(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
02 December 2021 (02.12.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/242845 A1

(51) International Patent Classification:
H04N 19/11 (2014.01) *H04N 19/463* (2014.01)
H04N 19/154 (2014.01) *H04N 19/593* (2014.01)
H04N 19/176 (2014.01)

(21) International Application Number:
PCT/US2021/034226

(22) International Filing Date:
26 May 2021 (26.05.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/030,792 27 May 2020 (27.05.2020) US

(72) Inventors; and

(71) Applicants: **BAE, Tae Meon** [KR/US]; 11091 Sunset Hills Road, Suite #510, Reston, Virginia 20190 (US). **DINAN, Esmael Hejazi** [US/US]; 11091 Sunset Hills Road, Suite #510, Reston, Virginia 20190 (US).

(74) Agent: **NASABZADEH, Kavon** et al.; 11091 Sunset Hills Road, Suite 510, Reston, Virginia 20190 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

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

(54) Title: INTRA PREDICTION

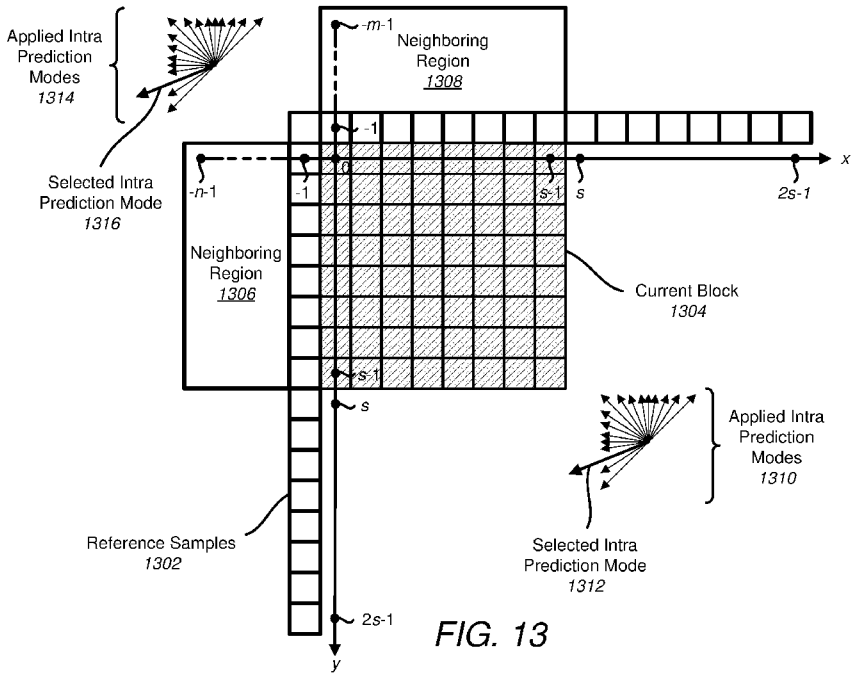


FIG. 13

(57) Abstract: A decoder generates, for each of a plurality of intra prediction modes determined based on a most probable mode (MPM) list for intra prediction of a block, a prediction of at least one neighboring region of the block. The decoder determines, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes. The decoder decodes a second intra prediction mode for the block based on the first intra prediction mode.

WO 2021/242845 A1

[Continued on next page]

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

TITLE

Intra Prediction

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 63/030,792, filed May 27, 2020, which is hereby incorporated by reference in its entirety.

BRIEF DESCRIPTION OF THE DRAWINGS

[0002] Examples of several of the various embodiments of the present disclosure are described herein with reference to the drawings.

[0003] FIG. 1 illustrates an exemplary video coding/decoding system in which embodiments of the present disclosure may be implemented.

[0004] FIG. 2 illustrates an exemplary encoder in which embodiments of the present disclosure may be implemented.

[0005] FIG. 3 illustrates an exemplary decoder in which embodiments of the present disclosure may be implemented.

[0006] FIG. 4 illustrates an example quadtree partitioning of a coding tree block (CTB) in accordance with embodiments of the present disclosure.

[0007] FIG. 5 illustrates a corresponding quadtree of the example quadtree partitioning of the CTB in FIG. 4 in accordance with embodiments of the present disclosure.

[0008] FIG. 6 illustrates example binary and ternary tree partitions in accordance with embodiments of the present disclosure.

[0009] FIG. 7 illustrates an example quadtree + multi-type tree partitioning of a CTB in accordance with embodiments of the present disclosure.

[0010] FIG. 8 illustrates a corresponding quadtree + multi-type tree of the example quadtree + multi-type tree partitioning of the CTB in FIG. 7 in accordance with embodiments of the present disclosure.

[0011] FIG. 9 illustrates an example set of reference samples determined for intra prediction of a current block being encoded or decoded in accordance with embodiments of the present disclosure.

[0012] FIG. 10A illustrates the 35 intra prediction modes supported by HEVC in accordance with embodiments of the present disclosure.

[0013] FIG. 10B illustrates the 67 intra prediction modes supported by HEVC in accordance with embodiments of the present disclosure.

[0014] FIG. 11 illustrates the current block and reference samples from FIG. 9 in a two-dimensional x, y plane in accordance with embodiments of the present disclosure.

[0015] FIG. 12 illustrates an example angular mode prediction of the current block from FIG. 9 in accordance with embodiments of the present disclosure.

[0016] FIG. 13 illustrates example neighboring regions for predicting a selected intra prediction mode for a current block being encoded by an encoder in accordance with embodiments of the present disclosure.

[0017] FIG. 14 illustrates an example angular mode prediction of the neighboring region from FIG. 13 in accordance with embodiments of the present disclosure.

[0018] FIG. 15A illustrates an example comparison of a selected intra prediction mode for at least one neighboring region and a selected intra prediction mode determined for a current block in accordance with embodiments of the present disclosure.

[0019] FIG. 15B illustrates an example comparison of a selected intra prediction mode for at least one neighboring region and a selected intra prediction mode determined for a current block in accordance with embodiments of the present disclosure.

[0020] FIG. 16 illustrates an example comparison of a selected intra prediction mode for at least one neighboring region and a selected intra prediction mode determined for a current block in accordance with embodiments of the present disclosure.

[0021] FIG. 17 illustrates a flowchart of a method for signaling an intra prediction mode for a block in accordance with embodiments of the present disclosure.

[0022] FIG. 18 illustrates a flowchart of a method for signaling an intra prediction mode for a block in accordance with embodiments of the present disclosure.

[0023] FIG. 19 illustrates a block diagram of an example computer system in which embodiments of the present disclosure may be implemented.

DETAILED DESCRIPTION

[0024] In the following description, numerous specific details are set forth in order to provide a thorough understanding of the disclosure. However, it will be apparent to those skilled in the art that the disclosure, including structures, systems, and methods, may be practiced without these specific details. The description and representation herein are the common means used by those experienced or skilled in the art to most effectively convey the substance of their work to others skilled in the art. In other instances, well-known methods, procedures, components, and circuitry have not been described in detail to avoid unnecessarily obscuring aspects of the disclosure.

[0025] References in the specification to "one embodiment," "an embodiment," "an example embodiment," etc., indicate that the embodiment described may include a particular feature,

structure, or characteristic, but every embodiment may not necessarily include the particular feature, structure, or characteristic. Moreover, such phrases are not necessarily referring to the same embodiment. Further, when a particular feature, structure, or characteristic is described in connection with an embodiment, it is submitted that it is within the knowledge of one skilled in the art to affect such feature, structure, or characteristic in connection with other embodiments whether or not explicitly described.

[0026] Also, it is noted that individual embodiments may be described as a process which is depicted as a flowchart, a flow diagram, a data flow diagram, a structure diagram, or a block diagram. Although a flowchart may describe the operations as a sequential process, many of the operations can be performed in parallel or concurrently. In addition, the order of the operations may be re-arranged. A process is terminated when its operations are completed, but could have additional steps not included in a figure. A process may correspond to a method, a function, a procedure, a subroutine, a subprogram, etc. When a process corresponds to a function, its termination can correspond to a return of the function to the calling function or the main function.

[0027] The term "computer-readable medium" includes, but is not limited to, portable or non-portable storage devices, optical storage devices, and various other mediums capable of storing, containing, or carrying instruction(s) and/or data. A computer-readable medium may include a non-transitory medium in which data can be stored and that does not include carrier waves and/or transitory electronic signals propagating wirelessly or over wired connections. Examples of a non-transitory medium may include, but are not limited to, a magnetic disk or tape, optical storage media such as compact disk (CD) or digital versatile disk (DVD), flash memory, memory or memory devices. A computer-readable medium may have stored thereon code and/or machine-executable instructions that may represent a procedure, a function, a subprogram, a program, a routine, a subroutine, a module, a software package, a class, or any combination of instructions, data structures, or program statements. A code segment may be coupled to another code segment or a hardware circuit by passing and/or receiving information, data, arguments, parameters, or memory contents. Information, arguments, parameters, data, etc. may be passed, forwarded, or transmitted via any suitable means including memory sharing, message passing, token passing, network transmission, or the like.

[0028] Furthermore, embodiments may be implemented by hardware, software, firmware, middleware, microcode, hardware description languages, or any combination thereof. When implemented in software, firmware, middleware or microcode, the program code or code segments to perform the necessary tasks (e.g., a computer-program product) may be stored in a

computer-readable or machine-readable medium. A processor(s) may perform the necessary tasks.

[0029] Representing a video sequence in digital form may require a large number of bits. The data size of a video sequence in digital form may be too large for storage and/or transmission in many applications. Video encoding may be used to compress the size of a video sequence to provide for more efficient storage and/or transmission. Video decoding may be used to decompress a compressed video sequence for display and/or other forms of consumption.

[0030] FIG. 1 illustrates an exemplary video coding/decoding system 100 in which embodiments of the present disclosure may be implemented. Video coding/decoding system 100 comprises a source device 102, a transmission medium 104, and a destination device 106. Source device 102 encodes a video sequence 108 into an encoded bitstream 110 for more efficient storage and/or transmission. Source device 102 may store and/or transmit encoded bitstream 110 to destination device 106 via transmission medium 104. Destination device 106 decodes encoded bitstream 110 to display video sequence 108. Destination device 106 may receive encoded bit stream 110 from source device 102 via transmission medium 104. Source device 102 and destination device 106 may be any one of a number of different devices, including a desktop computer, laptop computer, tablet computer, smart phone, wearable device, television, camera, video gaming console, set-top box, or video streaming device.

[0031] To encode video sequence 108 into encoded bitstream 110, source device 102 may comprise a video source 112, an encoder 114, and an output interface 116. Video source 112 may provide or generate video sequence 108 from a capture of a natural scene and/or a synthetically generated scene. A synthetically generated scene may be a scene comprising computer generated graphics or screen content. Video source 112 may comprise a video capture device (e.g., a video camera), a video archive comprising previously captured natural scenes and/or synthetically generated scenes, a video feed interface to receive captured natural scenes and/or synthetically generated scenes from a video content provider, and/or a processor to generate synthetic scenes.

[0032] A shown in FIG. 1, a video sequence, such as video sequence 108, may comprise a series of pictures (also referred to as frames). A video sequence may achieve the impression of motion when a constant or variable time is used to successively present pictures of the video sequence. A picture may comprise one or more sample arrays of intensity values. The intensity values may be taken at a series of regularly spaced locations within a picture. A color picture typically comprises a luminance sample array and two chrominance sample arrays. The luminance sample array may comprise intensity values representing the brightness (or luma

component, Y) of a picture. The chrominance sample arrays may comprise intensity values that respectively represent the blue and red components of a picture (or chroma components, Cb and Cr) separate from the brightness. Other color picture sample arrays are possible based on different color schemes (e.g., an RGB color scheme). For color pictures, a pixel may refer to all three intensity values for a given location in the three sample arrays used to represent color pictures. A monochrome picture comprises a single, luminance sample array. For monochrome pictures, a pixel may refer to the intensity value at a given location in the single, luminance sample array used to represent monochrome pictures.

[0033] Encoder 114 may encode video sequence 108 into encoded bitstream 110. To encode video sequence 108, encoder 114 may apply one or more prediction techniques to reduce redundant information in video sequence 108. Redundant information is information that may be predicted at a decoder and therefore may not be needed to be transmitted to the decoder for accurate decoding of the video sequence. For example, encoder 114 may apply spatial prediction (e.g., intra-frame or intra prediction), temporal prediction (e.g., inter-frame prediction or inter prediction), inter-layer prediction, and/or other prediction techniques to reduce redundant information in video sequence 108. Before applying the one or more prediction techniques, encoder 114 may partition pictures of video sequence 108 into rectangular regions referred to as blocks. Encoder 114 may then encode a block using one or more of the prediction techniques.

[0034] For temporal prediction, encoder 114 may search for a block similar to the block being encoded in another picture (also referred to as a reference picture) of video sequence 108. The block determined during the search (also referred to as a prediction block) may then be used to predict the block being encoded. For spatial prediction, encoder 114 may form a prediction block based on data from reconstructed neighboring samples of the block to be encoded within the same picture of video sequence 108. A reconstructed sample refers to a sample that was encoded and then decoded. Encoder 114 may determine a prediction error (also referred to as a residual) based on the difference between a block being encoded and a prediction block. The prediction error may represent non-redundant information that may be transmitted to a decoder for accurate decoding of a video sequence.

[0035] Encoder 114 may apply a transform to the prediction error (e.g. a discrete cosine transform (DCT)) to generate transform coefficients. Encoder 114 may form encoded bitstream 110 based on the transform coefficients and other information used to determine prediction blocks (e.g., prediction types, motion vectors, and prediction modes). In some examples, encoder 114 may perform one or more of quantization and entropy coding of the transform

coefficients and/or the other information used to determine prediction blocks before forming encoded bitstream 110 to further reduce the number of bits needed to store and/or transmit video sequence 108.

[0036] Output interface 116 may be configured to write and/or store encoded bitstream 110 onto transmission medium 104 for transmission to destination device 106. In addition or alternatively, output interface 116 may be configured to transmit, upload, and/or stream encoded bitstream 110 to destination device 106 via transmission medium 104. Output interface 116 may comprise a wired and/or wireless transmitter configured to transmit, upload, and/or stream encoded bitstream 110 according to one or more proprietary and/or standardized communication protocols, such as Digital Video Broadcasting (DVB) standards, Advanced Television Systems Committee (ATSC) standards, Integrated Services Digital Broadcasting (ISDB) standards, Data Over Cable Service Interface Specification (DOCSIS) standards, 3rd Generation Partnership Project (3GPP) standards, Institute of Electrical and Electronics Engineers (IEEE) standards, Internet Protocol (IP) standards, and Wireless Application Protocol (WAP) standards.

[0037] Transmission medium 104 may comprise a wireless, wired, and/or computer readable medium. For example, transmission medium 104 may comprise one or more wires, cables, air interfaces, optical discs, flash memory, and/or magnetic memory. In addition or alternatively, transmission medium 104 may comprise one more networks (e.g., the Internet) or file servers configured to store and/or transmit encoded video data.

[0038] To decode encoded bitstream 110 into video sequence 108 for display, destination device 106 may comprise an input interface 118, a decoder 120, and a video display 122. Input interface 118 may be configured to read encoded bitstream 110 stored on transmission medium 104 by source device 102. In addition or alternatively, input interface 118 may be configured to receive, download, and/or stream encoded bitstream 110 from source device 102 via transmission medium 104. Input interface 118 may comprise a wired and/or wireless receiver configured to receive, download, and/or stream encoded bitstream 110 according to one or more proprietary and/or standardized communication protocols, such as those mentioned above.

[0039] Decoder 120 may decode video sequence 108 from encoded bit stream 110. To decode video sequence 108, decoder 120 may generate prediction blocks for pictures of video sequence 108 in a similar manner as encoder 114 and determine prediction errors for the blocks. Decoder 120 may generate the prediction blocks using prediction types, prediction modes, and/or motion vectors received in encoded bit stream 110 and determine the prediction errors using transform coefficients also received in encoded bit stream 110. Decoder 120 may determine the prediction

errors by weighting transform basis functions using the transform coefficients. Decoder 120 may combine the prediction blocks and prediction errors to decode video sequence 108. In some examples, decoder 120 may decode a video sequence that approximates video sequence 108 due to, for example, lossy compression of video sequence 108 by encoder 114 and/or errors introduced into encoded bit stream 110 during transmission to destination device 106.

[0040] Video display 122 may display video sequence 108 to a user. Video display 122 may comprise a cathode rate tube (CRT) display, liquid crystal display (LCD), a plasma display, light emitting diode (LED) display, or any other display device suitable for displaying video sequence 108.

[0041] It should be noted that video encoding/decoding system 100 is presented by way of example and not limitation. In the example of FIG. 1, video encoding/decoding system 100 may have other components and/or arrangements. For example, video source 112 may be external to source device 102. Similarly, video display device 122 may be external to destination device 106 or omitted altogether where video sequence is intended for consumption by a machine and/or storage device. In another example, source device 102 may further comprise a video decoder and destination device 104 may comprise a video encoder. In such an example, source device 102 may be configured to further receive an encoded bit stream from destination device 106 to support two-way video transmission between the devices.

