residual processor 320 may derive transform coefficients by performing dequantization based on the quantized transform coefficients derived based on the residual information, and the inverse transformer 322 of the residual processor may derive the residual samples for the current block
20 by performing dequantization for the transform coefficients. S640 may be performed by the adder 340 or the reconstructor of the decoding apparatus.

[116] Specifically, the decoding apparatus may derive the intra prediction mode for the current block based on the received prediction information (S600). The decoding apparatus may derive neighboring reference samples of the current block (S610). The decoding apparatus
25 generates prediction samples in the current block by performing the intra prediction based on

the intra prediction mode and the neighboring reference samples (S620). In this case, the decoding apparatus may perform the prediction sample filtering procedure. The prediction sample filtering may be called post-filtering. Some or all of the prediction samples may be filtered by the prediction sample filtering procedure. In some cases, the prediction sample filtering procedure may be omitted.

[117] The decoding apparatus generates the residual samples for the current block based on the received residual information (S630). The decoding apparatus may generate reconstructed samples for the current block based on the prediction samples and the residual samples, and may derive a reconstructed block including the reconstructed samples (S640). A reconstructed picture for the current picture may be generated based on the reconstructed block.

[118] Here, the intra predictor 331 of the decoding apparatus may include a prediction mode/type determiner 331_1, a reference sample deriver 331_2, and a prediction sample deriver 331_3. The prediction mode/type determiner 331_1 may determine the intra prediction mode for the current block based on prediction mode information obtained by the entropy decoder 310 of the decoding apparatus, the reference sample deriver 331_2 may derive the neighboring reference samples of the current block, and the prediction sample deriver 331_3 may derive the prediction sample of the current block. Meanwhile, in case that the above-described prediction sample filtering procedure is performed, the intra predictor 331 may further include a prediction sample filter (not illustrated).

[119] The prediction information may include intra prediction mode information and/or intra prediction type information. For example, the intra prediction mode information may include flag information (e.g., `intra_luma_mpm_flag`) representing whether a most probable mode (MPM) or a remaining mode is applied to the current block, and in case that the MPM is applied to the current block, the prediction mode information may further include index information (e.g., `intra_luma_mpm_idx`) representing one of intra prediction mode candidates (MPM

candidates). The intra prediction mode candidates (MPM candidates) may constitute an MPM candidate list or an MPM list. Further, in case that the MPM is not applied to the current block the intra prediction mode information may further include remaining mode information (e.g., intra_luma_mpm_remainder) indicating one of remaining intra prediction modes excluding the intra prediction mode candidates (MPM candidates). The decoding apparatus may determine the intra prediction mode of the current block based on the intra prediction mode information. For the above-described MIP, a separate MPM list may be configured.

[120] Further, the intra prediction type information may be implemented in various forms. As an example, the intra prediction type information may include intra prediction type index information indicating one of the intra prediction types. As another example, the intra prediction type information may include at least one of reference sample line information (e.g., intra_luma_ref_idx) representing whether the MRL is applied to the current block and which reference sample line is used in case that the MRL is applied, ISP flag information (e.g., intra_subpartitions_mode_flag) representing whether the ISP is applied to the current block, ISP type information (e.g., intra_subpartitions_split_flag) indicating the partitioning type of subpartitions in case that the ISP is applied, flag information representing whether the PDCP is applied, or flag information representing whether the LIP is applied. Further, the intra prediction type information may include an MIP flag representing whether the MIP is applied to the current block.

[121] The intra prediction mode information and/or the intra prediction type information may be encoded/decoded through the coding method described in the present document. For example, the intra prediction mode information and/or the intra prediction type information may be encoded/decoded through entropy coding (e.g., CABAC or CAVLC) based on a truncated (rice) binary code.

[122] Meanwhile, schematically for example, the video/image encoding procedure based on

the inter prediction may include the followings.

[123] FIG. 8 illustrates an example of video/image encoding method based on inter prediction, and FIG. 9 schematically illustrates an inter predictor in an encoding apparatus.

[124] Referring to FIGS. 8 and 9, the encoding apparatus performs inter prediction for the current block (S800). The encoding apparatus may derive an inter prediction mode and motion information of the current block, and may generate prediction samples of the current block. Here, the procedures of determining the inter prediction mode, deriving motion information, and generating the prediction samples may be performed simultaneously, and any one procedure may be performed prior to another procedure. For example, the inter predictor 221 of the encoding apparatus may include a prediction mode determiner 221_1, a motion information deriver 221_2, and a prediction sample deriver 221_3, and the prediction mode determiner 221_1 may determine the prediction mode for the current block, the motion information deriver 221_2 may derive the motion information of the current block, and the prediction sample deriver 221_3 may derive the prediction samples for the current block. For example, the inter predictor of the encoding apparatus may search a block similar to the current block in a specific area (search area) of the reference pictures through motion estimation, and may derive the reference block having a difference from the current block that is the minimum or equal to or lower than a predetermined level. Based on this, a reference picture index indicating a reference picture in which the reference block is located is derived, and a motion vector is derived based on a difference in location between the reference block and the current block. The encoding apparatus may determine a mode being applied for the current block among various prediction modes. The encoding apparatus may compare rate-distortion (RD) costs for the various prediction modes, and may determine the optimum prediction mode for the current block.

[125] For example, in case that a skip mode or a merge mode is applied to the current block,

the encoding apparatus may configure a merge candidate list, and may derive the reference block having the difference from the current block that is the minimum or equal to or lower than the predetermined level among the reference blocks indicated by the merge candidates included in the merge candidate list. In this case, a merge candidate related to the derived
5 reference block may be selected, and merge index information indicating the selected merge candidate may be generated and signaled to the decoding apparatus. The motion information of the current block may be derived by using the motion information of the selected merge candidate.

[126] As another example, in case that an (A)MVP mode is applied to the current block, the
10 encoding apparatus may configure an (A)MVP candidate list, and may use the motion vector of the mvp candidate selected among motion vector predictor (mvp) candidates included in the (A)MVP candidate list as the mvp of the current block. In this case, for example, the motion vector indicating the reference block derived by the above-described motion estimation may be used as the motion vector of the current block, and the mvp candidate having the motion
15 vector having the smallest difference from the motion vector of the current block among the mvp candidates may be the selected mvp candidate. A motion vector difference (MVD), which is the difference obtained by subtracting the mvp from the motion vector of the current block, may be derived. In this case, information on the MVD may be signaled to the decoding apparatus. Further, in case that the (A)MVP mode is applied, the value of the reference picture
20 index may be configured as the reference picture index information, and may be separately signaled to the decoding apparatus.

[127] The encoding apparatus may derive residual samples based on the prediction samples (S810). The encoding apparatus may derive the residual samples through comparison of the prediction samples with the original samples of the current block.

[128] The encoding apparatus encodes image information including prediction information

and residual information (S820). The encoding apparatus may output the encoded image information in the form of a bitstream. The prediction information may be information related to the prediction procedure, and may include information related to the prediction mode information (e.g., skip flag, merge flag, or mode index) and the motion information. The information about the motion information may include candidate selection information (e.g., merge index, mvp flag, or mvp index) that is the information for deriving the motion vector. Further, the information about the motion information may include information about the above-described MVD and/or the reference picture index information. Further, the information about the motion information may include information representing whether L0 prediction, L1 prediction, or bi-prediction is applied. The residual information is information about the residual samples. The residual information may include information about quantized transform coefficients for the residual samples.

[129] The output bitstream may be stored in a (digital) storage medium to be transferred to the decoding apparatus, or may be transferred to the decoding apparatus through a network.

[130] Meanwhile, as described above, the encoding apparatus may generate a reconstructed picture (including reconstructed samples and reconstructed block) based on the reference samples and the residual samples. This is for the encoding apparatus to derive the same prediction result as that performed in the decoding apparatus, and through this, the coding efficiency can be enhanced. Accordingly, the encoding apparatus may store the reconstructed picture (or reconstructed samples or reconstructed block) in a memory, and may utilize the same as a reference picture for the inter prediction. As described above, an in-loop filtering procedure may be further applied to the reconstructed picture.

[131] Schematically for example, a video/image decoding procedure based on inter prediction may include the followings.

[132] FIG. 10 illustrates an example of a video/image decoding method based on inter

prediction, and FIG. 11 schematically illustrates an inter predictor in a decoding apparatus.

[133] The decoding apparatus may perform an operation corresponding to the operation performed by the encoding apparatus. The decoding apparatus may perform prediction for the current block based on the received prediction information, and may derive prediction samples.

5 **[134]** Specifically, referring to FIGS. 10 and 11, the decoding apparatus may determine a prediction mode for the current block based on prediction information received from a bitstream (S1000). The decoding apparatus may determine which inter prediction mode is applied to the current block based on prediction mode information in the prediction information.

[135] For example, it may be determined whether a merge mode is applied to the current
10 block based on a merge flag, or an (A)MVP mode is determined. Further, one of various inter prediction mode candidates may be selected based on the merge index. The inter prediction mode candidates may include various inter prediction modes, such as a skip mode, a merge mode, and/or an (A)MVP mode.