[0042] In the example of FIG. 1, encoder 114 and decoder 120 may operate according to any one of a number of proprietary or industry video coding standards. For example, encoder 114 and decoder 120 may operate according to one or more of International Telecommunications Union Telecommunication Standardization Sector (ITU-T) H.263, ITU-T H.264 and Moving Picture Expert Group (MPEG)-4 Visual (also known as Advanced Video Coding (AVC)), ITU-T H.265 and MPEG-H Part 2 (also known as High Efficiency Video Coding (HEVC), ITU-T H.265 and MPEG-I Part 3 (also known as Versatile Video Coding (VVC)), the WebM VP8 and VP9 codecs, and AOMedia Video 1 (AV1).

[0043] FIG. 2 illustrates an exemplary encoder 200 in which embodiments of the present disclosure may be implemented. Encoder 200 encodes a video sequence 202 into an encoded bitstream 204 for more efficient storage and/or transmission. Encoder 200 may be implemented in video coding/decoding system 100 in FIG. 1 or in any one of a number of different devices, including a desktop computer, laptop computer, tablet computer, smart phone, wearable device, television, camera, video gaming console, set-top box, or video streaming device. Encoder 200 comprises an inter prediction unit 206, an intra prediction unit 208, combiners 210 and 212, a

transform and quantization unit (TR + Q) unit 214, an inverse transform and quantization unit (iTR + iQ) 216, entropy coding unit 218, one or more filters 220, and a buffer 222.

[0044] Encoder 200 may partition the pictures of video sequence 202 into blocks and encode video sequence 202 on a block-by-block basis. Encoder 200 may perform a prediction technique on a block being encoded using either inter prediction unit 206 or intra prediction unit 208. Inter prediction unit 206 may perform inter prediction by searching for a block similar to the block being encoded in another, reconstructed picture (also referred to as a reference picture) of video sequence 202. A reconstructed picture refers to a picture that was encoded and then decoded. The block determined during the search (also referred to as a prediction block) may then be used to predict the block being encoded to remove redundant information. Inter prediction unit 206 may exploit temporal redundancy or similarities in scene content from picture to picture in video sequence 202 to determine the prediction block. For example, scene content between pictures of video sequence 202 may be similar except for differences due to motion or affine transformation of the screen content over time.

[0045] Intra prediction unit 208 may perform intra prediction by forming a prediction block based on data from reconstructed neighboring samples of the block to be encoded within the same picture of video sequence 202. A reconstructed sample refers to a sample that was encoded and then decoded. Intra prediction unit 208 may exploit spatial redundancy or similarities in scene content within a picture of video sequence 202 to determine the prediction block. For example, the texture of a region of scene content in a picture may be similar to the texture in the immediate surrounding area of the region of the scene content in the same picture.

[0046] After prediction, combiner 210 may determine a prediction error (also referred to as a residual) based on the difference between the block being encoded and the prediction block. The prediction error may represent non-redundant information that may be transmitted to a decoder for accurate decoding of a video sequence.

[0047] Transform and quantization unit 214 may transform and quantize the prediction error. Transform and quantization unit 214 may transform the prediction error into transform coefficients by applying, for example, a DCT to reduce correlated information in the prediction error. Transform and quantization unit 214 may quantize the coefficients by mapping data of the transform coefficients to a predefined set of representative values. Transform and quantization unit 214 may quantize the coefficients to reduce irrelevant information in encoded bitstream 204. Irrelevant information is information that may be removed from the coefficients without producing visible and/or perceptible distortion in video sequence 202 after decoding.

[0048] Entropy coding unit 218 may apply one or more entropy coding methods to the quantized transform coefficients to further reduce the bit rate. For example, entropy coding unit 218 may apply context adaptive variable length coding (CAVLC), context adaptive binary arithmetic coding (CABAC), and syntax-based context-based binary arithmetic coding (SBAC). The entropy coded coefficients are packed to form encoded bitstream 204.

[0049] Inverse transform and quantization unit 216 may inverse quantize and inverse transform the quantized transform coefficients to determine a reconstructed prediction error. Combiner 212 may combine the reconstructed prediction error with the prediction block to form a reconstructed block. Filter(s) 220 may filter the reconstructed block using, for example, a deblocking filter and/or a sample-adaptive offset (SAO) filter. Buffer 222 may store the reconstructed block for prediction of one or more other blocks in the same and/or different picture of video sequence 202.

[0050] Although not shown in FIG. 2, encoder 200 further comprises an encoder control unit configured to control one or more of the units of encoder 200 shown in FIG. 2. The encoder control unit may control the one or more units of encoder 200 such that encoded bitstream 204 is generated in conformance with the requirements of any one of a number of proprietary or industry video coding standards. For example, The encoder control unit may control the one or more units of encoder 200 such that encoded bitstream 204 is generated in conformance with one or more of ITU-T H.263, AVC, HEVC, VVC, VP8, VP9, and AV1 video coding standards.

[0051] Within the constraints of a proprietary or industry video coding standard, the encoder control unit may attempt to minimize or reduce the bitrate of encoded bitstream 204 and maximize or increase the reconstructed video quality. For example, the encoder control unit may attempt to minimize or reduce the bitrate of encoded bitstream 204 given a level that the reconstructed video quality may not fall below, or attempt to maximize or increase the reconstructed video quality given a level that the bit rate of encoded bitstream 204 may not exceed. The encoder control unit may determine/control one or more of: partitioning of the pictures of video sequence 202 into blocks, whether a block is inter predicted by inter prediction unit 206 or intra predicted by intra prediction unit 208, a motion vector for inter prediction of a block, an intra prediction mode among a plurality of intra prediction modes for intra prediction of a block, filtering performed by filter(s) 220, and one or more transform types and/or quantization parameters applied by transform and quantization unit 214. The encoder control unit may determine/control the above based on how the determination/control effects a rate-distortion measure for a block or picture being encoded. The encoder control unit may

determine/control the above to reduce the rate-distortion measure for a block or picture being encoded.

[0052] After being determined, the prediction type used to encode a block (intra or inter prediction), prediction information of the block (intra prediction mode if intra predicted, motion vector, etc.), and transform and quantization parameters, may be sent to entropy coding unit 218 to be further compressed to reduce the bit rate. The prediction type, prediction information, and transform and quantization parameters may be packed with the prediction error to form encoded bitstream 204.

[0053] It should be noted that encoder 200 is presented by way of example and not limitation. In other examples, encoder 200 may have other components and/or arrangements. For example, one or more of the components shown in FIG. 2 may be optionally included in encoder 200, such as entropy coding unit 218 and filters(s) 220.

[0054] FIG. 3 illustrates an exemplary decoder 300 in which embodiments of the present disclosure may be implemented. Decoder 300 decodes an encoded bitstream 302 into a decoded video sequence for display and/or some other form of consumption. Decoder 300 may be implemented in video coding/decoding system 100 in FIG. 1 or in any one of a number of different devices, including a desktop computer, laptop computer, tablet computer, smart phone, wearable device, television, camera, video gaming console, set-top box, or video streaming device. Decoder 300 comprises an entropy decoding unit 306, an inverse transform and quantization (iTR + iQ) unit 308, a combiner 310, one or more filters 312, a buffer 314, an inter prediction unit 316, and an intra prediction unit 318.

[0055] Although not shown in FIG. 3, decoder 300 further comprises a decoder control unit configured to control one or more of the units of decoder 300 shown in FIG. 3. The decoder control unit may control the one or more units of decoder 300 such that encoded bitstream 302 is decoded in conformance with the requirements of any one of a number of proprietary or industry video coding standards. For example, The decoder control unit may control the one or more units of decoder 300 such that encoded bitstream 302 is decoded in conformance with one or more of ITU-T H.263, AVC, HEVC, VVC, VP8, VP9, and AV1 video coding standards.

[0056] The decoder control unit may determine/control one or more of: whether a block is inter predicted by inter prediction unit 316 or intra predicted by intra prediction unit 318, a motion vector for inter prediction of a block, an intra prediction mode among a plurality of intra prediction modes for intra prediction of a block, filtering performed by filter(s) 312, and one or more inverse transform types and/or inverse quantization parameters to be applied by inverse

transform and quantization unit 308. One or more of the control parameters used by the decoder control unit may be packed in encoded bitstream 302.

[0057] Entropy decoding unit 306 may entropy decode the encoded bitstream 302. Inverse transform and quantization unit 308 may inverse quantize and inverse transform the quantized transform coefficients to determine a decoded prediction error. Combiner 310 may combine the decoded prediction error with a prediction block to form a decoded block. The prediction block may be generated by inter prediction unit 318 or inter prediction unit 316 as described above with respect to encoder 200 in FIG 2. Filter(s) 312 may filter the decoded block using, for example, a deblocking filter and/or a sample-adaptive offset (SAO) filter. Buffer 314 may store the decoded block for prediction of one or more other blocks in the same and/or different picture of the video sequence in encoded bitstream 302. Decoded video sequence 304 may be output from filter(s) 312 as shown in FIG. 3.

[0058] It should be noted that decoder 300 is presented by way of example and not limitation. In other examples, decoder 300 may have other components and/or arrangements. For example, one or more of the components shown in FIG. 3 may be optionally included in decoder 300, such as entropy decoding unit 306 and filters(s) 312.

[0059] It should be further noted that, although not shown in FIGS. 2 and 3, each of encoder 200 and decoder 300 may further comprise an intra block copy unit in addition to inter prediction and intra prediction units. The intra block copy unit may perform similar to an inter prediction unit but predict blocks within the same picture. For example, the intra block copy unit may exploit repeated patterns that appear in screen content. Screen content may include, for example, computer generated text, graphics, and animation.

[0060] As mentioned above, video encoding and decoding may be performed on a block-by-block basis. The process of partitioning a picture into blocks may be adaptive based on the content of the picture. For example, larger block partitions may be used in areas of a picture with higher levels of homogeneity to improve coding efficiency.

[0061] In HEVC, a picture may be partitioned into non-overlapping square blocks, referred to as coding tree blocks (CTBs), comprising samples of a sample array. A CTB may have a size of 2^nx2^n samples, where n may be specified by a parameter of the encoding system. For example, n may be 4, 5, or 6. A CTB may be further partitioned by a recursive quadtree partitioning into coding blocks (CBs) of half vertical and half horizontal size. The CTB forms the root of the quadtree. A CB that is not split further as part of the recursive quadtree partitioning may be referred to as a leaf-CB of the quadtree and otherwise as a non-leaf CB of the quadtree. A CB may have a minimum size specified by a parameter of the encoding system. For example, a CB

may have a minimum size of 4x4, 8x8, 16x16, 32x32, or 64x64 samples. For inter and intra prediction, a CB may be further partitioned into one or more prediction blocks (PBs) for performing inter and intra prediction. A PB may be a rectangular block of samples on which the same prediction type/mode may be applied. For transformations, a CB may be partitioned into one or more transform blocks (TBs). A TB may be a rectangular block of samples that may determine an applied transform size.

[0062] FIG. 4 illustrates an example quadtree partitioning of a CTB 400. FIG. 5 illustrates a corresponding quadtree 500 of the example quadtree partitioning of CTB 400 in FIG. 4. As shown in FIGS. 4 and 5, CTB 400 is first partitioned into four CBs of half vertical and half horizontal size. Three of the resulting CBs of the first level partitioning of CTB 400 are leaf-CBs. The three leaf CBs of the first level partitioning of CTB 400 are respectively labeled 7, 8, and 9 in FIGS. 4 and 5. The non-leaf CB of the first level partitioning of CTB 400 is partitioned into four sub-CBs of half vertical and half horizontal size. Three of the resulting sub-CBs of the second level partitioning of CTB 400 are leaf CBs. The three leaf CBs of the second level partitioning of CTB 400 are respectively labeled 0, 5, and 6 in FIGS. 4 and 5. Finally, the non-leaf CB of the second level partitioning of CTB 400 is partitioned into four leaf CBs of half vertical and half horizontal size. The four leaf CBs are respectively labeled 1, 2, 3, and 4 in FIGS. 4 and 5.

[0063] Altogether, CTB 400 is partitioned into 10 leaf CBs respectively labeled 0-9. The resulting quadtree partitioning of CTB 400 may be scanned using a z-scan (left-to-right, top-to-bottom) to form the sequence order for encoding/decoding the CB leaf nodes. The numeric label of each CB leaf node in FIGS. 4 and 5 may correspond to the sequence order for encoding/decoding, with CB leaf node 0 encoded/decoded first and CB leaf node 9 encoded/decoded last. Although not shown in FIGS. 4 and 5, it should be noted that each CB leaf node may comprise one or more PBs and TBs.

[0064] In VVC, a picture may be partitioned in a similar manner as in HEVC. A picture may be first partitioned into non-overlapping square CTBs. The CTBs may then be partitioned by a recursive quadtree partitioning into CBs of half vertical and half horizontal size. In VVC, a quadtree leaf node may be further partitioned by a binary tree or ternary tree partitioning into CBs of unequal sizes. FIG. 6 illustrates example binary and ternary tree partitions. A binary tree partition may divide a parent block in half in either the vertical direction 602 or horizontal direction 604. The resulting partitions may be half in size as compared to the parent block. A ternary tree partition may divide a parent block into three parts in either the vertical direction

606 or horizontal direction 608. The middle partition may be twice as large as the other two end partitions in a ternary tree partition.

[0065] Because of the addition of binary and ternary tree partitioning, in VVC the block partitioning strategy may be referred to as quadtree + multi-type tree partitioning. FIG. 7 illustrates an example quadtree + multi-type tree partitioning of a CTB 700. FIG. 8 illustrates a corresponding quadtree + multi-type tree 800 of the example quadtree + multi-type tree partitioning of CTB 700 in FIG. 7. In both FIGS. 7 and 8, quadtree splits are shown in solid lines and multi-type tree splits are shown in dashed lines. For ease of explanation, CTB 700 is shown with the same quadtree partitioning as CTB 400 described in FIG. 4. Therefore, description of the quadtree partitioning of CTB 700 is omitted. The description of the additional multi-type tree partitions of CTB 700 is made relative to three leaf-CBs shown in FIG. 4 that have been further partitioned using one or more binary and ternary tree partitions. The three leaf-CBs in FIG. 4 that are shown in FIG. 7 as being further partitioned are leaf-CBs 5, 8, and 9.

[0066] Starting with leaf-CB 5 in FIG. 4, FIG. 7 shows this leaf-CB partitioned into two CBs based on a vertical binary tree partitioning. The two resulting CBs are leaf-CBs respectively labeled 5 and 6 in FIGS. 7 and 8. With respect to leaf-CB 8 in FIG. 4, FIG. 7 shows this leaf-CB partitioned into three CBs based on a vertical ternary tree partition. Two of the three resulting CBs are leaf-CBs respectively labeled 9 and 14 in FIGS. 7 and 8. The remaining, non-leaf CB is partitioned first into two CBs based on a horizontal binary tree partition, one of which is a leaf-CB labeled 10 and the other of which is further partitioned into three CBs based on a vertical ternary tree partition. The resulting three CBs are leaf-CBs respectively labeled 11, 12, and 13 in FIGS. 7 and 8. Finally, with respect to leaf-CB 9 in FIG. 4, FIG. 7 shows this leaf-CB partitioned into three CBs based on a horizontal ternary tree partition. Two of the three CBs are leaf-CBs respectively labeled 15 and 19 in FIGS. 7 and 8. The remaining, non-leaf CB is partitioned into three CBs based on another horizontal ternary tree partition. The resulting three CBs are all leaf-CBs respectively labeled 16, 17, and 18 in FIGS. 7 and 8.

[0067] Altogether, CTB 700 is partitioned into 20 leaf CBs respectively labeled 0-19. The resulting quadtree + multi-type tree partitioning of CTB 700 may be scanned using a z-scan (left-to-right, top-to-bottom) to form the sequence order for encoding/decoding the CB leaf nodes. The numeric label of each CB leaf node in FIGS. 7 and 8 may correspond to the sequence order for encoding/decoding, with CB leaf node 0 encoded/decoded first and CB leaf node 19 encoded/decoded last. Although not shown in FIGS. 7 and 8, it should be noted that each CB leaf node may comprise one or more PBs and TBs.

[0068] In addition to specifying various blocks (e.g., CTB, CB, PB, TB), HEVC and VVC further define various units. While blocks may comprise a rectangular area of samples in a sample array, units may comprise the collocated blocks of samples from the different sample arrays (e.g., luma and chroma sample arrays) that form a picture as well as syntax elements and prediction data of the blocks. A coding tree unit (CTU) may comprise the collocated CTBs of the different sample arrays and may form a complete entity in an encoded bit stream. A coding unit (CU) may comprise the collocated CBs of the different sample arrays and syntax structures used to code the samples of the CBs. A prediction unit (PU) may comprise the collocated PBs of the different sample arrays and syntax elements used to predict the PBs. A transform unit (TU) may comprise TBs of the different samples arrays and syntax elements used to transform the TBs.