[136] The decoding apparatus derives motion information of the current block based on the
15 determined inter prediction mode (S1010). For example, in case that a skip mode or a merge mode is applied to the current block, the decoding apparatus may configure a merge candidate list to be described later, and may select one of merge candidates included in the merge candidate list. The selection may be performed based on the above-described selection information (merge index). The motion information of the current block may be derived by
20 using the motion information of the selected merge candidate. The motion information of the selected merge candidate may be used as the motion information of the current block.

[137] As another example, in case that the (A)MVP mode is applied to the current block, the decoding apparatus may configure the (A)MVP candidate list, and may use the motion vector of the mvp candidate selected among the motion vector predictor (mvp) candidates included in
25 the (A)MVP candidate list. The selection may be performed based on the above-described

selection information (mvp flag or mvp index). In this case, the MVD of the current block may be derived based on information on the MVD, and the motion vector of the current block may be derived based on the mvp and the MVD of the current block. Further, the reference picture index of the current block may be derived based on the reference picture index information.

5 The picture indicated by the reference picture index in the reference picture list for the current block may be derived as the reference picture being referred to for the inter prediction of the current block.

[138] Meanwhile, the motion information of the current block may be derived without the candidate list configuration, and in this case, the above-described candidate list configuration
10 may be omitted.

[139] The decoding apparatus may generate prediction samples for the current block based on the motion information of the current block (S1020). In this case, the reference picture may be derived based on the reference picture index of the current block, and the motion vector of the current block may derive the prediction samples of the current block by using samples of
15 the reference block indicated in the reference picture. In this case, to be described later, the prediction sample filtering procedure may be further performed for all or some of the prediction samples of the current block as needed.

[140] For example, the inter predictor 332 of the decoding apparatus may include a prediction mode determiner 332_1, a motion information deriver 332_2, and a prediction
20 sample deriver 332_3, and the prediction mode determiner 332_1 may determine the prediction mode for the current block based on the received prediction mode information, the motion information deriver 332_2 may derive the motion information (motion vector and/or reference picture index) of the current block based on the information on the received motion information, and the prediction sample deriver 332_3 may derive the prediction samples of the current block.

25 [141] The decoding apparatus generates residual samples for the current block based on the

received residual information (S1030). The decoding apparatus may generate reconstructed samples for the current block based on the prediction samples and the residual samples (S1040). Thereafter, the in-loop filtering procedure may be further applied to the reconstructed picture as described above.

5 **[142]** Meanwhile, as described above, a high level syntax (HLS) may be coded/signaled for video/image coding. The coded picture may be composed of one or more slices. A parameter describing the coded picture is signaled in the picture header, and a parameter describing the slice is signaled in the slice header. The picture header is carried in the form of a NAL unit itself. The slice header is present at a start portion of the NAL unit including a payload of the
10 slice (i.e., slice data).

[143] Each picture is related to a picture header. The picture may be composed of different types of slices (intra-coded slice (i.e., I slice) and inter-coded slice (i.e., P slice and B slice)). Accordingly, the picture header may include syntax elements necessary for the intra slice of the picture and the inter slice of the picture. For example, the syntax of the picture header may
15 be as in the following Table 1.

[144] [Table 1]

picture_header_rbsp() {	Descriptor
non_reference_picture_flag	u(1)
gdr_pic_flag	u(1)
no_output_of_prior_pics_flag	u(1)
if(gdr_pic_flag)	
recovery_poc_cnt	ue(v)
ph_pic_parameter_set_id	ue(v)
if(sps_poc_msb_flag) {	
ph_poc_msb_present_flag	u(1)
if(ph_poc_msb_present_flag)	
poc_msb_val	u(v)
}	
if(sps_subpic_id_present_flag && !sps_subpic_id_signalling_flag) {	

[145]

ph_subpic_id_signalling_present_flag	u(1)
if(ph_subpics_id_signalling_present_flag) {	
ph_subpic_id_len_minus1	ue(v)
for(i = 0; i <= sps_num_subpics_minus1; i++)	
ph_subpic_id[i]	u(v)
}	
}	
if(!sps_loop_filter_across_virtual_boundaries_disabled_present_flag) {	
ph_loop_filter_across_virtual_boundaries_disabled_present_flag	u(1)
if(ph_loop_filter_across_virtual_boundaries_disabled_present_flag) {	
ph_num_ver_virtual_boundaries	u(2)
for(i = 0; i < ph_num_ver_virtual_boundaries; i++)	
ph_virtual_boundaries_pos_x[i]	u(13)
ph_num_hor_virtual_boundaries	u(2)
for(i = 0; i < ph_num_hor_virtual_boundaries; i++)	
ph_virtual_boundaries_pos_y[i]	u(13)
}	
}	
if(separate_colour_plane_flag == 1)	
colour_plane_id	u(2)
if(output_flag_present_flag)	
pic_output_flag	u(1)
pic_rpl_present_flag	u(1)
if(pic_rpl_present_flag) {	

[146]

for(i = 0; i < 2; i++) {	
if(num_ref_pic_lists_in_sps[i] > 0 && !pps_ref_pic_list_sps_idc[i] && (i == 0 (i == 1 && rpl_idx_present_flag)))	
pic_rpl_sps_flag[i]	u(1)
if(pic_rpl_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 && (i == 0 (i == 1 && rpl_idx_present_flag)))	
pic_rpl_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i])	
for(j = 0; j < NumLtrpEntries[i][RplIdx[i]]; j++) {	
if(ltrp_in_slice_header_flag[i][RplIdx[i]])	
pic_poc_lsb_ltrp[i][j]	u(v)
pic_delta_poc_msb_present_flag[i][j]	u(1)
if(pic_delta_poc_msb_present_flag[i][j])	
pic_delta_poc_msb_cycle_H[i][j]	ue(v)
}	
}	
}	
if(partition_constraints_override_enabled_flag) {	
partition_constraints_override_flag	ue(v)
if(partition_constraints_override_flag) {	
pic_log2_diff_min_qt_min_cb_intra_slice_luma	ue(v)

pic_log2_diff_min_qt_min_cb_inter_slice	ue(v)
pic_max_mtt_hierarchy_depth_inter_slice	ue(v)
pic_max_mtt_hierarchy_depth_intra_slice_luma	ue(v)
if(pic_max_mtt_hierarchy_depth_intra_slice_luma != 0) {	
pic_log2_diff_max_bt_min_qt_intra_slice_luma	ue(v)
pic_log2_diff_max_tt_min_qt_intra_slice_luma	ue(v)
}	
if(pic_max_mtt_hierarchy_depth_inter_slice != 0) {	
pic_log2_diff_max_bt_min_qt_inter_slice	ue(v)
pic_log2_diff_max_tt_min_qt_inter_slice	ue(v)
}	
if(qtbtt_dual_tree_intra_flag) {	
pic_log2_diff_min_qt_min_cb_intra_slice_chroma	ue(v)
pic_max_mtt_hierarchy_depth_intra_slice_chroma	ue(v)
if(pic_max_mtt_hierarchy_depth_intra_slice_chroma != 0) {	
pic_log2_diff_max_bt_min_qt_intra_slice_chroma	ue(v)
pic_log2_diff_max_tt_min_qt_intra_slice_chroma	ue(v)
}	
}	
}	
if(cu_qp_delta_enabled_flag) {	
pic_cu_qp_delta_subdiv_intra_slice	ue(v)
pic_cu_qp_delta_subdiv_inter_slice	ue(v)
}	

[147]

if(pps_cu_chroma_qp_offset_list_enabled_flag) {	
pic_cu_chroma_qp_offset_subdiv_intra_slice	ue(v)
pic_cu_chroma_qp_offset_subdiv_inter_slice	ue(v)
}	
if(sps_temporal_mvp_enabled_flag)	
pic_temporal_mvp_enabled_flag	u(1)
if(!pps_mvd_l1_zero_idc)	
mvd_l1_zero_flag	u(1)
if(!pps_six_minus_max_num_merge_cand_plus1)	
pic_six_minus_max_num_merge_cand	ue(v)
if(sps_affine_enabled_flag)	
pic_five_minus_max_num_subblock_merge_cand	ue(v)
if(sps_fpel_mmvd_enabled_flag)	
pic_fpel_mmvd_enabled_flag	u(1)
if(sps_bdof_pic_present_flag)	
pic_disable_bdof_flag	u(1)
if(sps_dmvr_pic_present_flag)	
pic_disable_dmvr_flag	u(1)
if(sps_prof_pic_present_flag)	
pic_disable_prof_flag	u(1)

[148]

[149]

if(sps_triangle_enabled_flag && MaxNumMergeCand >= 2 && !pps_max_num_merge_cand_minus_max_num_triangle_cand_plus1)	
pic_max_num_merge_cand_minus_max_num_triangle_cand	ue(v)
if(sps_ibc_enabled_flag)	
pic_six_minus_max_num_ibc_merge_cand	ue(v)
if(sps_joint_cbr_enabled_flag)	
pic_joint_cbr_sign_flag	u(1)
if(sps_sao_enabled_flag) {	
pic_sao_enabled_present_flag	u(1)
if(pic_sao_enabled_present_flag) {	
pic_sao_luma_enabled_flag	u(1)
if(ChromaArrayType != 0)	
pic_sao_chroma_enabled_flag	u(1)
}	
}	
if(sps_alf_enabled_flag) {	
pic_alf_enabled_present_flag	u(1)
if(pic_alf_enabled_present_flag) {	
pic_alf_enabled_flag	u(1)
if(pic_alf_enabled_flag) {	
pic_num_alf_aps_ids_luma	u(3)
for(i = 0; i < pic_num_alf_aps_ids_luma; i++)	
pic_alf_aps_id_luma[i]	u(3)