[0069] It should be noted that the term block may be used to refer to any of a CTB, CB, PB, TB, CTU, CU, PU, or TU in the context of HEVC and VVC. It should be further noted that the term block may be used to refer to similar data structures in the context of other video coding standards. For example, the term block may refer to a macroblock in AVC, a macroblock or sub-block in VP8, a superblock or sub-block in VP9, or a superblock or sub-block in AV1.

[0070] In intra prediction, samples of a block to be encoded (also referred to as the current block) may be predicted from samples of the column immediately adjacent to the left-most column of the current block and samples of the row immediately adjacent to the top-most row of the current block. The samples from the immediately adjacent column and row may be jointly referred to as reference samples. Each sample of the current block may be predicted by projecting the position of the sample in the current block in a given direction (also referred to as an intra prediction mode) to a point along the reference samples. The sample may be predicted by interpolating between the two closest reference samples of the projection point if the projection does not fall directly on a reference sample. A prediction error (also referred to as a residual) may be determined for the current block based on differences between the predicted sample values and the original sample values of the current block.

[0071] At an encoder, this process of predicting samples and determining a prediction error based on a difference between the predicted samples and original samples may be performed for a plurality of different intra prediction modes, including non-directional intra prediction modes. The encoder may select one of the plurality of intra prediction modes and its corresponding prediction error to encode the current block. The encoder may send an indication of the selected prediction mode and its corresponding prediction error to a decoder for decoding of the current block. The decoder may decode the current block by predicting the samples of the current block

using the intra prediction mode indicated by the encoder and combining the predicted samples with the prediction error.

[0072] FIG. 9 illustrates an example set of reference samples 902 determined for intra prediction of a current block 904 being encoded or decoded. In FIG. 9, current block 904 corresponds to block 3 of partitioned CTB 700 in FIG. 7. As explained above, the numeric labels 0-19 of the blocks of partitioned CTB 700 may correspond to the sequence order for encoding/decoding the blocks and are used as such in the example of FIG. 9.

[0073] Given current block 904 is of w x h samples in size, reference samples 902 may extend over $2w$ samples of the row immediately adjacent to the top-most row of current block 904, $2h$ samples of the column immediately adjacent to the left-most column of current block 904, and the top left neighboring corner sample to current block 904. In the example of FIG. 9, current block 904 is square, so $w = h = s$. For constructing the set of reference samples 902, available samples from neighboring blocks of current block 904 may be used. Samples may not be available for constructing the set of reference samples 902 if, for example, the samples would lie outside the picture of the current block, the samples are part of a different slice of the current block (where the concept of slices are used), and/or the samples belong to blocks that have been inter coded and constrained intra prediction is indicated. When constrained intra prediction is indicated, intra prediction may not be dependent on inter predicted blocks.

[0074] In addition to the above, samples that may not be available for constructing the set of reference samples 902 include samples in blocks that have not already been encoded and reconstructed at an encoder or decoded at a decoder based on the sequence order for encoding/decoding. This restriction may allow identical prediction results to be determined at both the encoder and decoder. In FIG. 9, samples from neighboring blocks 0, 1, and 2 may be available to construct reference samples 902 given that these blocks are encoded and reconstructed at an encoder and decoded at a decoder prior to coding of current block 904. This assumes there are no other issues, such as those mentioned above, preventing the availability of samples from neighboring blocks 0, 1, and 2. However, the portion of reference samples 902 from neighboring block 6 may not be available due to the sequence order for encoding/decoding.

[0075] Unavailable ones of reference samples 902 may be filled with available ones of reference samples 902. For example, an unavailable reference sample may be filled with a nearest available reference sample determined by moving in a clock-wise direction through reference samples 902 from the position of the unavailable reference. If no reference samples

are available, reference samples 902 may be filled with the mid-value of the dynamic range of the picture being coded.

[0076] It should be noted that reference samples 902 may be filtered based on the size of current block 904 being coded and an applied intra prediction mode. It should be further noted that FIG. 9 illustrates only one exemplary determination of reference samples for intra prediction of a block. In some proprietary and industry video coding standards, reference samples may be determined in a different manner than discussed above. For example, multiple reference lines may be used in other instances, such as used in VVC.

[0077] After reference samples 902 are determined and optionally filtered, samples of current block 904 may be intra predicted based on reference samples 902. Most encoders/decoders support a plurality of intra prediction modes in accordance with one or more video coding standards. For example, HEVC supports 35 intra prediction modes, including a planar mode, a DC mode, and 33 angular modes. VVC supports 67 intra prediction modes, including a planar mode, a DC mode, and 65 angular modes. Planar and DC modes may be used to predict smooth and gradually changing regions of a picture. Angular modes may be used to predict directional structures in regions of a picture.

[0078] FIG. 10A illustrates the 35 intra prediction modes supported by HEVC. The 35 intra prediction modes are identified by indices 0 to 34. Prediction mode 0 corresponds to planar mode. Prediction mode 1 corresponds to DC mode. Prediction modes 2-34 correspond to angular modes. Prediction modes 2-18 may be referred to as horizontal prediction modes because the principal source of prediction is in the horizontal direction. Prediction modes 19-34 may be referred to as vertical prediction modes because the principal source of prediction is in the vertical direction.

[0079] FIG. 10B illustrates the 67 intra prediction modes supported by HEVC. The 67 intra prediction modes are identified by indices 0 to 66. Prediction mode 0 corresponds to planar mode. Prediction mode 1 corresponds to DC mode. Prediction modes 2-66 correspond to angular modes. Prediction modes 2-34 may be referred to as horizontal prediction modes because the principal source of prediction is in the horizontal direction. Prediction modes 35-66 may be referred to as vertical prediction modes because the principal source of prediction is in the vertical direction. Because blocks in VVC may be non-square, some of the intra prediction modes illustrated in FIG. 10B may be adaptively replaced by wide-angle directions.

[0080] To further describe the application of intra prediction modes to determine a prediction of a current block, reference is made to FIGS. 11 and 12. In FIG. 11, current block 904 and reference samples 902 from FIG. 9 are shown in a two-dimensional x, y plane. Current block

904 is referred to as *Cb*, where *Cb*(*x*, *y*) denotes the predicted value of current block 904 at the coordinates (*x*, *y*). Reference samples 902 are referred to as *r*, where *r*(*x*, *y*) denotes the reference sample of reference samples 902 at the coordinates (*x*, *y*).

[0081] For planar mode, a sample in *Cb* may be predicted by calculating the mean of two interpolated values. The first of the two interpolated values may be based on a horizontal linear interpolation of the predicted sample in *Cb*. The second of the two interpolated values may be based on a vertical linear interpolation of the predicted sample in *Cb*. The predicted value of the sample in *Cb* may be calculated as

$$Cb(x,y) = \frac{1}{2 \cdot s}(h(x,y) + v(x,y) + s) \tag{1}$$

where

$$h(x,y) = (s - x - 1) \cdot r(-1,y) + (x + 1) \cdot r(s,-1) \tag{2}$$

may be the horizonal linear interpolation of the predicted sample in *Cb* and

$$v(x,y) = (s - y - 1) \cdot r(x,-1) + (y + 1) \cdot r(-1,s) \tag{3}$$

may be the vertical linear interpolation of the predicted sample in *Cb*.

[0082] For DC mode, a sample in *Cb* may be predicted by the mean of the reference samples. The predicted value of the sample in *Cb* may be calculated as

$$Cb(x,y) = \frac{1}{2 \cdot s} \cdot \left(\sum_{x=0}^{s-1} r(x,-1) + \sum_{y=0}^{s-1} r(-1,y) \right) \tag{4}$$

A boundary filter may be applied to boundary samples in *Cb* to smooth the transition between the boundary samples and their respective adjacent neighboring reference sample(s) in *r*.

[0083] For angular modes, a sample in *Cb* may be predicted by projecting the position of the sample in a direction specified by a given angular mode to a point on the horizontal or vertical axis comprising the reference samples *r*. The sample may be predicted by interpolating between the two closest reference samples in *r* of the projection point if the projection does not fall directly on a reference sample in *r*. The direction specified by the angular mode may be given by an angle φ defined relative to the *y*-axis for vertical prediction modes (e.g., modes 19-34 in HEVC and modes 35-66 in VVC) and relative to the x-axis for horizontal prediction modes (e.g., modes 2-18 in HEVC and modes 2-34 in VVC).

[0084] FIG. 12 illustrates a sample in *Cb* predicted for a vertical prediction mode. For vertical prediction modes, the position (*x*, *y*) of the sample in *Cb* is projected onto the horizontal axis comprising reference samples *r*. Because the projection falls between two reference samples *r*1

and $r2$ in the example of FIG. 12, the predicted value of the sample in Cb may be calculated as the linear interpolation between the two reference samples $r1$ and $r2$ as

$$Cb(x, y) = (1 - \Delta) \cdot r1 + \Delta \cdot r2 \quad (5)$$

where

$$r1 = r(x + \lfloor (y + 1) \cdot \tan \varphi \rfloor, -1), \quad (6)$$

$$r2 = r(x + \lfloor (y + 1) \cdot \tan \varphi \rfloor + 1, -1), \quad (7)$$

$$\Delta = ((y + 1) \cdot \tan \varphi) - \lfloor (y + 1) \cdot \tan \varphi \rfloor, \text{ and} \quad (8)$$

$$\lfloor \cdot \rfloor \text{ is an integer floor.} \quad (9)$$

[0085] It should be noted that the weighting factors $(1 - \Delta)$ and Δ may be calculated with some predefined level of precision, such as 1/32 pixel precision. To avoid floating point operations while preserving the specified precision, the weighting factors $(1 - \Delta)$ and Δ may be multiplied by the reciprocal of the specified precision used and then divided by the reciprocal using, for example, right shift operations. It should be further noted that supplementary reference samples may be constructed for the case where the position (x, y) of a sample Cb to predicted is projected to a negative x coordinate, which happens with negative angles φ. The supplementary reference samples may be constructed by projecting the reference samples in r on the vertical axis to the horizontal axis using the angle φ. Finally, it should be further noted that a sample in Cb may be predicted for a horizontal prediction mode in a similar manner as discussed above for vertical prediction modes. For horizontal prediction modes, the position (x, y) of the sample in Cb may be projected onto the vertical axis comprising reference samples r and the angle φ may be defined relative to the x-axis. Supplemental reference samples may be similarly constructed for horizontal prediction modes by projecting the reference samples in r on the horizontal axis to the vertical axis using the angle φ.

[0086] An encoder may predict the samples of a current block being encoded, such as current block 904, for a plurality of intra prediction modes as explained above. For example, the encoder may predict the samples of the current block for each of the 35 intra prediction modes in HEVC or 67 intra prediction modes in VVC. For each intra prediction mode applied, the encoder may determine a prediction error for the current block based on a difference (e.g., sum of squared differences (SSD), sum of absolute differences (SAD), or sum of absolute transformed differences (SATD)) between the prediction samples determined for the intra prediction mode and the original samples of the current block. The encoder may select one of the intra prediction modes to encode the current block based on the determined prediction errors. For example, the encoder may select an intra prediction mode that results in the smallest

prediction error for the current block. In another example, the encoder may select the intra prediction mode to encode the current block based on a rate-distortion measure determined using the prediction errors. The encoder may send an indication of the selected intra prediction mode and its corresponding prediction error to a decoder for decoding of the current block.

[0087] Although the description above was primarily made with respect to intra prediction modes in HEVC and VVC, it will be understood that the techniques of the present disclosure described above and further below may be applied to other intra prediction modes, including those of other video coding standards like VP8, VP9, AV1, and the like.

[0088] In existing technologies, an encoder may signal an intra prediction mode selected to encode a block using a fixed-length binarization (e.g., an unsigned integer $0 \leq x < M$ given by its binary representation of length $\lceil \log_2 M \rceil$ bits) of the mode index. For example, an encoder may signal the intra prediction mode selected among the 35 intra prediction modes in HEVC to encode a block using a fixed-length binarization of the mode index. With 35 different intra prediction modes available, the fixed-length binarization of the mode index requires 6-bits. This signaling overhead reduces the compression gain achieved by removing redundant information for a block based on the selected intra prediction mode. In newer video coding standards, the reduction in compression gain may be further reduced by support for a larger number of intra prediction modes. For example, in VVC, the successor to HEVC, 67 intra prediction modes are supported. With 67 different intra prediction modes available, the fixed-length binarization of the mode index requires 7-bits. In some video coding standards, like HEVC and VVC, a smaller list of most probable modes (MPMs) may be constructed at an encoder and decoder to reduce intra prediction mode signaling overhead. If a selected intra prediction mode is within the MPM list, an encoder may signal the selected intra prediction mode using a smaller number of bits (e.g., 2 or 3 bits) based on the smaller size of the MPM list. However, even with the use of the MPM list, an encoder may often have to fall back to signaling a selected intra-prediction mode using a fixed-length binarization of the mode index based on the selected mode not being in the MPM list.

[0089] Embodiments of the present disclosure are related to an approach for decreasing the overhead of signaling an intra prediction mode selected by an encoder to encode a current block. Embodiments of the present disclosure may determine a prediction of the selected intra prediction mode at an encoder and decoder and use the prediction to more efficiently signal the selected intra prediction mode. The prediction of the selected intra prediction mode may be determined based on an intra prediction of a neighboring region of the current block. The intra prediction of the neighboring region of the current block may be generated from reference

samples determined for intra prediction of the current block. These and other features of the present disclosure are described further below.

[0090] FIG. 13 illustrates example neighboring regions for predicting a selected intra prediction mode for a current block being encoded by an encoder in accordance with embodiments of the present disclosure. FIG. 13 illustrates a set of reference samples 1302, a current block 1304 being encoded using intra prediction from reference samples 1302, and neighboring regions 1306 and 1308 for predicting a selected intra prediction mode for current block 1304.

[0091] The encoder may determine reference samples 1302 for intra prediction of current block 1304. Given current block 1304 is of w x h samples in size, reference samples 1304 may extend over $2w$ samples of the row immediately adjacent to the top-most row of current block 1304, $2h$ samples of the column immediately adjacent to the left-most column of current block 1304, and the top left neighboring corner sample to current block 1304. In FIG. 13, current block 1304 is square, so $w = h = s$.

[0092] For constructing the set of reference samples 1302, the encoder may use available samples from neighboring blocks of current block 1304 as discussed above with respect to FIG. 9. The encoder may fill unavailable ones of reference samples 1302 with available ones of reference samples 1302. For example, the encoder may fill an unavailable reference sample with a nearest available reference sample determined by moving in a clock-wise direction through reference samples 1302 from the position of the unavailable reference. If no reference samples are available, the encoder may fill reference samples 1302 with the mid-value of the dynamic range of the picture being coded. It should be noted that FIG. 13 illustrates only one exemplary determination of reference samples for intra prediction of current block 1304. In other examples, the encoder may determine reference samples in a different manner (e.g., using multiple reference samples as in VVC, or the reference samples may not extend over $2w$ samples of the row immediately adjacent to the top-most row of current block 1304 or $2h$ samples of the column immediately adjacent to the left-most column of current block 1304).

[0093] After the encoder determines and optionally filters reference samples 1302, samples of current block 1304 may be intra predicted based on reference samples 1302. For example, the encoder may predict the samples of current block 1304 for each of the 35 intra prediction modes in HEVC or 67 intra prediction modes in VVC. For each intra prediction mode 1310 applied, the encoder may determine a prediction error for current block 1304 based on a difference (e.g., sum of squared differences (SSD), sum of absolute differences (SAD), or sum of absolute transformed differences (SATD)) between the prediction samples determined for the intra prediction mode and the original samples of current block 1304. The encoder may select an

intra prediction mode 1312 from the intra prediction modes 1310 to encode current block 1304 based on the determined prediction errors. For example, the encoder may select an intra prediction mode that results in the smallest prediction error for current block 1304. In another example, the encoder may select the intra prediction mode to encode current block 1304 based on a rate-distortion measure determined using the prediction errors. The encoder may send an indication of selected intra prediction mode 1312 and its corresponding prediction error to a decoder for decoding of current block 1304.

[0094] To reduce the overhead of signaling selected intra prediction mode 1312 to a decoder, the encoder may determine a prediction of selected intra prediction mode 1312 and use the prediction to more efficiently signal selected intra prediction mode 1312. The encoder may determine the prediction of selected intra prediction mode 1312 based on an intra prediction of at least one of neighboring regions 1306 and 1308 of current block 1304.

[0095] Neighboring regions 1306 and 1308 may comprise a region of available samples. Samples may not be available for constructing neighboring regions 1306 and 1308 if, for example, the samples would lie outside the picture of current block 1304, the samples are part of a different slice of the current block (where the concept of slices are used), and/or the samples belong to blocks that have been inter coded and constrained intra prediction is indicated. When constrained intra prediction is indicated, intra prediction may not be dependent on inter predicted blocks. In addition to the above, samples that may not be available for constructing neighboring regions 1306 and 1308 may include samples in blocks that have not already been encoded and reconstructed at an encoder or decoded at a decoder based on the sequence order for encoding/decoding blocks. This restriction may allow identical prediction results for the selected intra prediction mode to be determined at both the encoder and decoder based on neighboring regions 1306 and 1308.

[0096] In an embodiment, neighboring regions 1306 and 1308 may be rectangular and may have a size determined based on current block 1304. For example, neighboring region 1306 may have a size of n x s samples as shown in FIG. 13, where the height of neighboring region 1306 is equal to the height of current block 1304. In other embodiments, the height of neighboring region 1306 may be of different size than the height of current block 1304. The width of neighboring region 1306 may be determined to be any reasonable value. For example, the width of neighboring region 1306 may be determined based on a number of available samples for constructing neighboring region 1306 and/or the width of current block 1304. Similarly, neighboring region 1308 may have a size of s x m samples as shown in FIG. 13, where the width of neighboring region 1308 is equal to the width of current block 1304. In other

embodiments, the width of neighboring region 1308 may be of different size than the width of current block 1304. The height of neighboring region 1308 may be determined to be any reasonable value. For example, the height of neighboring region 1308 may be determined based on a number of available samples for constructing neighboring region 1308 and/or the height of current block 1304.

[0097] In an embodiment, neighboring regions 1306 and 1308 may be directly adjacent to reference samples 1302 as shown in FIG. 13. In other embodiments, one or both of neighboring regions 1306 and 1308 may not be directly adjacent to reference samples 1302. For example, neighboring region 1306 may be one or more columns of samples removed from the samples of reference samples 1302 along the *y* axis. Similarly, neighboring region 1308 may be one or more rows of samples removed from the samples of reference samples 1302 along the *x* axis.

[0098] The encoder may generate, for each of a plurality of intra prediction modes 1314, a prediction of at least one of neighboring regions 1306 and 1308 from reference samples 1302. The encoder may generate the predictions in a different or similar manner to the predictions generated for current block 1304 from reference samples 1302. To further describe an example application of intra prediction modes 1314 to determine a prediction of at least one of neighboring regions 1306 and 1308, neighboring regions 1306 and 1308 may be referred to as *Nr*, where *Nr*(*x*, *y*) denotes the predicted value of neighboring regions 1306 and 1308 at the coordinates (*x*, *y*). Reference samples 1302 may be further referred to as *r*, where *r*(*x*, *y*) denotes the reference sample of reference samples 1302 at the coordinates (*x*, *y*).

[0099] For planar mode, a sample in *Nr* may be predicted based on a horizontal linear interpolation (or extrapolation) $h(x, y)$ or a vertical linear interpolation (or extrapolation) $v(x, y)$. For example, the predicted value of the sample in *Nr* may be calculated as

$$h(x,y) = (s - x - 1) \cdot r(-1,-1) + (x + 1) \cdot r(s,-1), \qquad if\ x < 0 \tag{10}$$

and

$$v(x,y) = (s - y - 1) \cdot r(-1,-1) + (y + 1) \cdot r(-1,s), \qquad if\ x \geq 0. \tag{11}$$

[0100] For DC mode, a sample in *Nr* may be predicted by the mean of the reference samples. The predicted value of the sample in *Nr* may be calculated as

$$Nr(x,y) = \frac{1}{2 \cdot s} \cdot \left(\sum_{x=0}^{s-1} r(x,-1) + \sum_{y=0}^{s-1} r(-1,y) \right) \tag{12}$$

[0101] For angular modes, a sample in *Nr* may be predicted by projecting the position of the sample in a direction specified by a given angular mode to a point on the horizontal or vertical

axis comprising the reference samples *r*. The sample may be predicted by interpolating (or extrapolating) between the two closest reference samples in *r* of the projection point if the projection does not fall directly on a reference sample in *r*. The direction specified by the angular mode may be given by an angle φ defined relative to the *y*-axis for vertical prediction modes (e.g., modes 19-34 in HEVC and modes 35-66 in VVC) and relative to the x-axis for horizontal prediction modes (e.g., modes 2-18 in HEVC and modes 2-34 in VVC).

[0102] FIG. 14 illustrates a sample in *Nr* predicted for a vertical prediction mode. For vertical prediction modes, the position (*x*, *y*) of the sample in *Nr* is projected onto the horizontal axis comprising reference samples *r*. Because the projection falls between two reference samples *r*1 and *r*2, the predicted value of the sample in *Nr* may be calculated as the linear interpolation between the two reference samples *r*1 and *r*2 as

$$Nr(x,y) = (1-\Delta)\cdot r1 + \Delta\cdot r2 \tag{13}$$

where

$$r1 = r(x + \lfloor -(y+1)\cdot \tan\varphi \rfloor, -1), \tag{14}$$

$$r2 = r(x + \lfloor -(y+1)\cdot \tan\varphi \rfloor + 1, -1), \tag{15}$$

$$\Delta = (-(y+1)\cdot \tan\varphi) - \lfloor -(y+1)\cdot \tan\varphi \rfloor, \text{ and} \tag{16}$$

$$\lfloor\cdot\rfloor \text{ is an integer floor.} \tag{17}$$

[0103] It should be noted that the weighting factors $(1-\Delta)$ and Δ may be calculated with some predefined level of precision, such as 1/32 pixel precision. To avoid floating point operations while preserving the specified precision, the weighting factors $(1-\Delta)$ and Δ may be multiplied by the reciprocal of the specified pixel precision and then divided by the reciprocal using, for example, right shift operations. It should be further noted that supplementary reference samples may be constructed for the case where the position (*x*, *y*) of a sample *Nr* to predicted is projected to a negative *x* coordinate, which happens with negative angles φ. The supplementary reference samples may be constructed by projecting the reference samples in *r* on the vertical axis to the horizontal axis using the angle φ. Finally, it should be further noted that a sample in *Nr* may be predicted for a horizontal prediction mode in a similar manner as discussed above for vertical prediction modes. For horizontal prediction modes, the position (*x*, *y*) of the sample in *Nr* may be projected onto the vertical axis comprising reference samples *r* and the angle φ may be defined relative to the *x*-axis. Supplementary reference samples may be similarly constructed for the horizontal prediction modes by projecting the reference samples in *r* on the horizontal axis to the vertical axis using the angle φ.

[0104] It should be noted that the above calculations for DC, planar, and angular modes are provided by way of example and not limitation. Other calculations for these intra prediction modes may be used.

[0105] Referring back to FIG. 13, for each applied intra prediction mode 1314, the encoder may determine a prediction error for the at least one of neighboring regions 1306 and 1308 based on a difference (e.g., sum of squared differences (SSD), sum of absolute differences (SAD), or sum of absolute transformed differences (SATD)) between the prediction samples determined for the intra prediction mode and the reconstructed samples of the at least one neighboring regions 1306 and 1308. The encoder may select an intra prediction mode 1316 from the applied intra prediction modes 1314 as a prediction for selected intra prediction mode 1312 based on the determined prediction errors. For example, the encoder may select an intra prediction mode 1316 from the applied intra prediction modes 1314 that results in the smallest prediction error for the at least one of neighboring regions 1306 and 1308. It should be noted that other selection criteria may be used to select intra prediction mode 1316 from the applied intra prediction modes 1314.

[0106] After selecting intra prediction mode 1316, the encoder may compare selected intra prediction mode 1316 determined for the at least one of neighboring regions 1306 and 1308 to selected intra prediction mode 1312 determined for current block 1304. Based on the comparison, the encoder may signal selected intra prediction mode 1312 for current block 1304 to the decoder in a more efficient manner than prior signaling techniques.

[0107] For example, based on selected intra prediction mode 1316 determined for the at least one of neighboring regions 1306 and 1308 being the same as selected intra prediction mode 1312 determined for current block 1304 as shown in FIG. 15A, the encoder may signal selected intra prediction mode 1312 using one or more fewer bits than used to signal a fixed-length binarization of the selected intra prediction mode 1312. In an embodiment, the encoder may signal selected intra prediction mode 1312 to a decoder using a single, one-bit flag based on selected intra prediction mode 1316 determined for the at least one of neighboring regions 1306 and 1308 being the same as selected intra prediction mode 1312 determined for current block 1304. The signaling for selected intra prediction mode 1312 may indicate to the decoder that selected intra prediction mode 1316 determined for the at least one of neighboring regions 1306 and 1308 is the same as selected intra prediction mode 1312 determined for current block 1304.

[0108] The decoder may decode intra prediction mode 1312 based on the signaling from the encoder and the decoder's own independent determination of selected intra prediction mode 1316. For example, the decoder may construct reference samples 1302 and perform intra

prediction of the at least one of neighboring regions 1306 and 1308 in the same manner as the encoder described above to independently determine intra prediction mode 1316 from among the plurality of intra prediction modes 1314. The encoder may then decode intra prediction mode 1312 based on the signaling from the encoder and the decoder's independent determination of selected intra prediction mode 1316. For example, based on the signaling for selected intra prediction mode 1312 indicating that selected intra prediction mode 1316 is the same as selected intra prediction mode 1312, the decoder may decode intra prediction mode 1312 as the decoder's independent determination of intra prediction mode 1316. The decoder may generate a prediction for current block 1304 based on the decoded intra prediction mode 1312. The decoder may combine the prediction with a prediction error of current block 1304 received from the encoder to decode current block 1304.

[0109] In another example, based on selected intra prediction mode 1316 determined for the at least one of neighboring regions 1306 and 1308 being different than selected intra prediction mode 1312 determined for current block 1304 as shown in FIG. 15B, the encoder may signal selected intra prediction mode 1312 based on a difference between an index of selected intra prediction mode 1316 and an index of selected intra prediction mode 1312. For example, the signaling for selected intra prediction mode 1312 may indicate to a decoder the difference between an index of selected intra prediction mode 1316 and an index of selected intra prediction mode 1312.

[0110] Based on the signaling for selected intra prediction mode 1312 indicating the difference between an index of selected intra prediction mode 1316 and an index of selected intra prediction mode 1312, the decoder may add or subtract the difference from the index of the decoder's independent determination of intra prediction mode 1316. The decoder may use the resulting value of the addition or subtraction as the index of the decoded intra prediction mode 1312.

[0111] In another example, based on selected intra prediction mode 1316 being different than selected intra prediction mode 1312, the encoder may fall back to signaling selected intra prediction mode 1312 using a fixed-length binarization of the index of selected intra prediction mode 1312. The decoder may then decode the signaling for selected intra prediction mode 1312 without any independent determination of intra prediction mode 1316. In an embodiment, the decoder may not perform intra prediction of the at least one of neighboring regions 1306 and 1308 based on the encoder signaling selected intra prediction mode 1312 using a fixed-length binarization.

[0112] It should be noted that the encoder may entropy encode each of the above mentioned signaling approaches of selected intra prediction mode 1312 prior to signaling selected intra prediction mode 1312 to the decoder. For example, the encoder may apply context adaptive variable length coding (CAVLC), context adaptive binary arithmetic coding (CABAC), and/or syntax-based context-based binary arithmetic coding (SBAC).

[0113] To reduce coding complexity at the encoder and/or the decoder, applied intra prediction modes 1314 for intra prediction of the at least one of neighboring regions 1306 and 1308 may comprise a subset of applied intra prediction modes 1310 for intra prediction of current block 1304 as shown in FIG. 16. For example, applied intra prediction modes 1314 may comprise only a subset of the angular intra prediction modes in applied intra prediction modes 1310. For example, applied intra prediction modes 1314 may comprise only one of the horizontal or vertical angular prediction modes from applied intra prediction modes 1310. In another example, applied intra prediction modes 1314 for may be determined based on applied intra prediction modes 1310 that are included in an MPM list constructed for intra prediction of current block 1304. For example, applied intra prediction modes 1314 may comprise only the subset of applied intra prediction modes 1310 that are included in an MPM list constructed for intra prediction of current block 1304. The MPM list may be adaptively generated for current block 1304 based on, for example, the availability and indices of intra prediction modes of the top and left neighboring blocks of current block 1304. The MPM list may be constructed in accordance with HEVC, VVC, or any other video coding standard. More general than an MPM list, applied intra prediction modes 1314 may comprise a candidate list of intra prediction modes determined according to any reasonable criteria, such as a list of intra prediction modes determined to be most likely to be used for intra prediction of current block 1304. In another example to reduce complexity, only one of neighboring regions 1306 and 1308 may be used to determine selected intra prediction mode 1316.

[0114] FIG. 17 illustrates a flowchart 1700 of a method for signaling an intra prediction mode for a block in accordance with embodiments of the present disclosure. The method of flowchart 1700 may be implemented by an encoder, such as encoder 200 in FIG. 2.

[0115] The method of flowchart 1700 begins at 1702. At 1702, reference samples for intra prediction of a block (also referred to as a current block) may be determined. The block may be a picture within a video sequence. For constructing the set of reference samples, available samples from neighboring blocks of the current block may be used as discussed above with respect to FIG. 9. Unavailable ones of the reference samples may be filled with available ones of the reference samples. For example, an unavailable reference sample may be filled with a

nearest available reference sample determined by moving in a clock-wise direction through the reference samples from the position of the unavailable reference. If no reference samples are available, the reference samples may be filled with the mid-value of the dynamic range of the picture being coded. In other examples, the reference samples may be determined in a different manner.

[0116] At 1704, a prediction of at least one neighboring region of the current block may be generated for each of a plurality of intra prediction modes. A neighboring region may comprise a region of available samples as described above with respect to FIG. 13 above. The plurality of intra prediction modes may comprise, for example, the 35 intra prediction modes of HEVC, 67 intra prediction modes of VVC, the intra prediction modes of some other proprietary or industry video coding standard, or a subset of one of these sets of intra prediction modes. In another example, the plurality of intra prediction modes may be determined based on an MPM list constructed for intra prediction of the current block. For example, the plurality of intra prediction modes may comprise only the intra prediction modes included in an MPM list constructed for intra prediction of the current block.

[0117] At step 1706, a first intra prediction mode from the plurality of intra prediction modes may be determined for the at least one neighboring region based on the predictions of the at least one neighboring region generated at 1704. For example, for each of the plurality of intra prediction modes applied at 1704, a prediction error may be determined for the at least one neighboring region based on a difference (e.g., sum of squared differences (SSD), sum of absolute differences (SAD), or sum of absolute transformed differences (SATD)) between the prediction samples determined for the intra prediction mode and reconstructed samples of the at least one neighboring region. An intra prediction mode from the plurality of intra prediction modes may be determined based on the prediction errors. For example, an intra prediction mode from the plurality of intra prediction modes may be selected that results in the smallest prediction error for the at least one neighboring region. It should be noted that other selection criteria may be used to select intra prediction mode from the plurality of intra prediction modes.

[0118] At step 1708, a second intra prediction mode, determined based on the reference samples, for the block may be signaled based on the first intra prediction mode. The first intra prediction mode determined for the at least one neighboring region at 1707 may be compared to second intra prediction mode determined for the current block. Based on the comparison, the second intra prediction mode for the current block may be signaled in a more efficient manner than prior signaling techniques.

[0119] For example, based on the first intra prediction mode determined for the at least one neighboring region being the same as the second intra prediction mode determined for the current block, the second intra prediction mode may be signaled using one or more fewer bits than used to signal a fixed-length binarization of the second intra prediction mode. In an embodiment, the second intra prediction mode may be signaled using a single, one-bit flag based on the first intra prediction mode determined for the at least one neighboring region being the same as the second intra prediction mode determined for the current block. The signaling for the second intra prediction mode may indicate to a decoder that the first intra prediction mode determined for the at least one neighboring region is the same as the second intra prediction mode determined for the current block.

[0120] In another example, based on first intra prediction mode determined for the at least one neighboring region being different than the second intra prediction mode determined for the current block, the second intra prediction mode may be signaled based on a difference between an index of the first intra prediction mode and an index of the second intra prediction mode. For example, the signaling for the second intra prediction mode may indicate to a decoder the difference between an index of the first intra prediction mode and an index of the second intra prediction mode. In another example, based on the first intra prediction mode being different than second intra prediction mode, the second intra prediction mode may be signaled using a fixed-length binarization of the index of second intra prediction mode.

[0121] FIG. 18 illustrates a flowchart 1800 of a method for signaling an intra prediction mode for a block in accordance with embodiments of the present disclosure. The method of flowchart 1800 may be implemented by a decoder, such as decoder 300 in FIG. 3.

[0122] The method of flowchart 1800 begins at 1802. At 1802, reference samples for intra prediction of a block (also referred to as a current block) may be determined. The block may be a picture within a video sequence. For constructing the set of reference samples, available samples from neighboring blocks of the current block may be used as discussed above with respect to FIG. 9. Unavailable ones of the reference samples may be filled with available ones of the reference samples. For example, an unavailable reference sample may be filled with a nearest available reference sample determined by moving in a clock-wise direction through the reference samples from the position of the unavailable reference. If no reference samples are available, the reference samples may be filled with the mid-value of the dynamic range of the picture being coded. In other examples, the reference samples may be determined in a different manner.