[150]

if(ChromaArrayType != 0)	
pic_alf_chroma_idc	u(2)
if(pic_alf_chroma_idc)	
pic_alf_aps_id_chroma	u(3)
}	
}	
if(!pps_dep_quant_enabled_flag)	
pic_dep_quant_enabled_flag	u(1)
if(!pic_dep_quant_enabled_flag)	
sign_data_hiding_enabled_flag	u(1)
if(deblocking_filter_override_enabled_flag) {	
pic_deblocking_filter_override_present_flag	u(1)
if(pic_deblocking_filter_override_present_flag) {	
pic_deblocking_filter_override_flag	u(1)
if(pic_deblocking_filter_override_flag) {	
pic_deblocking_filter_disabled_flag	u(1)
if(!pic_deblocking_filter_disabled_flag) {	
pic_beta_offset_div2	se(v)
pic_tc_offset_div2	se(v)
}	
}	
}	

}	
if(sps_lmcs_enabled_flag) {	
pic_lmcs_enabled_flag	u(1)
if(pic_lmcs_enabled_flag) {	
pic_lmcs_aps_id	u(2)
if(ChromaArrayType != 0)	
pic_chroma_residual_scale_flag	u(1)
}	
}	
if(sps_scaling_list_enabled_flag) {	
pic_scaling_list_present_flag	u(1)
if(pic_scaling_list_present_flag)	
pic_scaling_list_aps_id	u(3)
}	
if(picture_header_extension_present_flag) {	
ph_extension_length	ue(v)
for(i = 0; i < ph_extension_length; i++)	
ph_extension_data_byte[i]	u(8)
}	
rbsp_trailing_bits()	
}	

[151]

[152] Among syntax elements of Table 1, syntax elements including "intra_slice" in their titles (e.g., pic_log2_diff_min_qt_min_cb_intra_slice_luma) are syntax elements being used in I slice of the corresponding picture, and syntax elements (e.g., pic_temporal_mvp_enabled_flag) related to syntax elements including "inter_slice" in their titles (e.g., pic_log2_diff_min_qt_min_cb_inter_slice, mvp, mvd, mmvd, and merge) are syntax elements being used in P slice and/or B slice of the corresponding picture.

[153] That is, the picture header includes all of syntax elements necessary for the intra-coded slice and syntax elements necessary for the inter-coded slice for every single picture. However, this is useful only with respect to the picture including mixed type slices (picture including all of the intra-coded slice and the inter-coded slice). In general, since the picture does not include the mixed type slices (i.e., the general picture includes only the intra-coded slices or only the inter-coded slices), it is unnecessary to perform signaling of all data (syntax elements being used in the intra-coded slice and syntax elements being used in the inter-coded slice).

[154] The following drawings have been prepared to explain a detailed example of the

present document. Since the name of a detailed device or the name of detailed signal/information is exemplarily presented, the technical features of the present document are not limited to the detailed names used in the following drawing.

[155] The present document provides the following methods in order to solve the above-described problem. Items of each method may be individually applied, or may be applied in combination.

[156] 1. A flag in picture header to specify whether syntax elements that are needed only by intra coded slices are present in the picture header may be signaled. The flag may be referred to as `intra_signaling_present_flag`.

[157] a) When `intra_signaling_present_flag` is equal to 1, syntax elements that are needed by intra coded slices are present in the picture header. Likewise, when `intra_signaling_present_flag` is equal to 0, syntax elements that are needed by intra coded slices are not present in the picture header.

[158] b) The value of `intra_signaling_present_flag` in a picture header shall be equal to 1 on the picture associated with the picture header has at least one intra coded slice.

[159] c) The value of `intra_signaling_present_flag` in a picture header may be equal to 1 even when the picture associated with the picture header does not have intra coded slice.

[160] d) When a picture has one or more subpicture(s) containing intra coded slices only and it is anticipated that one or more of the subpicture(s) may be extracted and merged with subpictures which contains one or more inter coded slices, the value of `intra_signaling_present_flag` should be set equal to 1.

[161] 2. A flag in picture header to specify whether syntax elements that are needed only by inter coded slices are present in the picture header may be signaled. The flag may be referred to as `inter_signaling_present_flag`.

[162] a) When `inter_signaling_present_flag` is equal to 1, syntax elements that are needed by

inter coded slices are present in the picture header. Likewise, when `inter_signaling_present_flag` is equal to 0, syntax elements that are needed by inter coded slices are not present in the picture header.

[163] b) The value of `inter_signaling_present_flag` in a picture header shall be equal to 1 on the picture associated with the picture header has at least one inter coded slice.

[164] c) The value of `inter_signaling_present_flag` in a picture header may be equal to 1 even when the picture associated with the picture header does not have inter coded slice.

[165] d) When a picture has one or more subpicture(s) containing inter coded slices only and it is anticipated that one or more of the subpicture(s) may be extracted and merged with subpictures which contains one or more intra coded slices, the value of `inter_signaling_present_flag` should be set equal to 1.

[166] 3. The above flags (`intra_signaling_present_flag` and `inter_signaling_present_flag`) may be signaled in other parameter set such as picture parameter set (PPS) instead of in picture header).

[167] 4. Another alternative for signaling the above flags can be as follow.

[168] a) Two variables `IntraSignalingPresentFlag` and `InterSignalingPresentFlag` which specify whether syntax elements needed by intra coded slices and syntax element needed by inter coded slices, respectively, present in the picture header or not may be defined.

[169] b) A flag called `mixed_slice_types_present_flag` in the picture header may be signaled.

When `mixed_slice_types_present_flag` is equal to 1, the value of `IntraSignalingPresentFlag` and `InterSignalingPresentFlag` are set to be equal to 1.

[170] c) When `mixed_slice_types_present_flag` is equal to 0, additional flag called `intra_slice_only_flag` may be signaled in the picture header and the following applies. If `intra_slice_only_flag` is equal to 1, the value of `IntraSignalingPresentFlag` is set equal to 1 and the value of `InterSignalingPresentFlag` is set equal to 0. Otherwise, the value of

IntraSignalingPresentFlag is set equal to 0 and the value of InterSignalingPresentFlag is set equal to 1.

[171] 5. A fixed length syntax element in picture header, which may be called slice_types_idc, which specifies the following information may be signaled.

5 **[172]** a) Whether the picture associated with the picture header contain intra coded slices only. For this type, the value of slice_types_idc may be set equal to 0.

[173] b) Whether the picture associated with the picture header contain inter coded slices only. The value of slice_types_idc may be set equal to 1.

[174] c) Whether the picture associated with the picture header may contain intra coded
10 slices and inter coded slices. The value of slice_types_idc may be set equal to 2.

[175] Note that when slice_types_idc has value equal to 2, it is still possible that the picture contains intra coded slices only or inter coded slices only.

[176] d) Other values of slice_types_idc may be reserved for future use.

[177] 6. For slice_types_idc semantics in a picture header, the following constraints may be
15 further specified.

[178] a) When the picture associated with the picture header has one or more intra coded slice, the value of slice_types_idc shall not be equal to 1.

[179] b) When the picture associated with the picture header has one or more inter coded slice, the value of slice_types_idc shall not be equal to 0.

20 **[180]** 7. slice_types_idc may be signaled in other parameter set such as picture parameter set (PPS) instead of in picture header.

[181] As an embodiment, the encoding apparatus and the decoding apparatus may use the following Table 2 and Table 3 as the syntax and semantics of the picture header based on the methods of 1 and 2 as described above.

25 **[182]** [Table 2]

picture_header_rbsp() {	Descriptor
...	
intra_signalling_present_flag	u(1)
inter_signalling_present_flag	u(1)
if(partition_constraints_override_enabled_flag) {	
partition_constraints_override_flag	ue(v)
if(partition_constraints_override_flag) {	
if(intra_signalling_present_flag){	
pic_log2_diff_min_qt_min_cb_intra_slice_luma	ue(v)
pic_max_mtt_hierarchy_depth_intra_slice_luma	ue(v)
if(pic_max_mtt_hierarchy_depth_intra_slice_luma != 0) {	
pic_log2_diff_max_bt_min_qt_intra_slice_luma	ue(v)
pic_log2_diff_max_tt_min_qt_intra_slice_luma	ue(v)
}	
if(qtbtt_dual_tree_intra_flag) {	
pic_log2_diff_min_qt_min_cb_intra_slice_chroma	ue(v)
pic_max_mtt_hierarchy_depth_intra_slice_chroma	ue(v)
if(pic_max_mtt_hierarchy_depth_intra_slice_chroma != 0) {	
pic_log2_diff_max_bt_min_qt_intra_slice_chroma	ue(v)
pic_log2_diff_max_tt_min_qt_intra_slice_chroma	ue(v)
}	
}	
}	
if(inter_signalling_preset_flag){	
pic_log2_diff_min_qt_min_cb_inter_slice	ue(v)
pic_max_mtt_hierarchy_depth_inter_slice	ue(v)
if(pic_max_mtt_hierarchy_depth_inter_slice != 0) {	
pic_log2_diff_max_bt_min_qt_inter_slice	ue(v)
pic_log2_diff_max_tt_min_qt_inter_slice	ue(v)
}	
}	
}	
}	
if(intra_signalling_present_flag){	
if(cu_qp_delta_enabled_flag)	
pic_cu_qp_delta_subdiv_intra_slice	ue(v)
if(pps_cu_chroma_qp_offset_list_enabled_flag)	
pic_cu_chroma_qp_offset_subdiv_intra_slice	ue(v)
}	
if(inter_signalling_present_flag){	
if(cu_qp_delta_enabled_flag)	
pic_cu_qp_delta_subdiv_inter_slice	ue(v)
if(pps_cu_chroma_qp_offset_list_enabled_flag)	
pic_cu_chroma_qp_offset_subdiv_inter_slice	ue(v)
if(sps_temporal_mvp_enabled_flag)	
pic_temporal_mvp_enabled_flag	u(1)