[0123] At 1804, a prediction of at least one neighboring region of the current block may be generated for each of a plurality of intra prediction modes. A neighboring region may comprise a region of available samples as described above with respect to FIG. 13 above. The plurality of intra prediction modes may comprise, for example, the 35 intra prediction modes of HEVC, 67 intra prediction modes of VVC, the intra prediction modes of some other proprietary or industry video coding standard, or a subset of one of these sets of intra prediction modes. In another example, the plurality of intra prediction modes may be determined based on an MPM list constructed for intra prediction of the current block. For example, the plurality of intra prediction modes may comprise only the intra prediction modes included in an MPM list constructed for intra prediction of the current block.

[0124] At step 1806, a first intra prediction mode from the plurality of intra prediction modes may be determined for the at least one neighboring region based on the predictions of the at least one neighboring region generated at 1804. For example, for each of the plurality of intra prediction modes applied at 1804, a prediction error may be determined for the at least one neighboring region based on a difference (e.g., sum of squared differences (SSD), sum of absolute differences (SAD), or sum of absolute transformed differences (SATD)) between the prediction samples determined for the intra prediction mode and reconstructed samples of the at least one neighboring region. An intra prediction mode from the plurality of intra prediction modes may be determined based on the prediction errors. For example, an intra prediction mode from the plurality of intra prediction modes may be selected that results in the smallest prediction error for the at least one neighboring region. It should be noted that other selection criteria may be used to select intra prediction mode from the plurality of intra prediction modes.

[0125] At step 1808, a second intra prediction mode for the current block may be decoded based on: signaling information received in an encoded bit stream; and the first intra prediction mode determined for the at least one neighboring region at 1806. For example, based on the signaling information indicating that the first intra prediction mode is the same as the second intra prediction mode, the second intra prediction mode may be decoded as the first intra prediction mode. In another example, based on the signaling information indicating a difference between an index of the first intra prediction mode and an index of the second intra prediction mode, the difference may be added or subtracted from the first intra prediction mode. The resulting value of the addition or subtraction may be used as the index of the second intra prediction mode. After the second intra prediction mode is decoded, a prediction for the current block may generated based on the decoded second intra prediction mode. The prediction may be

combined with a prediction error of the current block received in the encoded video bit stream to decode the current block.

[0126] Embodiments of the present disclosure may be implemented in hardware using analog and/or digital circuits, in software, through the execution of instructions by one or more general purpose or special-purpose processors, or as a combination of hardware and software. Consequently, embodiments of the disclosure may be implemented in the environment of a computer system or other processing system. An example of such a computer system 1900 is shown in FIG. 19. Blocks depicted in the figures above, such as the blocks in FIGS. 1, 2, and 3, may execute on one or more computer systems 1900. Furthermore, each of the steps of the flowcharts depicted in this disclosure may be implemented on one or more computer systems 1900.

[0127] Computer system 1900 includes one or more processors, such as processor 1904. Processor 1904 may be, for example, a special purpose processor, general purpose processor, microprocessor, or digital signal processor. Processor 1904 may be connected to a communication infrastructure 1902 (for example, a bus or network). Computer system 1900 may also include a main memory 1906, such as random access memory (RAM), and may also include a secondary memory 1908.

[0128] Secondary memory 1908 may include, for example, a hard disk drive 1910 and/or a removable storage drive 1912, representing a magnetic tape drive, an optical disk drive, or the like. Removable storage drive 1912 may read from and/or write to a removable storage unit 1916 in a well-known manner. Removable storage unit 1916 represents a magnetic tape, optical disk, or the like, which is read by and written to by removable storage drive 1912. As will be appreciated by persons skilled in the relevant art(s), removable storage unit 1916 includes a computer usable storage medium having stored therein computer software and/or data.

[0129] In alternative implementations, secondary memory 1908 may include other similar means for allowing computer programs or other instructions to be loaded into computer system 1900. Such means may include, for example, a removable storage unit 1918 and an interface 1914. Examples of such means may include a program cartridge and cartridge interface (such as that found in video game devices), a removable memory chip (such as an EPROM or PROM) and associated socket, a thumb drive and USB port, and other removable storage units 1918 and interfaces 1914 which allow software and data to be transferred from removable storage unit 1918 to computer system 1900.

[0130] Computer system 1900 may also include a communications interface 1920. Communications interface 1920 allows software and data to be transferred between computer

system 1900 and external devices. Examples of communications interface 1920 may include a modem, a network interface (such as an Ethernet card), a communications port, etc. Software and data transferred via communications interface 1920 are in the form of signals which may be electronic, electromagnetic, optical, or other signals capable of being received by communications interface 1920. These signals are provided to communications interface 1920 via a communications path 1922. Communications path 1922 carries signals and may be implemented using wire or cable, fiber optics, a phone line, a cellular phone link, an RF link, and other communications channels.

[0131] As used herein, the terms "computer program medium" and "computer readable medium" are used to refer to tangible storage media, such as removable storage units 1916 and 1918 or a hard disk installed in hard disk drive 1910. These computer program products are means for providing software to computer system 1900. Computer programs (also called computer control logic) may be stored in main memory 1906 and/or secondary memory 1908. Computer programs may also be received via communications interface 1920. Such computer programs, when executed, enable the computer system 1900 to implement the present disclosure as discussed herein. In particular, the computer programs, when executed, enable processor 1904 to implement the processes of the present disclosure, such as any of the methods described herein. Accordingly, such computer programs represent controllers of the computer system 1900.

[0132] In another embodiment, features of the disclosure may be implemented in hardware using, for example, hardware components such as application-specific integrated circuits (ASICs) and gate arrays. Implementation of a hardware state machine to perform the functions described herein will also be apparent to persons skilled in the relevant art(s).

CLAIMS

1. A method for decoding video data, comprising:

determining a most probable mode (MPM) list for intra prediction of a block;

generating, for each of a plurality of intra prediction modes determined based on the MPM list, a prediction of at least one neighboring region of the block based on reference samples;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes; and

decoding a second intra prediction mode for the block based on:

the first intra prediction mode; and

signaling information in an encoded video bitstream.

generating, for the second intra prediction mode, a prediction of the block from the reference samples.

2. A method comprising:

generating, for each of a plurality of intra prediction modes determined based on a most probable mode (MPM) list for intra prediction of a block, a prediction of at least one neighboring region of the block;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes; and

decoding a second intra prediction mode for the block based on the first intra prediction mode.

3. A method comprising:

generating, for each of a plurality of intra prediction modes determined based on a most probable mode (MPM) list for intra prediction of a block, a prediction of at least one neighboring region of the block;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes; and

decoding a second intra prediction mode for the block based on:

the first intra prediction mode; and

signaling information in an encoded video bitstream.

4. The method of claim 3, further comprising determining reference samples for intra prediction of the block.

5. The method of claim 3, wherein the determining the first intra prediction mode further comprises determining the first intra prediction mode, from the plurality of intra prediction modes, that results in a smallest prediction residual for the at least one neighboring region.

6. The method of claim 3, wherein the signaling information indicates a difference between an index of the first intra prediction mode and an index of the second intra prediction mode.

7. The method of claim 3, wherein the signaling information indicates that the second intra prediction mode is the same as the first intra prediction mode.

8. The method of claim 7, wherein the signaling information is a single, one-bit flag.

9. The method of claim 1, further comprising generating, for the second intra prediction mode, a prediction of the block from reference samples.

10. The method of claim 9, further comprising reconstructing the block from the prediction of the block.

11. The method of claim 3, wherein the generating, for each of the plurality of intra prediction modes, the prediction of the at least one neighboring region comprises determining a prediction for a sample of the at least one neighboring region from reference samples based on a location of the sample in the at least one neighboring region and the intra prediction mode.

12. The method of claim 11, wherein the prediction for the sample is determined based on an interpolation between values of two of the reference samples.

13. The method of claim 12, wherein weighting factors, determined based on the intra prediction mode and the location of the sample in the at least one neighboring region, are applied to the values of the two of the reference samples to perform the interpolation.

14. The method of claim 3, wherein the reference samples comprise samples determined from at least one of:

the column of samples immediately adjacent to the left of the block; or

the row of samples immediately adjacent to the top of the block

15. The method of claim 3, wherein the at least one neighboring region comprises at least one of:

a first neighboring region adjacent to samples in the top of samples of the block; or

a second neighboring region adjacent to samples in the left column of samples of the block.

16. The method of claim 3, wherein the generating is performed based on the signaling information.

17. A decoder comprising:

one or more processors; and

memory storing instructions that, when executed by the one or more processors, cause the decoder to perform the method of any one of claims 1 to 16.

18. A non-transitory computer-readable medium comprising instructions that, when executed by one or more processors, cause the one or more processors to perform the method of any of claims 1 to 16.

19. A method for encoding video data, comprising:

determining a most probable mode (MPM) list for intra prediction of a block;

generating, for each of a plurality of intra prediction modes determined based on the MPM list, a prediction of at least one neighboring region of the block based on reference samples;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes; and

signaling a second intra prediction mode, determined based on the reference samples, for the block based on the first intra prediction mode.

20. A method comprising:

generating, for each of a plurality of intra prediction modes determined based on a most probable mode (MPM) list for intra prediction of a block, a prediction of at least one neighboring region of the block;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode; and

signaling a second intra prediction mode for the block based on the first intra prediction mode.

21. A method comprising:

determining a most probable mode (MPM) list for intra prediction of a block;

generating, for each of a plurality of intra prediction modes determined based on the MPM list, a prediction of at least one neighboring region of the block;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes; and

signaling a second intra prediction mode for the block based on the first intra prediction mode.

22. The method of claim 21, further comprising determining reference samples for intra prediction of the block.

23. The method of claim 21, wherein the determining the first intra prediction mode further comprises determining the first intra prediction mode, from the plurality of intra prediction modes, that results in a smallest prediction residual for the at least one neighboring region.

24. The method of claim 21, wherein the signaling further comprises signaling the second intra prediction mode using a single, one-bit flag based on the first intra prediction mode being the same as the second intra prediction mode.

25. The method of claim 21, wherein the signaling further comprises signaling the second intra prediction mode based on a difference between an index of the first intra prediction mode and an index of the second intra prediction mode.

26. The method of claim 21, wherein the signaling further comprises signaling the second intra prediction mode as a fixed-length binarization of the second intra prediction mode based on the first intra prediction mode being different than the second intra prediction mode.

27. The method of claim 21, further comprising:

generating, for each of the plurality of intra prediction modes, a prediction of the block from reference samples; and

determining, based on the predictions of the block, the second intra prediction mode from the plurality of intra prediction modes.

28. The method of claim 21, wherein the determining the second intra prediction mode further comprises determining the second intra prediction mode, from the plurality of intra prediction modes, that results in a smallest prediction residual for the block.

29. The method of claim 21, wherein the determining the second intra prediction mode further comprises determining the second intra prediction mode, from the plurality of intra prediction modes and one or more additional intra prediction modes.

30. The method of claim 21, wherein the generating, for each of the plurality of intra prediction modes, the prediction of the at least one neighboring region comprises determining a prediction for a sample of the at least one neighboring region from reference samples based on a location of the sample in the at least one neighboring region and the intra prediction mode.

31. The method of claim 30, wherein the prediction for the sample is determined based on an interpolation between values of two of the reference samples.

32. The method of claim 31, wherein weighting factors, determined based on the intra prediction mode and the location of the sample in the at least one neighboring region, are applied to the values of the two of the reference samples to perform the interpolation.

33. The method of claim 21, further comprising determining reference samples for intra prediction of the block. wherein the reference samples comprise samples determined from at least one of:

the column of samples immediately adjacent to the left of the block; or

the row of samples immediately adjacent to the top of the block

34. The method of claim 21, wherein the at least one neighboring region comprises at least one of:

a first neighboring region adjacent to samples in the top row of samples of the block; or

a second neighboring region adjacent to samples in the left column of samples of the block.

35. An encoder comprising:

one or more processors; and

memory storing instructions that, when executed by the one or more processors, cause the encoder to perform the method of any one of claims 19 to 34.

36. A non-transitory computer-readable medium comprising instructions that, when executed by one or more processors, cause the one or more processors to perform the method of any of claims 19 to 34.

37. A method for decoding video data, comprising:

determining reference samples for intra prediction of a block;

generating, for each of a plurality of intra prediction modes, a prediction of at least one neighboring region of the block from the reference samples;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes for the at least one neighboring region;

decoding a second intra prediction mode for the block based on:

the first intra prediction mode; and

signaling information in an encoded video bitstream; and

generating, for the second intra prediction mode, a prediction of the block from the reference samples.

38. A method comprising:

generating, for each of a plurality of intra prediction modes, a prediction of at least one neighboring region of a block from reference samples determined for intra prediction of the block;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode; and

decoding a second intra prediction mode for the block based on the first intra prediction mode.

39. A method comprising:

generating, for each of a plurality of intra prediction modes, a prediction of at least one neighboring region of a block from reference samples;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes;

decoding a second intra prediction mode for the block based on:

the first intra prediction mode; and

signaling information in an encoded video bitstream; and

generating, for the second intra prediction mode, a prediction of the block from the reference samples.

40. The method of claim 39, further comprising determining the reference samples for intra prediction of the block.

41. The method of claim 39, wherein the determining the first intra prediction mode further comprises determining the first intra prediction mode, from the plurality of intra prediction modes, that results in a smallest prediction residual for the at least one neighboring region.

42. The method of claim 39, wherein the signaling information indicates a difference between an index of the first intra prediction mode and an index of the second intra prediction mode.

43. The method of claim 39, wherein the signaling information indicates that the second intra prediction mode is the same as the first intra prediction mode.

44. The method of claim 43, wherein the signaling information is a single, one-bit flag.

45. The method of claim 39, wherein the plurality of intra prediction modes are determined based on a most probable mode list determined for intra prediction of the block.

46. The method of claim 39, further comprising reconstructing the block from the prediction of the block.

47. The method of claim 39, wherein the generating, for each of the plurality of intra prediction modes, the prediction of the at least one neighboring region from the reference samples

comprises determining a prediction for a sample of the at least one neighboring region from the reference samples based on a location of the sample in the at least one neighboring region and the intra prediction mode.

48. The method of claim 47, wherein the prediction for the sample is determined based on an interpolation between values of two of the reference samples.

49. The method of claim 48, wherein weighting factors, determined based on the intra prediction mode and the location of the sample in the at least one neighboring region, are applied to the values of the two of the reference samples to perform the interpolation.

50. The method of claim 39, wherein the reference samples comprise samples determined from at least one of:

the column of samples immediately adjacent to the left of the block; or

the row of samples immediately adjacent to the top of the block

51. The method of claim 39, wherein the at least one neighboring region comprises at least one of:

a first neighboring region adjacent to samples in the top of samples of the block; or

a second neighboring region adjacent to samples in the left column of samples of the block.

52. The method of claim 39, wherein the generating is performed based on the signaling information.

53. A decoder comprising:

one or more processors; and

memory storing instructions that, when executed by the one or more processors, cause the decoder to perform the method of any one of claims 39 to 52.

54. A non-transitory computer-readable medium comprising instructions that, when executed by one or more processors, cause the one or more processors to perform the method of any of claims 39 to 52.

55. A method for encoding video data, comprising:

determining reference samples for intra prediction of a block;

generating, for each of a plurality of intra prediction modes, a prediction of at least one neighboring region of the block from the reference samples;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes for the at least one neighboring region; and

signaling a second intra prediction mode, determined based on the reference samples, for the block based on the first intra prediction mode.

56. A method for encoding video data, comprising:

generating, for each of a plurality of intra prediction modes, a prediction of at least one neighboring region of a block from reference samples;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode; and

signaling a second intra prediction mode, determined based on the reference samples, for the block based on the first intra prediction mode.

57. A method for encoding video data, comprising:

generating, for each of a plurality of intra prediction modes, a prediction of at least one neighboring region of the block from the reference samples;

determining, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes for the at least one neighboring region; and

signaling a second intra prediction mode, determined based on the reference samples, for the block based on the first intra prediction mode.

58. The method of claim 57, further comprising determining the reference samples for intra prediction of a block.

59. The method of claim 57, wherein the determining the first intra prediction mode further comprises determining the first intra prediction mode, from the plurality of intra prediction modes, that results in a smallest prediction residual for the at least one neighboring region.

60. The method of claim 57, wherein the signaling further comprises signaling the second intra prediction mode using a single, one-bit flag based on the first intra prediction mode being the same as the second intra prediction mode.

61. The method of claim 57, wherein the signaling further comprises signaling the second intra prediction mode based on a difference between an index of the first intra prediction mode and an index of the second intra prediction mode.

62. The method of claim 57, wherein the signaling further comprises signaling the second intra prediction mode as a fixed-length binarization of the second intra prediction mode based on the first intra prediction mode being different than the second intra prediction mode.

63. The method of claim 57, wherein the plurality of intra prediction modes are determined based on a most probable mode list determined for intra prediction of the block.

64. The method of claim 57, further comprising:

generating, for each of the plurality of intra prediction modes, a prediction of the block from the reference samples; and

determining, based on the predictions of the block, the second intra prediction mode from the plurality of intra prediction modes.