[183]

if(!pps_mvd_l1_zero_idc)	
mvd_l1_zero_flag	u(1)
if(!pps_six_minus_max_num_merge_cand_plus1)	
pic_six_minus_max_num_merge_cand	ue(v)
if(sps_affine_enabled_flag)	
pic_five_minus_max_num_subblock_merge_cand	ue(v)
if(sps_fpel_mmvd_enabled_flag)	
pic_fpel_mmvd_enabled_flag	u(1)
if(sps_bdof_pic_present_flag)	
pic_disable_bdof_flag	u(1)
if(sps_dmvr_pic_present_flag)	
pic_disable_dmvr_flag	u(1)
if(sps_prof_pic_present_flag)	
pic_disable_prof_flag	u(1)
if(sps_triangle_enabled_flag && MaxNumMergeCand >= 2 && !pps_max_num_merge_cand_minus_max_num_triangle_cand_minus1)	
pic_max_num_merge_cand_minus_max_num_triangle_cand	ue(v)
}	
if(sps_ibc_enabled_flag)	
pic_six_minus_max_num_ibc_merge_cand	ue(v)
if(sps_joint_cbr_enabled_flag)	
pic_joint_cbr_sign_flag	u(1)
if(sps_sao_enabled_flag) {	
pic_sao_enabled_present_flag	u(1)
if(pic_sao_enabled_present_flag) {	
pic_sao_luma_enabled_flag	u(1)
if(ChromaArrayType != 0)	
pic_sao_chroma_enabled_flag	u(1)
}	
}	
...	
}	

[184]

[185] [Table 3]

intra_signalling_present_flag equal to 1 specifies that syntax elements used only by intra coded slices are present in the picture header. intra_signalling_present_flag equal to 0 specifies that syntax elements used only by intra coded slices are not present in the picture header. It is a requirement of bitstream conformance that, the value of intra_signalling_present_flag shall be equal to 1 when the picture associated with the picture header has one or more slice with slice_type equal to I. inter_signalling_present_flag equal to 1 specifies that syntax elements used only by inter coded slices are present in the picture header. inter_signalling_present_flag equal to 0 specifies that syntax elements used only by inter coded slices are not present in the picture header. It is a requirement of bitstream conformance that, the value of inter_signalling_present_flag shall be equal to 1 when the picture associated with the picture header has one or more slice with slice_type equal to P or B. NOTE – : The values of both intra_signalling_present_flag and inter_signalling_present_flag should be set equal to 1 in the picture header associated with picture containing one or more subpictures containing intra coded slice(s) which may be merged with one or more subpicture(s) containing inter coded slices(s)

[186] Referring to Table 2 and Table 3, if the value of **intra_signaling_present_flag** is 1, this

may represent that the syntax element being used only in the intra-coded slice is present in the picture header. If the value of the `intra_signaling_present_flag` is 0, this may represent that the syntax element being used only in the intra-coded slice is not present in the picture header. Accordingly, if the picture related to the picture header includes one or more slices having the slice type of I slice, the value of the `intra_signaling_present_flag` becomes 1. Further, if the picture related to the picture header does not include the slices having the slice type of I slice, the value of the `intra_signaling_present_flag` becomes 0.

[187] If the value of the `inter_signaling_present_flag` is 1, this may represent that the syntax element being used only in the inter-coded slice is present in the picture header. If the value of the `inter_signaling_present_flag` is 0, this may represent that the syntax element being used only in the inter-coded slice is not present in the picture header. Accordingly, if the picture related to the picture header includes one or more slices having the slice type of P slice and/or B slice, the value of the `intra_signaling_present_flag` becomes 1. Further, if the picture related to the picture header does not include the slices having the slice type of P slice and/or B slice, the value of the `intra_signaling_present_flag` becomes 0.

[188] Further, in case of the picture including one or more subpictures including the intra-coded slices which can be merged with one or more subpictures including the inter-coded slices, both the value of the `intra_signaling_present_flag` and the value of the `inter_signaling_present_flag` are set to be 1.

[189] For example, in case that only the inter-coded slices (P slice and/or B slice) are included in the current picture, the encoding apparatus may determine the value of the `inter_signaling_present_flag` as 1, and the value of the `intra_signaling_present_flag` as 0.

[190] As another example, in case that only the intra-coded slice (I slice) is included in the current picture, the encoding apparatus may determine the value of the `inter_signaling_present_flag` as 0, and the value of the `intra_signaling_present_flag` as 1.

[191] As still another example, in case that at least one inter-coded slice or at least one intra-coded slice is included in the current picture, the encoding apparatus may determine the value of the `inter_signaling_present_flag` and the value of the `intra_signaling_present_flag` as 1 in all.

[192] In case that the value of the `intra_signaling_present_flag` is determined as 0, the encoding apparatus may generate image information in which the syntax elements necessary for the intra slice are excluded or omitted, and only the syntax elements necessary for the inter slice is included in the picture header. If the value of the `inter_signaling_present_flag` is determined as 0, the encoding apparatus may generate the image information in which the syntax elements necessary for the inter slice are excluded or omitted, and only the syntax elements necessary for the intra slice is included in the picture header.

[193] If the value of the `inter_signaling_present_flag` obtained from the picture header in the image information is 1, the decoding apparatus may determine that at least one inter-coded slice is included in the corresponding picture, and may parse the syntax elements necessary for the intra prediction from the picture header. If the value of the `inter_signaling_present_flag` is 0, the decoding apparatus may determine that only the intra-coded slice is included in the corresponding picture, and may parse the syntax elements necessary for the intra prediction from the picture header. If the value of the `intra_signaling_present_flag` obtained from the picture header in the image information is 1, the decoding apparatus may determine that at least one intra-coded slice is included in the corresponding picture, and may parse the syntax elements necessary for the intra prediction from the picture header. If the value of the `intra_signaling_present_flag` is 0, the decoding apparatus may determine that only the inter-coded slice is included in the corresponding picture, and may parse the syntax elements necessary for the inter prediction from the picture header.

[194] As another embodiment, the encoding apparatus and the decoding apparatus may use the following Table 4 and Table 5 as the syntax and semantics of the picture header based on

the above methods of 5 and 6.

[195] [Table 4]

picture_header_rbsp() {	Descriptor
...	
slice_type_idc	u(1)
if(partition_constraints_override_enabled_flag) {	
partition_constraints_override_flag	ue(v)
if(partition_constraints_override_flag) {	
if(slice_type_idc != 1){	
pic_log2_diff_min_qt_min_cb_intra_slice_luma	ue(v)
pic_max_mtt_hierarchy_depth_intra_slice_luma	ue(v)
if(pic_max_mtt_hierarchy_depth_intra_slice_luma != 0){	
pic_log2_diff_max_bt_min_qt_intra_slice_luma	ue(v)
pic_log2_diff_max_tt_min_qt_intra_slice_luma	ue(v)
}	
if(qtbtt_dual_tree_intra_flag) {	
pic_log2_diff_min_qt_min_cb_intra_slice_chroma	ue(v)
pic_max_mtt_hierarchy_depth_intra_slice_chroma	ue(v)
if(pic_max_mtt_hierarchy_depth_intra_slice_chroma != 0){	
pic_log2_diff_max_bt_min_qt_intra_slice_chroma	ue(v)
pic_log2_diff_max_tt_min_qt_intra_slice_chroma	ue(v)
}	
}	
}	
if(slice_type_idc != 0){	
pic_log2_diff_min_qt_min_cb_inter_slice	ue(v)
pic_max_mtt_hierarchy_depth_inter_slice	ue(v)
if(pic_max_mtt_hierarchy_depth_inter_slice != 0){	
pic_log2_diff_max_bt_min_qt_inter_slice	ue(v)
pic_log2_diff_max_tt_min_qt_inter_slice	ue(v)
}	
}	
}	
}	

[196]

[197]

if(slice_type_idc != 1){	
if(cu_qp_delta_enabled_flag)	
pic_cu_qp_delta_subdiv_intra_slice	ue(v)
if(pps_cu_chroma_qp_offset_list_enabled_flag)	
pic_cu_chroma_qp_offset_subdiv_intra_slice	ue(v)
}	
if(slice_type_idc != 0){	
if(cu_qp_delta_enabled_flag)	
pic_cu_qp_delta_subdiv_inter_slice	ue(v)
if(pps_cu_chroma_qp_offset_list_enabled_flag)	
pic_cu_chroma_qp_offset_subdiv_inter_slice	ue(v)
if(sps_temporal_mvp_enabled_flag)	
pic_temporal_mvp_enabled_flag	u(1)
if(!pps_mvd_l1_zero_idc)	
mvd_l1_zero_flag	u(1)
if(!pps_six_minus_max_num_merge_cand_plus1)	
pic_six_minus_max_num_merge_cand	ue(v)
if(sps_affine_enabled_flag)	
pic_five_minus_max_num_subblock_merge_cand	ue(v)
if(sps_fpel_mmvd_enabled_flag)	

[198]

pic_fpel_mmvd_enabled_flag	u(1)
if(sps_bdof_pic_present_flag)	
pic_disable_bdof_flag	u(1)
if(sps_dmvr_pic_present_flag)	
pic_disable_dmvr_flag	u(1)
if(sps_prof_pic_present_flag)	
pic_disable_prof_flag	u(1)
if(sps_triangle_enabled_flag && MaxNumMergeCand >= 2 && !pps_max_num_merge_cand_minus_max_num_triangle_cand_minus1)	
pic_max_num_merge_cand_minus_max_num_triangle_cand	ue(v)
}	
if(sps_ibc_enabled_flag)	
pic_six_minus_max_num_ibc_merge_cand	ue(v)
if(sps_joint_cbr_enabled_flag)	
pic_joint_cbr_sign_flag	u(1)
if(sps_sao_enabled_flag) {	
pic_sao_enabled_present_flag	u(1)
if(pic_sao_enabled_present_flag) {	
pic_sao_luma_enabled_flag	u(1)
if(ChromaArrayType != 0)	
pic_sao_chroma_enabled_flag	u(1)
}	
}	
...	
}	

[199] [Table 5]

slice_types_idc equal to 0 specifies that all slices of the picture associated with the picture header are I slices. **slice_types_idc** equal to 1 specifies that all slices of the picture associated with the picture header are either P or B slices. **slice_types_idc** equal to 2 specifies that slices of the picture associated with the picture header may be I, P, and / or B slices. The value **slice_types_idc** equal to 3 is reserved.

It is a requirement of bitstream conformance that, the value of **slice_types_idc** shall be not be equal to either 0 or 1 when the picture associated with the picture header has intra coded slice(s) and inter coded slice(s).

NOTE -- : The value **slice_types_idc** should be set equal to 2 in the picture header associated with picture containing one or more subpictures containing intra coded slice(s) which may be merged with one or more subpicture(s) containing inter coded slices(s)

[200] Referring to Table 4 and Table 5, if the value of **slice_types_idc** is 0, this represents that the type of all slices in the picture related to the picture header is I slice. If the value of the **slice_types_idc** is 1, this represents that the type of all slices in the picture related to the picture header is P or B slice. If the value of the **slice_types_idc** is 2, this represents that the slice type of the slices in the picture related to the picture header is I, P, and/or B slice.

[201] For example, if only the intra-coded slice is included in the current picture, the encoding apparatus may determine the value of the **slice_types_idc** as 0, and may include only the syntax elements necessary for decoding of the intra slice in the picture header. That is, in this case, the syntax elements necessary for the inter slice are not included in the picture header.

[202] As another example, if only the inter-coded slice is included in the current picture, the encoding apparatus may determine the value of the **slice_types_idc** as 1, and may include only the syntax elements necessary for decoding of the inter slice in the picture header. That is, in this case, the syntax elements necessary for the intra slice is not included in the picture header.

[203] As still another example, if at least one inter-coded slice and at least one intra-coded slice are included in the current picture, the encoding apparatus may determine the value of the **slice_types_idc** as 2, and may include all of the syntax elements necessary for the decoding of the inter slice and the syntax elements necessary for the decoding of the intra slice in the picture header.

[204] If the value of the **slice_types_idc** obtained from the picture header in the image information is 0, the decoding apparatus may determine that only the intra-coded slice is

included in the corresponding picture, and may parse the syntax elements necessary for the decoding of the intra-coded slice from the picture header. If the value of the slice_types_idc is 1, the decoding apparatus may determine that only the inter-coded slice is included in the corresponding picture, and may parse the syntax elements necessary for the decoding of the inter-coded slice from the picture header. If the value of the slice_types_idc is 2, the decoding apparatus may determine that at least one intra-coded slice and at least one inter-coded slice are included in the corresponding picture, and may parse the syntax elements necessary for the decoding of the intra-coded slice and the syntax elements necessary for the decoding of the inter-coded slice from the picture header.

[205] As still another embodiment, the encoding apparatus and the decoding apparatus may use one flag representing whether the picture includes the intra- and inter-coded slices. If the flag is true, that is, if the value of the flag is 1, all of the intra slice and the inter slice may be included in the corresponding picture. IN this case, the following Table 6 and Table 7 may be used as the syntax and the semantics of the picture header.

[206] [Table 6]

picture_header_rbsp() {	Descriptor
...	
mixed_slice_signalling_present_flag	u(1)
if(!mixed_slice_signalling_present_flag)	
intra_slice_only_flag	u(1)
if(partition_constraints_override_enabled_flag) {	
partition_constraints_override_flag	ue(v)
if(partition_constraints_override_flag) {	
if(IntraSignallingPresentFlag){	
pic_log2_diff_min_qt_min_cb_intra_slice_luma	ue(v)
pic_max_mtt_hierarchy_depth_intra_slice_luma	ue(v)
if(pic_max_mtt_hierarchy_depth_intra_slice_luma != 0) {	
pic_log2_diff_max_bt_min_qt_intra_slice_luma	ue(v)
pic_log2_diff_max_tt_min_qt_intra_slice_luma	ue(v)
}	
if(qtbtt_dual_tree_intra_flag) {	
pic_log2_diff_min_qt_min_cb_intra_slice_chroma	ue(v)
pic_max_mtt_hierarchy_depth_intra_slice_chroma	ue(v)
if(pic_max_mtt_hierarchy_depth_intra_slice_chroma != 0) {	
pic_log2_diff_max_bt_min_qt_intra_slice_chroma	ue(v)
pic_log2_diff_max_tt_min_qt_intra_slice_chroma	ue(v)

}	
}	
}	
if(InterSignallingPresentFlag){	
pic_log2_diff_min_qt_min_cb_inter_slice	ue(v)
pic_max_mtt_hierarchy_depth_inter_slice	ue(v)
if(pic_max_mtt_hierarchy_depth_inter_slice != 0) {	
pic_log2_diff_max_bt_min_qt_inter_slice	ue(v)
pic_log2_diff_max_tt_min_qt_inter_slice	ue(v)
}	
}	
}	
if(IntraSignallingPresentFlag){	
if(cu_qp_delta_enabled_flag)	
pic_cu_qp_delta_subdiv_intra_slice	ue(v)
if(pps_cu_chroma_qp_offset_list_enabled_flag)	
pic_cu_chroma_qp_offset_subdiv_intra_slice	ue(v)
}	
if(InterSignallingPresentFlag){	
if(cu_qp_delta_enabled_flag)	
pic_cu_qp_delta_subdiv_inter_slice	ue(v)
if(pps_cu_chroma_qp_offset_list_enabled_flag)	
pic_cu_chroma_qp_offset_subdiv_inter_slice	ue(v)
if(sps_temporal_mvp_enabled_flag)	
pic_temporal_mvp_enabled_flag	u(1)

[207]

if(!pps_mvd_l1_zero_idc)	
mvd_l1_zero_flag	u(1)
if(!pps_six_minus_max_num_merge_cand_plus1)	
pic_six_minus_max_num_merge_cand	ue(v)
if(sps_affine_enabled_flag)	
pic_five_minus_max_num_subblock_merge_cand	ue(v)
if(sps_fpel_mmvd_enabled_flag)	
pic_fpel_mmvd_enabled_flag	u(1)
if(sps_bdof_pic_present_flag)	
pic_disable_bdof_flag	u(1)
if(sps_dmvr_pic_present_flag)	
pic_disable_dmvr_flag	u(1)
if(sps_prof_pic_present_flag)	
pic_disable_prof_flag	u(1)
if(sps_triangle_enabled_flag && MaxNumMergeCand >= 2 && !pps_max_num_merge_cand_minus_max_num_triangle_cand_minus1)	
pic_max_num_merge_cand_minus_max_num_triangle_cand	ue(v)
}	
if(sps_ibc_enabled_flag)	
pic_six_minus_max_num_ibc_merge_cand	ue(v)

[208]

if(sps_joint_cbr_enabled_flag)	
pic_joint_cbr_sign_flag	u(1)
if(sps_sao_enabled_flag) {	
pic_sao_enabled_present_flag	u(1)
if(pic_sao_enabled_present_flag) {	
pic_sao_luma_enabled_flag	u(1)
if(ChromaArrayType != 0)	
pic_sao_chroma_enabled_flag	u(1)
}	
}	
...	
}	

[209]

[210] [Table 7]

mixed_slice_signalling_present_flag equal to 1 specifies that the picture associated with the picture header may have one or more slices with different types. **mixed_slice_signalling_present_flag** equal to 0 specifies that the picture associated with the picture header contains data pertaining to only a single slice type.