65. The method of claim 57, wherein the determining the second intra prediction mode further comprises determining the second intra prediction mode, from the plurality of intra prediction modes, that results in a smallest prediction residual for the block.

66. The method of claim 57, wherein the determining the second intra prediction mode further comprises determining the second intra prediction mode, from the plurality of intra prediction modes and one or more additional intra prediction modes.

67. The method of claim 57, wherein the generating, for each of the plurality of intra prediction modes, the prediction of the at least one neighboring region from the reference samples comprises determining a prediction for a sample of the at least one neighboring region from the

reference samples based on a location of the sample in the at least one neighboring region and the intra prediction mode.

68. The method of claim 67, wherein the prediction for the sample is determined based on an interpolation between values of two of the reference samples.

69. The method of claim 68, wherein weighting factors, determined based on the intra prediction mode and the location of the sample in the at least one neighboring region, are applied to the values of the two of the reference samples to perform the interpolation.

70. The method of claim 57, wherein the reference samples comprise samples determined from at least one of:

the column of samples immediately adjacent to the left of the block; or

the row of samples immediately adjacent to the top of the block

71. The method of claim 57, wherein the at least one neighboring region comprises at least one of:

a first neighboring region adjacent to samples in the top row of samples of the block; or

a second neighboring region adjacent to samples in the left column of samples of the block.

72. An encoder comprising:

one or more processors; and

memory storing instructions that, when executed by the one or more processors, cause the encoder to perform the method of any one of claims 57 to 71.

73. A non-transitory computer-readable medium comprising instructions that, when executed by one or more processors, cause the one or more processors to perform the method of any of claims 57 to 71.

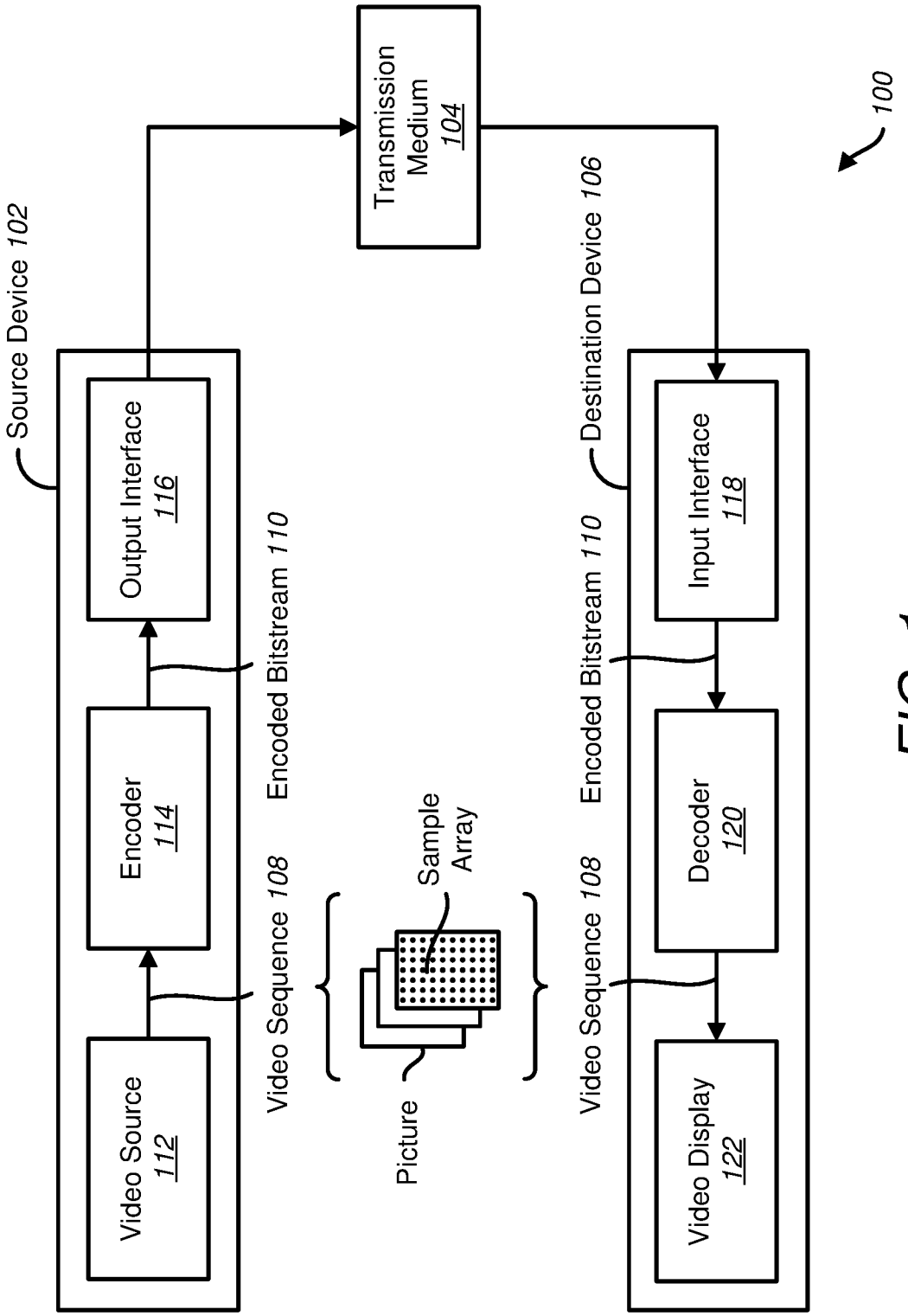

100
Source Device 102
Video Source 112
Encoder 114
Output Interface 116
Transmission Medium 104
Destination Device 106
Input Interface 118
Decoder 120
Video Display 122
Video Sequence 108
Encoded Bitstream 110
Sample Array
Picture

FIG. 1

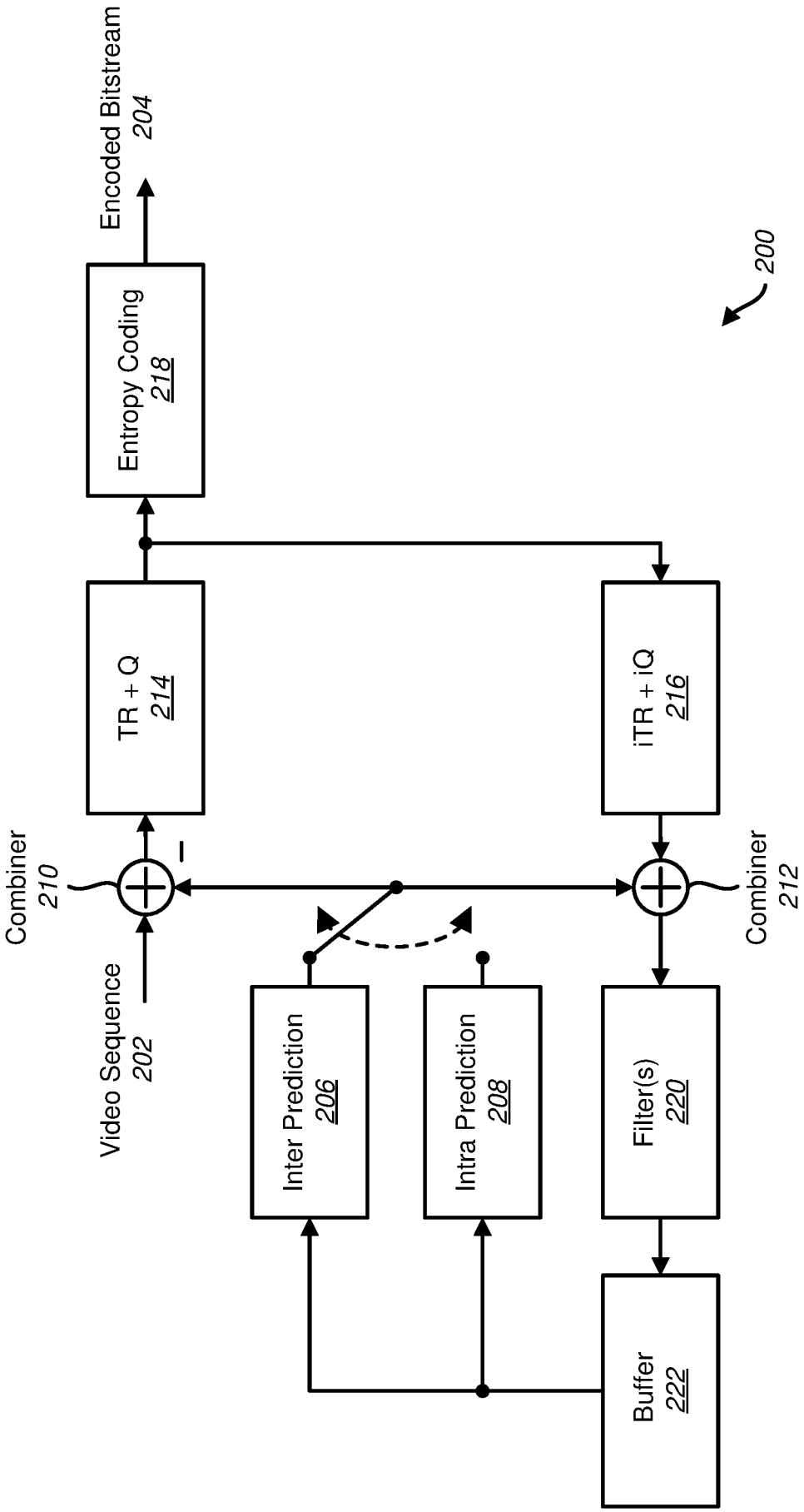

Encoded Bitstream
204
Entropy Coding
218
200
TR + Q
214
iTR + iQ
216
Combiner
210
Combiner
212
Video Sequence
202
Inter Prediction
206
Intra Prediction
208
Filter(s)
220
Buffer
222

FIG. 2

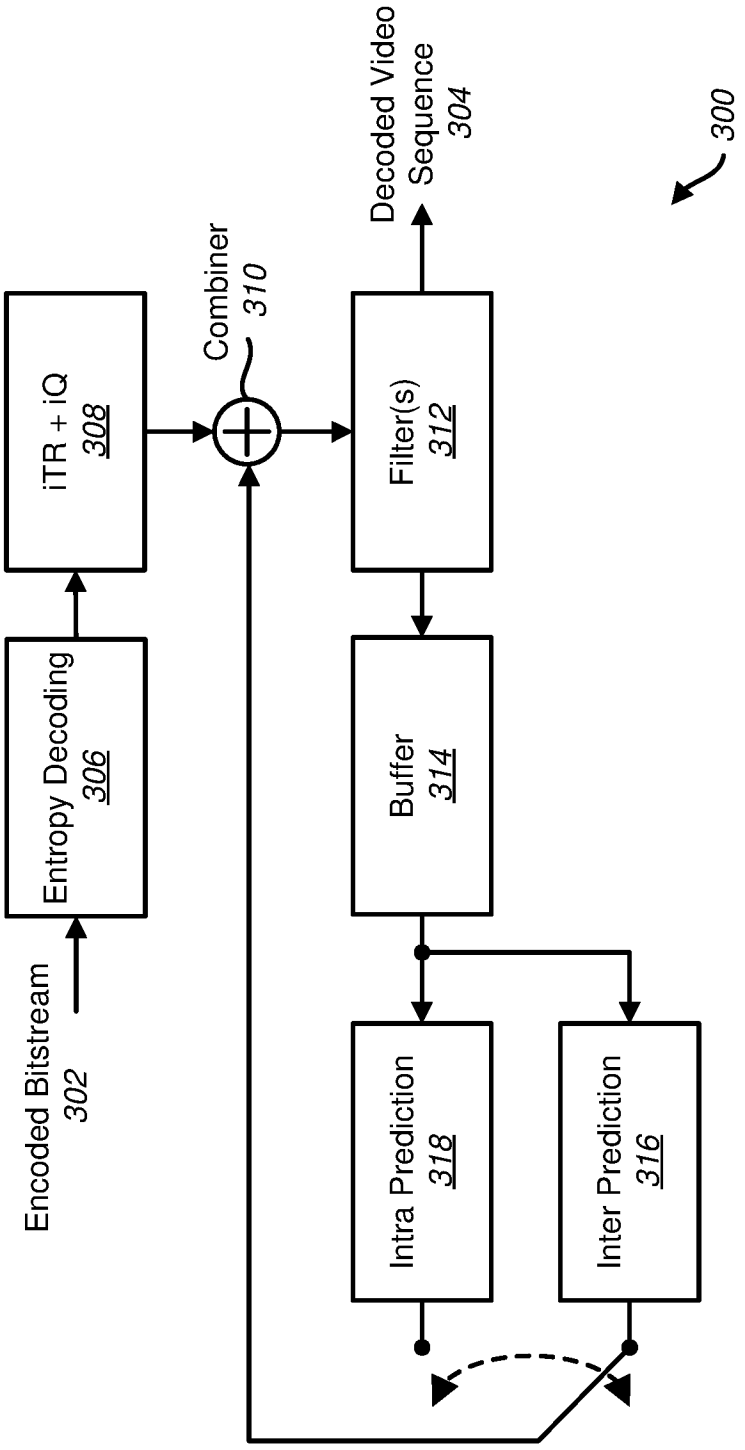

300
Encoded Bitstream
302
Entropy Decoding
306
iTR + iQ
308
Combiner
310
Filter(s)
312
Decoded Video
Sequence
304
Buffer
314
Intra Prediction
318
Inter Prediction
316

FIG. 3

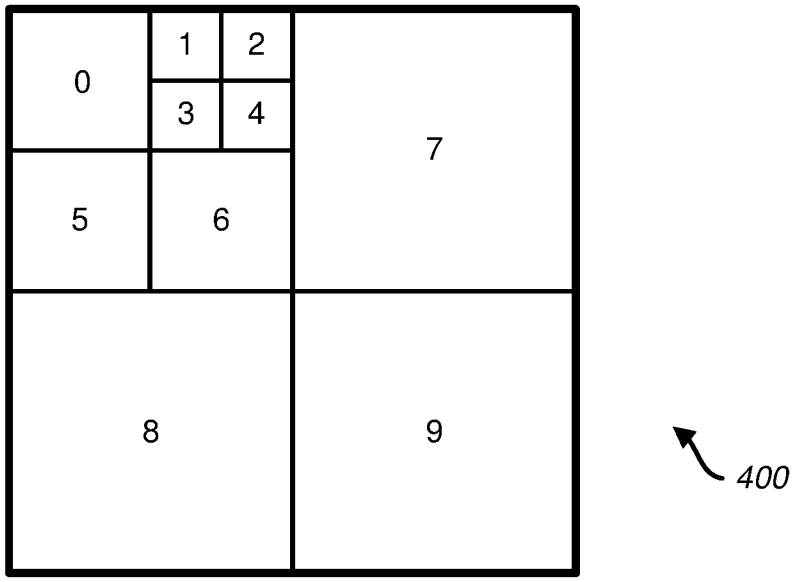


FIG. 4

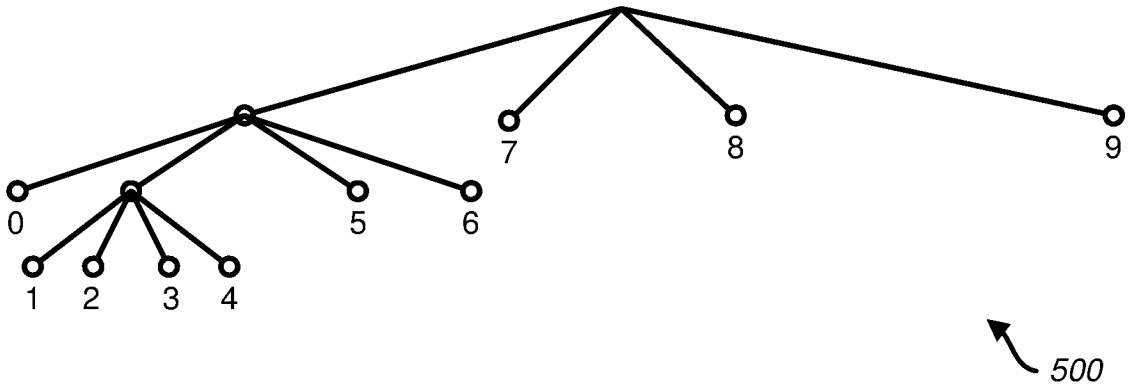
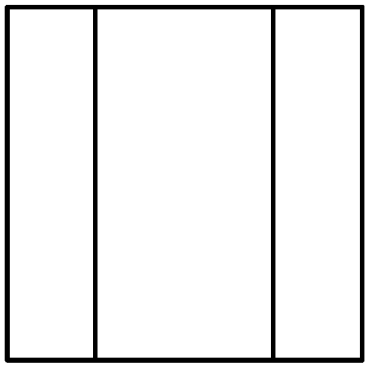
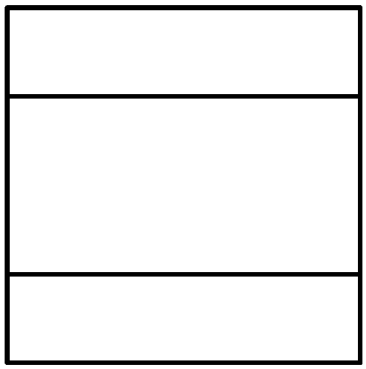