The variables **InterSignallingPresentFlag** and **IntraSignallingPresentFlag** are defined to specify whether syntax elements needed by intra coded slices and syntax elements needed by inter coded slices respectively, are present in the picture or not. When **mixed_slice_signalling_present_flag** is equal to 1, **IntraSignallingPresentFlag** and **InterSignallingPresentFlag** are set to be equal to 1.

intra_slice_only_flag equal to 1 specifies that the value of **IntraSignallingPresentFlag** is set to be equal to 1 and the value of **InterSignallingPresentFlag** is set to be equal to 0. **intra_slice_only_flag** equal to 0 specifies that **IntraSignallingPresentFlag** is set equal to 0 and the value of **InterSignallingPresentFlag** is set equal to 1.

It is a requirement of bitstream conformance that, the value of **IntraSignallingPresentFlag** shall be equal to 1 when the picture associated with the picture header has one or more slice with **slice_type** equal to I.

It is a requirement of bitstream conformance that, the value of **InterSignallingPresentFlag** shall be equal to 1 when the picture associated with the picture header has one or more slice with **slice_type** equal to P or B.

NOTE – : The values of both intra signalling present flag and inter signalling present flag should be set equal to 1 in the picture header associated with picture containing one or more subpictures containing intra coded slice(s) which may be merged with one or more subpicture(s) containing inter coded slices(s)

[211] Referring to Table 6 and Table 7, if the value of **mixed_slice_signaling_present_flag** is 1, this may represent that the picture related to the corresponding picture header has one or more slices having different types. If the value of the **mixed_slice_signaling_present_flag** is 0, this may mean that the picture related to the corresponding picture header includes data related to only a single slice type.

[212] Variables **InterSignalingPresentFlag** and **IntraSignalingPresentFlag** represent whether the syntax element necessary for the intra-coded slice and the syntax element necessary for the inter-coded slice are present in the corresponding picture header, respectively. If the value of the **mixed_slice_signaling_present_flag** is 1, the values of the **IntraSignalingPresentFlag** and the **InterSignalingPresentFlag** are set to 1.

[213] If the value of **intra_slice_only_flag** is set to 1, it represents that the value of **IntraSignalingPresentFlag** is set to 1, and the value of the **InterSignalingPresentFlag** is set to 0. If the value of the **intra_slice_only_flag** is 0, it represents that the value of the **IntraSignalingPresentFlag** is set to 0, and the value of the **InterSignalingPresentFlag** is set to 1.

[214] If the picture related to the picture header has one or more slices having the slice type of I slice, the value of the **IntraSignalingPresentFlag** is set to 1. If the picture related to the

picture header has one or more slices having the slice type of P or B slice, the value of the InterSignalingPresentFlag is set to 1.

[215] For example, if only the intra-coded slice is included in the current picture, the encoding apparatus may determine the value of mixed_slice_signaling_present_flag as 0, may
5 determine the value of intra_slice_only_flag as 1, may determine the value of IntraSignalingPresentFlag as 1, and may determine the value of InterSignalingPresentFlag as 0.

[216] As another example, if only the inter-coded slice is included in the current picture, the encoding apparatus may determine the value of the mixed_slice_signaling_present_flag as 0,
10 may determine the value of the intra_slice_only_flag as 0, may determine the value of the IntraSignalingPresentFlag as 0, and may determine the value of the InterSignalingPresentFlag as 1.

[217] As still another example, if at least one intra-coded slice and at least one inter-coded slice are included in the current picture, the encoding apparatus may determine the values of
15 the mixed_slice_signaling_present_flag, the IntraSignalingPresentFlag, and the InterSignalingPresentFlag as 1, respectively.

[218] If the value of the mixed_slice_signaling_present_flag obtained from the picture header in the image information is 0, the decoding apparatus may determine that only the intra-coded slice or inter-coded slice is included in the corresponding picture. In this case, if the
20 value of the intra_slice_only_flag obtained from the picture header is 0, the decoding apparatus may parse only the syntax elements necessary for decoding of the inter-coded slice from the picture header. If the value of the intra_slice_only_flag is 1, the decoding apparatus may parse only the syntax element necessary for decoding of the intra-coded slice from the picture header.

[219] If the value of the mixed_slice_signaling_present_flag obtained from the picture
25 header in the image information is 1, the decoding apparatus may determine that at least one

intra-coded slice and at least one inter-coded slice are included in the corresponding picture, and may parse the syntax elements necessary for the decoding of the inter-coded slice and the syntax elements necessary for the decoding of the intra-coded slice from the picture header.

[220] FIGS. 12 and 13 schematically illustrate a video/image encoding method and an example of related components according to an embodiment of the present document.

[221] The video/image encoding method disclosed in FIG. 12 may be performed by the (video/image) encoding apparatus 200 disclosed in FIGS. 2 and 13. Specifically, for example, S1200 and S1210 of FIG. 12 may be performed by the predictor 220 of the encoding apparatus 200, and S1220 and S1230 may be performed by the entropy encoder 240 of the encoding apparatus 200. The video/image encoding method disclosed in FIG. 12 may include the above-described embodiments of the present document.

[222] Specifically, referring to FIGS. 12 and 13, the predictor 220 of the encoding apparatus may determine the prediction mode of the current block in the current picture (S1200). The current picture may include a plurality of slices. The predictor 220 of the encoding apparatus may generate prediction samples (predicted block) for the current block based on the prediction mode (S1210). Here, the prediction mode may include an inter prediction mode and an intra prediction mode. If the prediction mode of the current block is the inter prediction mode, the prediction samples may be generated by the inter predictor 221 of the predictor 220. If the prediction mode of the current block is the intra prediction mode, the prediction samples may be generated by the intra predictor 222 of the predictor 220.

[223] The residual processor 230 of the encoding apparatus may generate residual samples and residual information based on the prediction samples and the original picture (original block and original samples). Here, the residual information is information about the residual samples, and may include information about (quantized) transform coefficients for the residual samples.

[224] The adder (or reconstructor) of the encoding apparatus may generate reconstructed samples (reconstructed picture, reconstructed block, or reconstructed sample array) by adding the residual samples generated by the residual processor 230 and the prediction samples generated by the inter predictor 221 or the intra predictor 222 to each other.

5 **[225]** Meanwhile, the entropy encoder 240 of the encoding apparatus may generate first information representing whether information necessary for the inter prediction operation for the decoding process is present in the picture header related to the current picture or second information representing whether information necessary for the intra prediction operation for the decoding process is present in the picture header related to the current picture (S1220).

10 Here, the first information and the second information are information included in the picture header of the image information, and may correspond to the above-described `intra_signaling_present_flag`, `inter_signaling_present_flag`, `slice_type_idc`, `mixed_slice_signaling_present_flag`, `intra_slice_only_flag`, `IntraSignalingPresentFlag`, and/or `InterSignalingPresentFlag`.

15 **[226]** As an example, in case that the information necessary for the inter prediction operation for the decoding process is included in the picture header related to the current picture as the inter-coded slice is included in the current picture, the entropy encoder 240 of the encoding apparatus may determine the value of the first information as 1. Further, in case that the information necessary for the intra prediction operation for the decoding process is included in
20 the corresponding picture header as the intra-coded slice is included in the current picture, the entropy encoder 240 of the encoding apparatus may determine the value of the second information as 1. In this case, the first information may correspond to the `inter_signaling_present_flag`, and the second information may correspond to the `intra_signaling_present_flag`. The first information may be called a first flag, information on
25 whether syntax elements being used for inter slices are present in the picture header, a flag for

whether syntax elements being used for the inter slices are present in the picture header, information on whether slices in the current picture are inter slices, or a flag for whether the slices are inter slices. The second information may be called a second flag, information on whether syntax elements being used for intra slices are present in the picture header, a flag for whether syntax elements being used for the intra slices are present in the picture header, information on whether slices in the current picture are intra slices, or a flag for whether the slices are intra slices.

[227] Meanwhile, in case that only the information necessary for the intra prediction operation is included in the corresponding picture header as only the intra-coded slice is included in the picture, the entropy encoder 240 of the encoding apparatus may determine the value of the first information as 0, and may determine the value of the second information as 1. Further, in case that only the information necessary for the inter prediction operation is included in the corresponding picture header as only the inter-coded slice is included in the picture, the value of the first information may be determined as 1, and the value of the second information may be determined as 0. Accordingly, if the value of the first information is 0, all slices in the current picture may have I slice type. If the value of the second information is 0, all the slices in the current picture may have the P slice type or the B slice type. Here, the information necessary for the intra prediction operation may include the syntax element being used for decoding of the intra slice, and the information necessary for the inter prediction operation may include the syntax element being used for decoding of the inter slice.

[228] As another example, if all the slices in the current picture have the I slice type, the entropy encoder 240 of the encoding apparatus may determine the value of the information about the slice type as 0, and if all the slices in the current picture have the P slice type or the B slice type, the entropy encoder 240 of the encoding apparatus may determine the value of the information about the slice type as 1. If all the slices in the current picture have I slice type,

P slice type, and/or B slice type (i.e., the slice types of the slices in the picture are mixed), the entropy encoder 240 of the encoding apparatus may determine the value of the information about the slice type as 2. In this case, the information about the slice type may correspond to the slice_type_idc.

5 **[229]** As still another example, if all the slices in the current picture have the same slice type, the entropy encoder 240 of the encoding apparatus may determine the value of the information about the slice type as 0, and if the slices in the current picture have different slice types, the entropy encoder 240 of the encoding apparatus may determine the value of the information about the slice type as 1. In this case, the information about the slice type may correspond to
10 the mixed_slice_signaling_present_flag.

[230] If the value of the information about the slice type is determined as 0, information on whether the intra slice is included in the slices may be included in the corresponding picture header. The information on whether the intra slice is included in the slices may correspond to the intra_slice_only_flag. If all the slices in the picture have the I slice type, the entropy encoder
15 240 of the encoding apparatus may determine the value of the information on whether the intra slice is included in the slices as 1, determine the value of the information on whether the syntax elements being used for the intra slice are present in the picture header as 1, and determine the value of the information on whether the syntax elements being used for the inter slice is present in the picture header as 0. If the slice type of all the slices in the picture is the P slice and/or B
20 slice type, the entropy encoder 240 of the encoding apparatus may determine the value of the information on whether the intra slice is included in the slices as 0, determine the value of the information on whether the syntax elements being used for the intra slice are present in the picture header as 0, and determine the value of the information on whether the syntax elements being used for the inter slice is present in the picture header as 1.

25 **[231]** The entropy encoder 240 of the encoding apparatus may encode the image information

including the first information, the second information, and information about the slice type together with the residual information and prediction related information (S1230). For example, the image information may include the partitioning related information, information about the prediction mode, residual information, in-loop filtering related information, first information, second information, and information about the slice type, and may include various syntax elements thereof. As an example, the image information may include information about various parameter sets, such as an adaptation parameter set (APS), picture parameter set (PPS), sequence parameter set (SPS), or video parameter set (VPS). Further, the image information may include various pieces of information, such as a picture header syntax, picture header structure syntax, slice header syntax, and coding unit syntax. The above-described first information, second information, information about the slice type, information necessary for the intra prediction operation, and information necessary for the inter prediction operation may be included in the syntax in the picture header.

[232] The information encoded by the entropy encoder 240 of the encoding apparatus may be output in the form of a bitstream. The bitstream may be transmitted through a network or a storage medium.

[233] FIGS. 14 and 15 schematically illustrate a video/image decoding method and an example of related components according to an embodiment of the present document.

[234] The video/image decoding method disclosed in FIG. 14 may be performed by the (video/image) decoding apparatus 300 disclosed in FIGS. 3 and 15. Specifically, for example, S1400 and S1410 of FIG. 14 may be performed by the entropy decoder 310 of the decoding apparatus, S1420 may be performed by the predictor 330 of the decoding apparatus 300, and S1430 may be performed by the adder 340 of the decoding apparatus 300. The video/image decoding method disclosed in FIG. 14 may include the above-described embodiments of the present document.

[235] Referring to FIGS. 14 and 15, the entropy decoder 310 of the decoding apparatus may obtain image information from a bitstream (S1400). The image information may include a picture header related to the current picture. The current picture may include a plurality of slices.

5 [236] Meanwhile, the entropy decoder 310 of the decoding apparatus may parse, from the picture header, a first flag representing whether information necessary for an inter prediction operation for a decoding process is present in the picture header related to the current picture and/or a second flag representing whether information necessary for an intra prediction operation for the decoding process is present in the picture header related to the current picture
10 (S1410). Here, the first flag and the second flag may correspond to the above-described `intra_signaling_present_flag`, `inter_signaling_present_flag`, `slice_type_idc`, `mixed_slice_signaling_present_flag`, `intra_slice_only_flag`, `IntraSignalingPresentFlag`, and/or `InterSignalingPresentFlag`. The entropy decoder 310 of the decoding apparatus may parse syntax elements included in the picture header of the image information based on any one
15 picture header syntax in Tables 2, 4, and 6 as described above.

[237] The decoding apparatus may generate prediction samples by performing at least one of intra prediction or inter prediction for slices in the current picture based on the first flag, the second flag, and the information about the slice type (S1420).

[238] Specifically, the entropy decoder 310 of the decoding apparatus may parse (or obtain)
20 at least one of the information necessary for the intra prediction operation and or the information necessary for the inter prediction operation for the decoding process from the picture header related to the current picture based on the first flag, the second flag, and/or the information about the slice type. The predictor 330 of the decoding apparatus may generate prediction samples by performing intra prediction and/or inter prediction based on at least one
25 of the information necessary for the intra prediction operation or the information for the inter

prediction. Here, the information necessary for the intra prediction operation may include a syntax element being used for decoding of an intra slice, and the information necessary for the inter prediction operation may include a syntax element being used for decoding of an inter slice.

5 **[239]** As an example, if the value of the first flag is 0, the entropy decoder 310 of the decoding apparatus may determine (or decide) that the syntax elements being used for the inter prediction are not present in the picture header, and may parse only the information necessary for the intra prediction operation from the picture header. If the value of the first flag is 1, the entropy decoder 310 of the decoding apparatus may determine (or decide) that the syntax
10 elements being used for the inter prediction are present in the picture header, and may parse the information necessary for the inter prediction operation from the picture header. In this case, the first flag may correspond to the `inter_signaling_present_flag`.

[240] Further, if the value of the second flag is 0, the entropy decoder 310 of the decoding apparatus may determine (or decide) that the syntax elements being used for the intra prediction
15 are not present in the picture header, and may parse only the information necessary for the inter prediction operation from the picture header. If the value of the second flag is 1, the entropy decoder 310 of the decoding apparatus may determine (or decide) that the syntax elements being used for the intra prediction are present in the picture header, and may parse the information necessary for the intra prediction operation from the picture header. In this case,
20 the second flag may correspond to the `intra_signaling_present_flag`.

[241] If the value of the first flag is 0, the decoding apparatus may determine that all slices in the current picture have the type of I slice. If the value of the first flag is 1, the decoding apparatus may determine that 0 or more slices in the current picture have the type of P slice or B slice. In other words, if the value of the first flag is 1, the slice having the type of P slice or
25 B slice may be included or may not be included in the current picture.

[242] Further, if the value of the second flag is 0, the decoding apparatus may determine that all slices in the current picture have the type of P slice or B slice. If the value of the second flag is 1, the decoding apparatus may determine that 0 or more slices in the current picture have the type of I slice. In other words, if the value of the second flag is 1, the slice having the type of I slice may be included or may not be included in the current picture.

[243] As another example, if the value of the information about the slice type is 0, the entropy decoder 310 of the decoding apparatus may determine that all slices in the current picture have I slice type, and may parse only the information necessary for the intra prediction operation. If the information about the slice type is 1, the entropy decoder 310 of the decoding apparatus may determine that all slices in the corresponding picture have the P slice type or the B slice type, and may parse only the information necessary for the inter prediction operation from the picture header. If the value of the information for the slice type is 2, the entropy decoder 310 of the decoding apparatus may determine that the slices in the corresponding picture have the slice type in which the I slice type, the P slice type, and/or the B slice type are mixed, and may parse all of the information necessary for the inter prediction operation and the information necessary for the intra prediction operation from the picture header. In this case, the information about the slice type may correspond to the slice_type_idc.

[244] As still another example, the entropy decoder 310 of the decoding apparatus may determine that all slices in the current picture have the same slice type if the value of the information about the slice type is determined as 0, and may determine that slices in the current picture have different slice types if the value of the information about the slice type is determined as 1. In this case, the information about the slice type may correspond to the mixed_slice_signalling_present_flag.

[245] If the value of the information about the slice type is determined as 0, the entropy decoder 310 of the decoding apparatus may parse information on whether the intra slice is

included in the slices from the picture header. The information on whether the intra slice is included in the slices may correspond to the `intra_slice_only_flag` as described above. If the information on whether the intra slice is included in the slices is 1, all the slices in the picture may have I slice type.

5 **[246]** If the value of the information on whether the intra slice is included in the slices is 1, the entropy decoder 310 of the encoding apparatus may parse only the information necessary for the intra prediction operation from the picture header. If the value of the information on whether the intra slice is included in the slices is 0, the entropy decoder 310 of the decoding apparatus may parse only the information necessary for the inter prediction operation from the
10 picture header.

[247] If the value of the information about the slice type is 1, the entropy decoder 310 of the decoding apparatus may parse all of the information necessary for the inter prediction operation and the information necessary for the intra prediction operation from the picture header.

[248] Meanwhile, the residual processor 320 of the decoding apparatus may generate
15 residual samples based on the residual information obtained by the entropy decoder 310.

[249] The adder 340 of the decoding apparatus may generate reconstructed samples based on the prediction samples generated by the predictor 330 and the residual samples generated by the residual processor 320 (1430). Further, the adder 340 of the decoding apparatus may generate a reconstructed picture (reconstructed block) based on the reconstructed samples.

20 **[250]** Thereafter, as needed, in order to enhance the subjective/objective picture quality, in-loop filtering procedures, such as deblocking filtering, SAO, and/or AFL procedures, may be applied to the reconstructed picture.