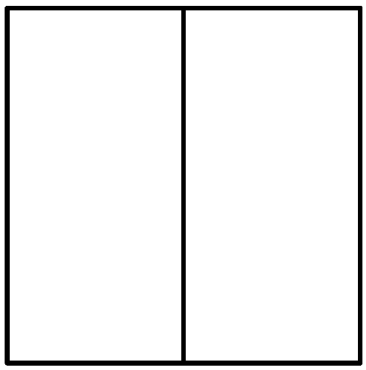
FIG. 5



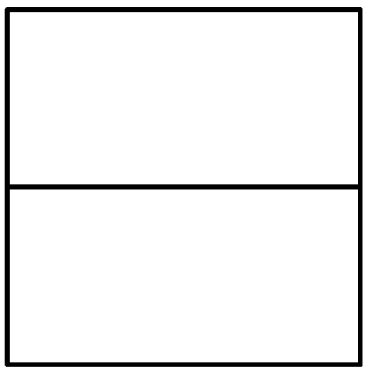
Horizontal Ternary Tree Partition 608



Vertical Ternary Tree Partition 606



Horizontal Binary Tree Partition 604



Vertical Binary Tree Partition 602

FIG. 6

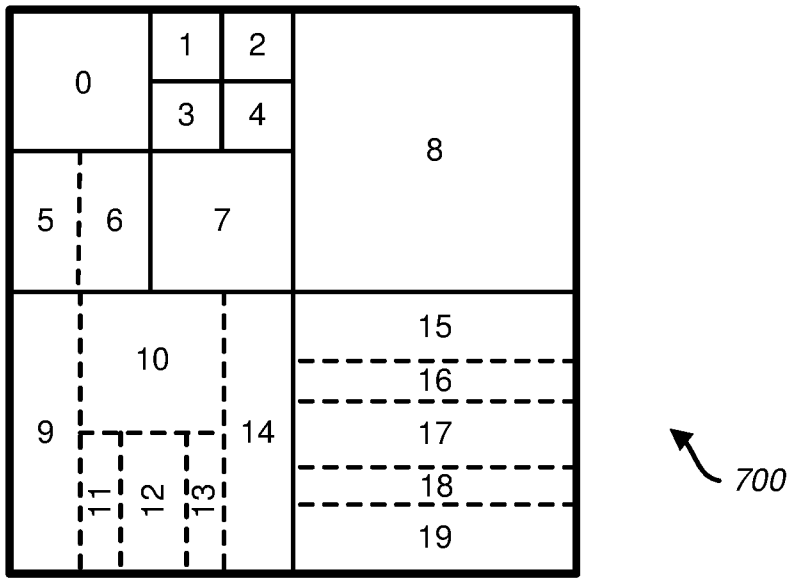

0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
700

FIG. 7

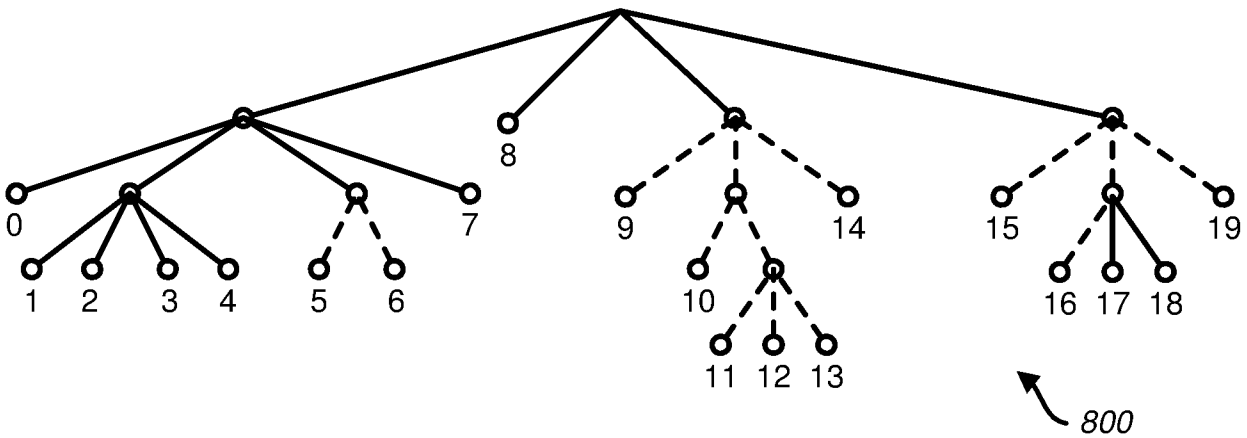

0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
800

FIG. 8

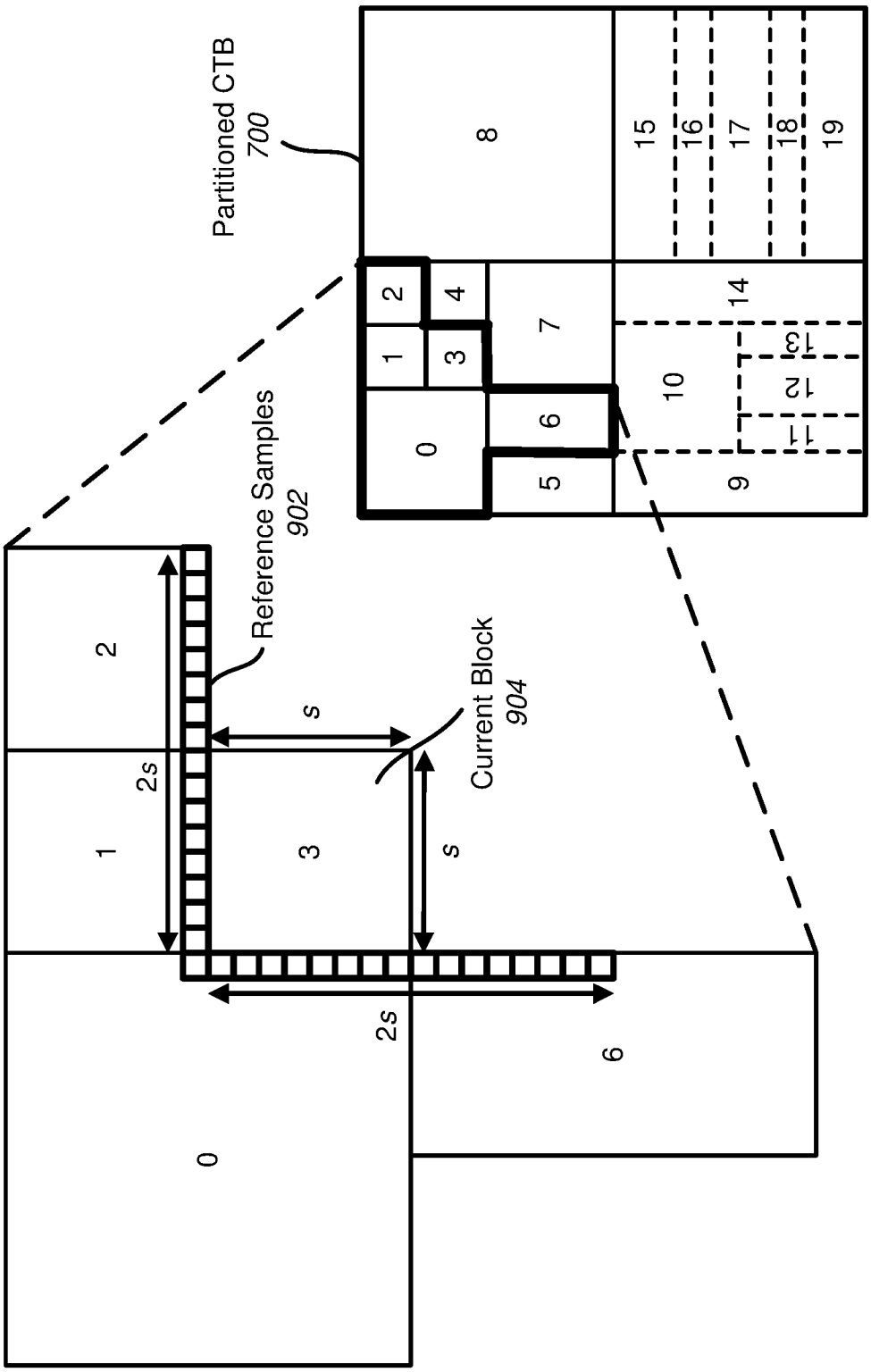

Partitioned CTB
700
Reference Samples
902
Current Block
904
s
2s
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