[251] Although methods have been described on the basis of a flowchart in which steps or blocks are listed in sequence in the above-described embodiments, the steps of the present
25 disclosure are not limited to a certain order, and a certain step may be performed in a different

step or in a different order or concurrently with respect to that described above. Further, it will be understood by those ordinary skilled in the art that the steps of the flowcharts are not exclusive, and another step may be included therein or one or more steps in the flowchart may be deleted without exerting an influence on the scope of the present disclosure.

5 **[252]** The aforementioned method according to the present disclosure may be in the form of software, and the encoding apparatus and/or decoding apparatus according to the present disclosure may be included in a device for performing image processing, for example, a TV, a computer, a smart phone, a set-top box, a display device, or the like.

10 **[253]** When the embodiments of the present disclosure are implemented by software, the aforementioned method may be implemented by a module (process or function) which performs the aforementioned function. The module may be stored in a memory and executed by a processor. The memory may be installed inside or outside the processor and may be connected to the processor via various well-known means. The processor may include Application-Specific Integrated Circuit (ASIC), other chipsets, a logical circuit, and/or a data
15 processing device. The memory may include a Read-Only Memory (ROM), a Random Access Memory (RAM), a flash memory, a memory card, a storage medium, and/or other storage device. In other words, the embodiments according to the present disclosure may be implemented and executed on a processor, a micro-processor, a controller, or a chip. For example, functional units illustrated in the respective figures may be implemented and executed
20 on a computer, a processor, a microprocessor, a controller, or a chip. In this case, information on implementation (for example, information on instructions) or algorithms may be stored in a digital storage medium.

25 **[254]** In addition, the decoding apparatus and the encoding apparatus to which the embodiment(s) of the present disclosure is applied may be included in a multimedia broadcasting transceiver, a mobile communication terminal, a home cinema video device, a

digital cinema video device, a surveillance camera, a video chat device, and a real time communication device such as video communication, a mobile streaming device, a storage medium, a camcorder, a video on demand (VoD) service provider, an Over The Top (OTT) video device, an internet streaming service provider, a 3D video device, a Virtual Reality (VR) device, an Augment Reality (AR) device, an image telephone video device, a vehicle terminal (for example, a vehicle (including an autonomous vehicle) terminal, an airplane terminal, or a ship terminal), and a medical video device; and may be used to process an image signal or data. For example, the OTT video device may include a game console, a Blu-ray player, an Internet-connected TV, a home theater system, a smartphone, a tablet PC, and a Digital Video Recorder (DVR).

[255] In addition, the processing method to which the embodiment(s) of the present disclosure is applied may be produced in the form of a program executed by a computer and may be stored in a computer-readable recording medium. Multimedia data having a data structure according to the embodiment(s) of the present disclosure may also be stored in the computer-readable recording medium. The computer readable recording medium includes all kinds of storage devices and distributed storage devices in which computer readable data is stored. The computer-readable recording medium may include, for example, a Blu-ray disc (BD), a universal serial bus (USB), a ROM, a PROM, an EPROM, an EEPROM, a RAM, a CD-ROM, a magnetic tape, a floppy disk, and an optical data storage device. The computer-readable recording medium also includes media embodied in the form of a carrier wave (for example, transmission over the Internet). In addition, a bitstream generated by the encoding method may be stored in the computer-readable recording medium or transmitted through a wired or wireless communication network.

[256] In addition, the embodiment(s) of the present disclosure may be embodied as a computer program product based on a program code, and the program code may be executed

on a computer according to the embodiment(s) of the present disclosure. The program code may be stored on a computer-readable carrier.

[257] FIG. 16 represents an example of a contents streaming system to which the embodiment of the present disclosure may be applied.

5 **[258]** Referring to FIG. 16, the content streaming system to which the embodiments of the present disclosure is applied may generally include an encoding server, a streaming server, a web server, a media storage, a user device, and a multimedia input device.

[259] The encoding server functions to compress to digital data the contents input from the multimedia input devices, such as the smart phone, the camera, the camcorder and the like, to
10 generate a bitstream, and to transmit it to the streaming server. As another example, in a case in which the multimedia input device, such as, the smart phone, the camera, the camcorder or the like, directly generates a bitstream, the encoding server may be omitted.

[260] The bitstream may be generated by an encoding method or a bitstream generation method to which the embodiments of the present disclosure is applied. And the streaming
15 server may temporarily store the bitstream in a process of transmitting or receiving the bitstream.

[261] The streaming server transmits multimedia data to the user equipment on the basis of a user's request through the web server, which functions as an instrument that informs a user of what service there is. When the user requests a service which the user wants, the web server
20 transfers the request to the streaming server, and the streaming server transmits multimedia data to the user. In this regard, the contents streaming system may include a separate control server, and in this case, the control server functions to control commands/responses between respective equipment in the content streaming system.

[262] The streaming server may receive contents from the media storage and/or the encoding
25 server. For example, in a case the contents are received from the encoding server, the contents

may be received in real time. In this case, the streaming server may store the bitstream for a predetermined period of time to provide the streaming service smoothly.

[263] For example, the user equipment may include a mobile phone, a smart phone, a laptop computer, a digital broadcasting terminal, a personal digital assistant (PDA), a portable multimedia player (PMP), a navigation, a slate PC, a tablet PC, an ultrabook, a wearable device (e.g., a watch-type terminal (smart watch), a glass-type terminal (smart glass), a head mounted display (HMD)), a digital TV, a desktop computer, a digital signage or the like.

[264] Each of servers in the contents streaming system may be operated as a distributed server, and in this case, data received by each server may be processed in distributed manner.

What is claimed is:

1. A video decoding method performed by a video decoding apparatus, the method comprising:

5 obtaining image information from a bitstream, the image information including a picture header related to a current picture, and the current picture including a plurality of slices;

parsing, from the picture header, at least one of a first flag representing whether information necessary for an inter prediction operation for a decoding process is present in the picture header or a second flag representing whether information necessary for an intra
10 prediction operation for the decoding process is present in the picture header;

generating prediction samples by performing at least one of an intra prediction or an inter prediction for the slices in the current picture based on at least one of the first flag or the second flag; and

generating reconstructed samples based on the prediction samples.

15 2. The video decoding method of claim 1, wherein the information necessary for the inter prediction operation comprises a syntax element used for decoding of an inter slice, and

the information necessary for the intra prediction operation includes a syntax element used for decoding of an intra slice.

20 3. The video decoding method of claim 1, wherein the generating the prediction samples comprises:

determining that syntax elements used for the inter prediction are present in the picture header based on that a value of the first flag is 1; and

25 determining that the syntax elements used for the inter prediction are not present in the

picture header based on that the value of the first flag is 0.

4. The video decoding method of claim 2, wherein slices in the current picture have a type of I slice based on that a value of the first flag is 0.

5

5. The video decoding method of claim 1, wherein 0 or more slices in the current picture have a type of P slice or B slice based on that a value of the first flag is 1.

6. The video decoding method of claim 1, wherein the generating the prediction samples
10 comprises:

determining that syntax elements used for the intra prediction are present in the picture header based on that a value of the second flag is 1; and

determining that the syntax elements used for the intra prediction are not present in the picture header based on that the value of the second flag is 0.

15

7. The video decoding method of claim 1, wherein slices in the current picture have a type of P slice or B slice based on that a value of the second flag is 0.

8. The video decoding method of claim 1, wherein 0 or more slices in the current picture
20 have a type of I slice based on that a value of the first flag is 1.

9. A video encoding method performed by a video encoding apparatus, the method comprising:

determining a prediction mode of a current block in a current picture, the current picture
25 including a plurality of slices;

generating prediction samples for the current block based on the prediction mode;

generating at least one of first information representing whether information necessary for an inter prediction operation for a decoding process is present in a picture header related to the current picture or second information representing whether information necessary for an intra prediction operation for the decoding process is present in the picture header based on the prediction mode; and

encoding image information including at least one of the first information or the second information,

wherein the first information and the second information are included in the picture header of the image information.

10. The video encoding method of claim 9, wherein the information necessary for the inter prediction operation comprises a syntax element used for decoding of an inter slice, and the information necessary for the intra prediction operation includes a syntax element used for decoding of an intra slice.

11. The video encoding method of claim 9, wherein a value of a first flag is 0 based on that slices in the current picture have a type of I slice.

12. The video encoding method of claim 9, wherein a value of a first flag is 1 based on that 0 or more slices in the current picture have a type of P slice or B slice.

13. The video encoding method of claim 9, wherein a value of a second flag is 0 based on that slices in the current picture have a type of P slice or B slice.

14. The video encoding method of claim 9, wherein a value of a first flag is 1 based on that 0 or more slices in the current picture have a type of I slice.

15. A computer-readable digital storage medium including information causing a
5 decoding apparatus to perform a video decoding method, the video decoding method comprising:

obtaining image information from a bitstream, the image information including a picture header related to a current picture, and the current picture including a plurality of slices;

10 parsing, from the picture header, at least one of a first flag representing whether information necessary for an inter prediction operation for a decoding process is present in the picture header or a second flag representing whether information necessary for an intra prediction operation for the decoding process is present in the picture header;

generating prediction samples by performing at least one of an intra prediction or an inter prediction for the slices in the current picture based on at least one of the first flag or the second
15 flag; and

generating reconstructed samples based on the prediction samples.