FIG. 9

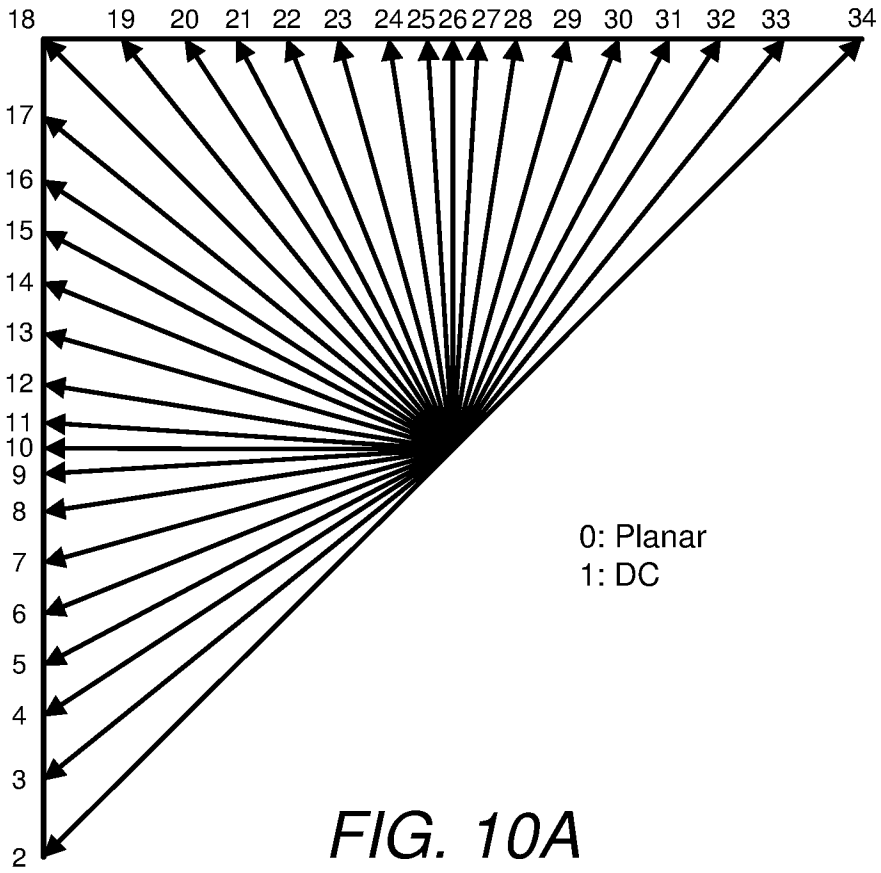


FIG. 10A

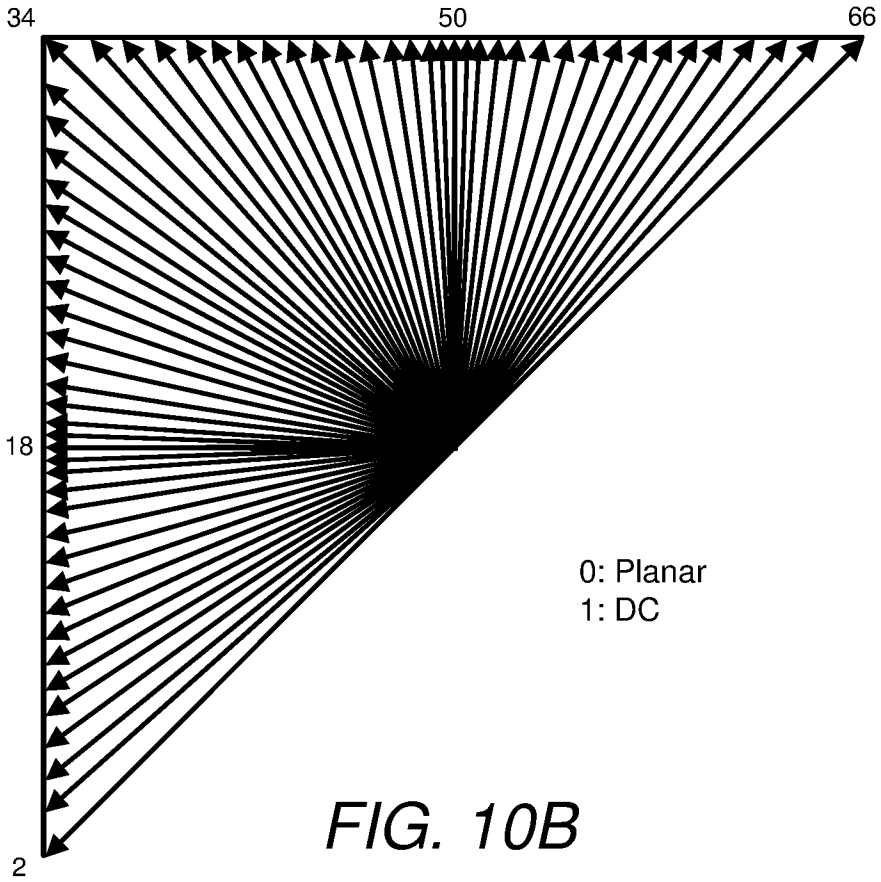


FIG. 10B

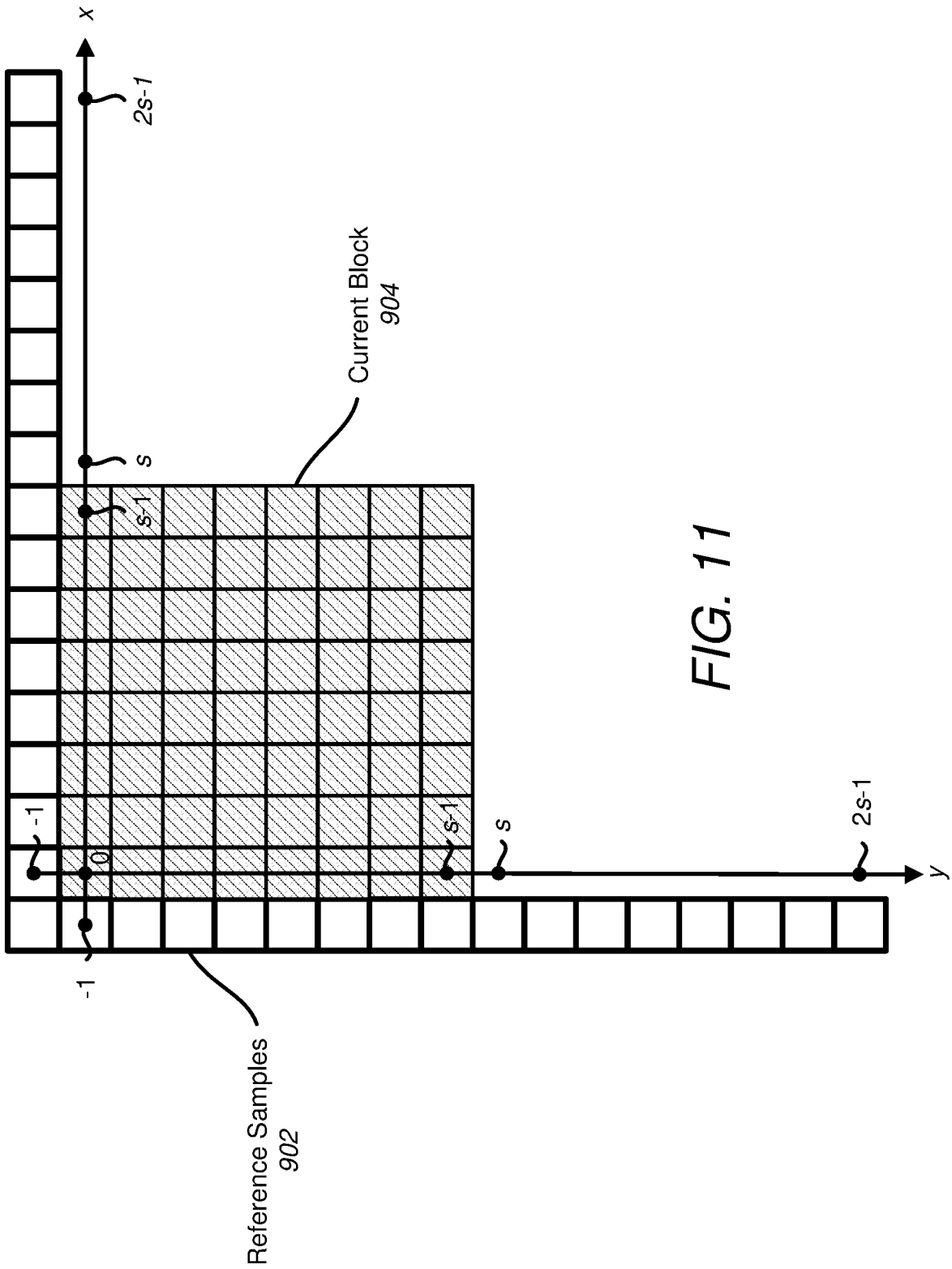

x
2s-1
Current Block
904
s
s-1
-1
0
-1
s-1
s
2s-1
y
Reference Samples
902

FIG. 11

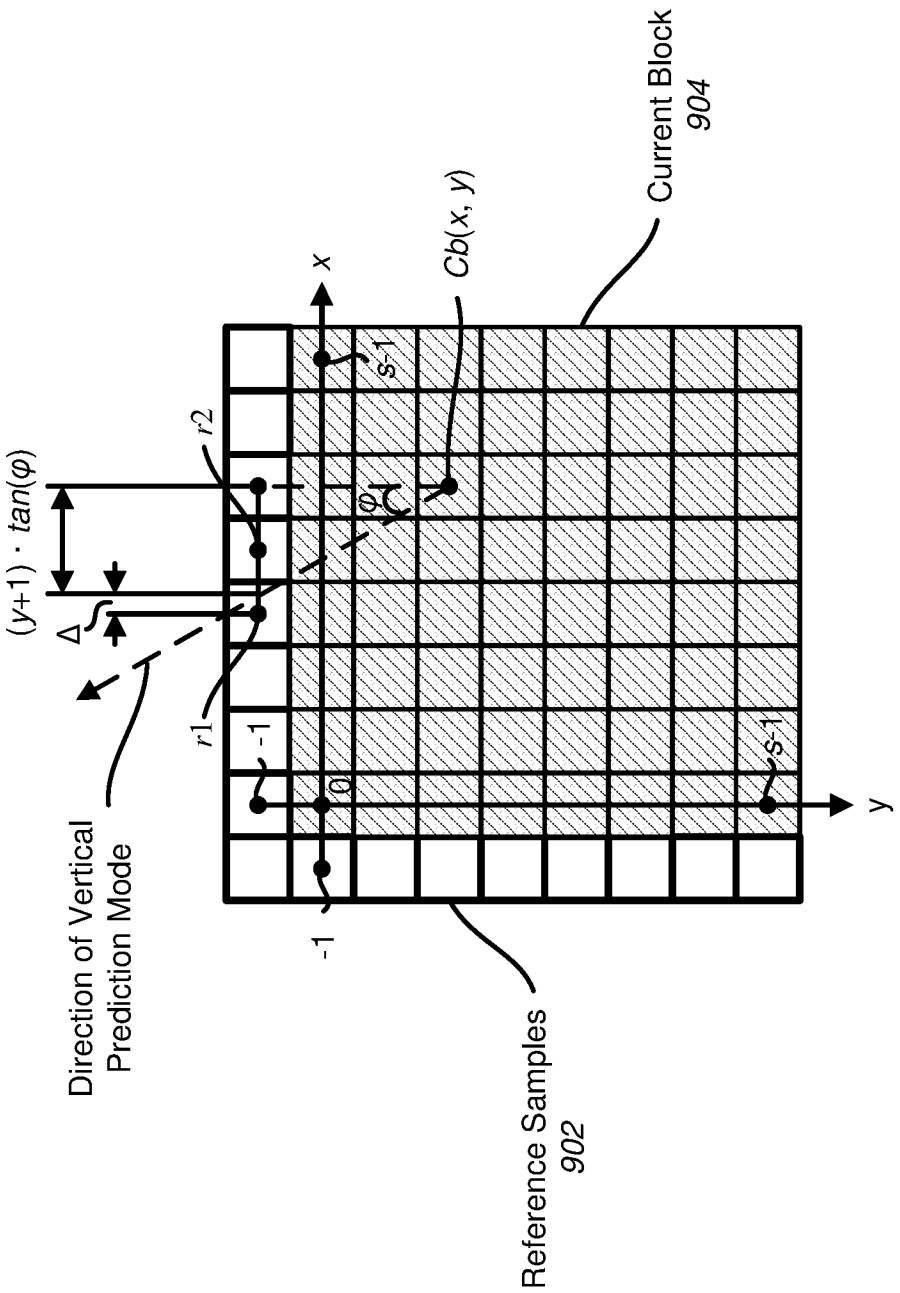

Current Block
904
Cb(x, y)
x
y
s-1
r2
(y+1) · tan(φ)
Δ
r1
-1
0
-1
s-1
Direction of Vertical
Prediction Mode
Reference Samples
902

FIG. 12

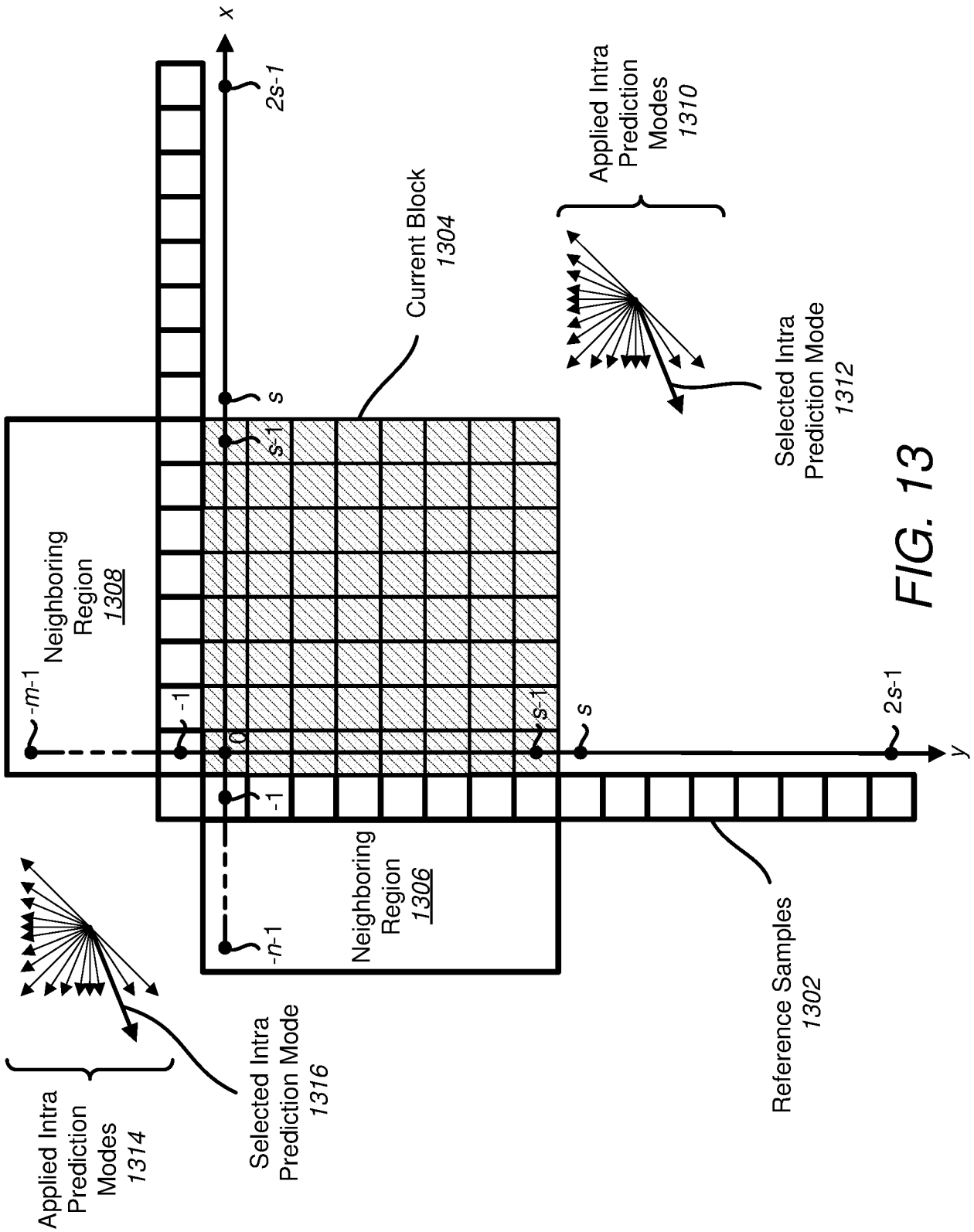

x
2s-1
s
s-1
Current Block
1304
Applied Intra
Prediction
Modes
1310
Selected Intra
Prediction Mode
1312
Neighboring
Region
1308
-m-1
-1
0
-1
s-1
s
2s-1
y
-n-1
Neighboring
Region
1306
Reference Samples
1302
Applied Intra
Prediction
Modes
1314
Selected Intra
Prediction Mode
1316

FIG. 13

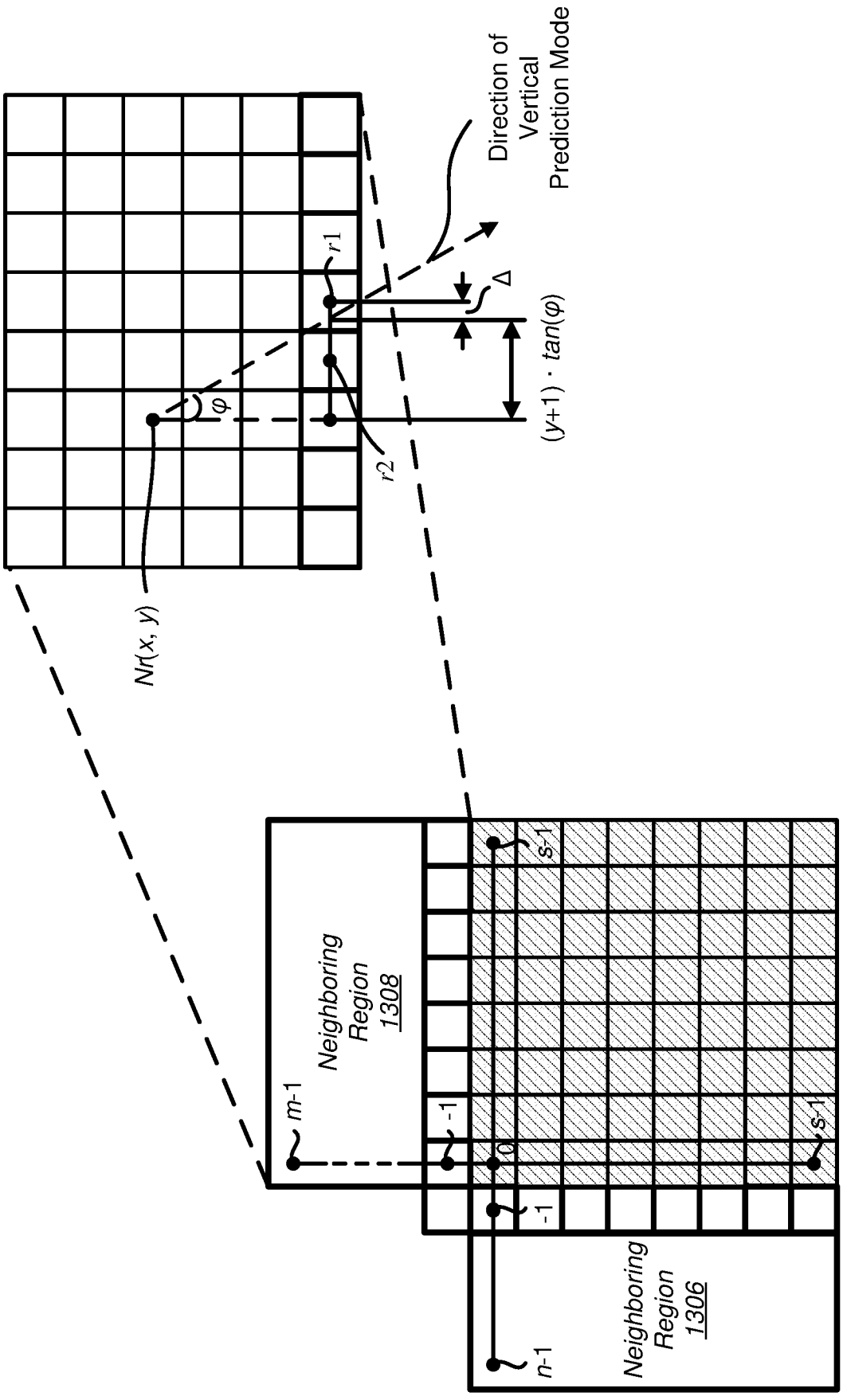

Direction of Vertical Prediction Mode
r1
r2
Δ
(y+1) · tan(φ)
φ
Nr(x, y)
Neighboring Region 1308
Neighboring Region 1306
m-1
-1
0
-1
n-1
s-1
s-1

FIG. 14

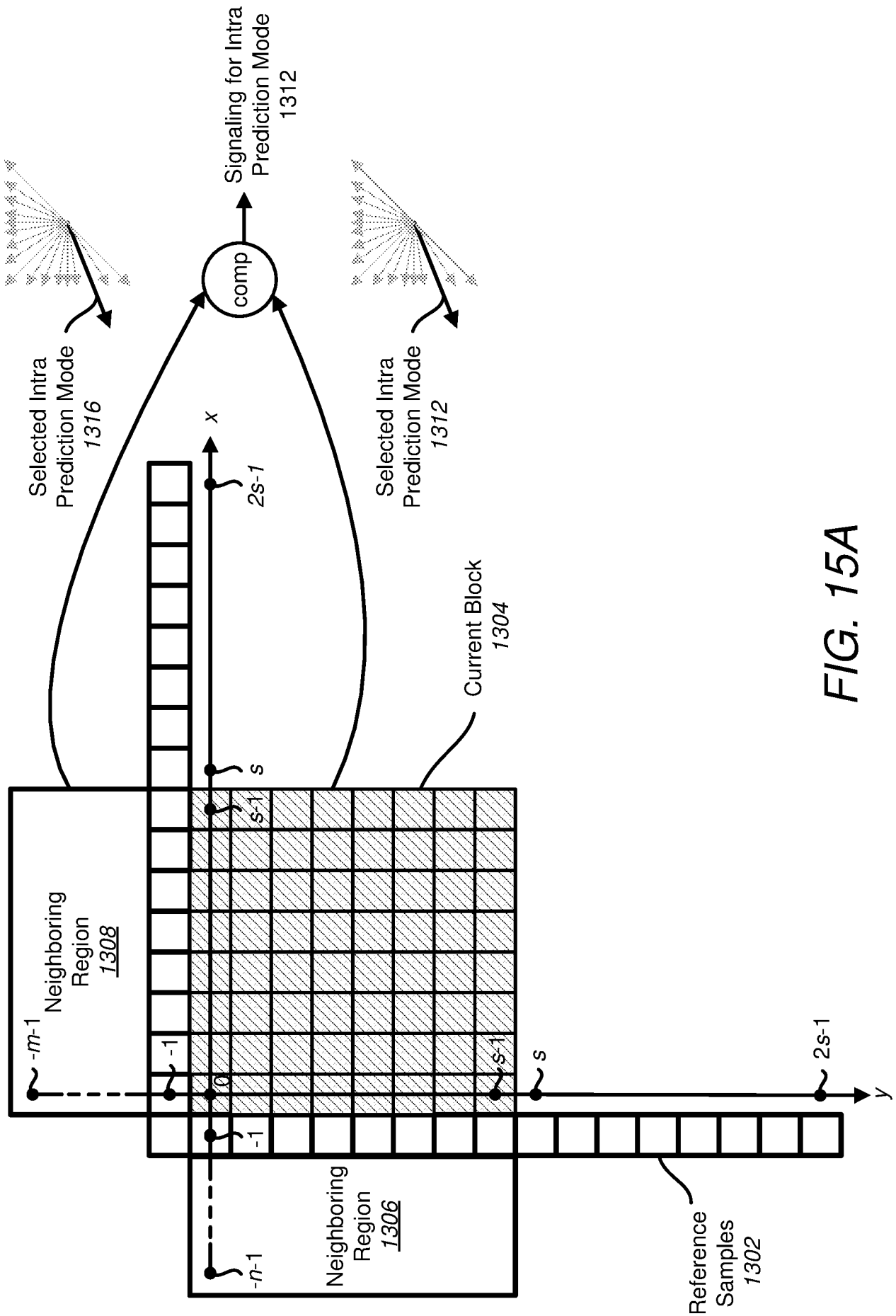

Signaling for Intra Prediction Mode 1312
comp
Selected Intra Prediction Mode 1316
Selected Intra Prediction Mode 1312
x
2s-1
s
s-1
Current Block 1304
Neighboring Region 1308
-m-1
-1
0
-1
s-1
s
2s-1
y
-n-1
Neighboring Region 1306
Reference Samples 1302

FIG. 15A

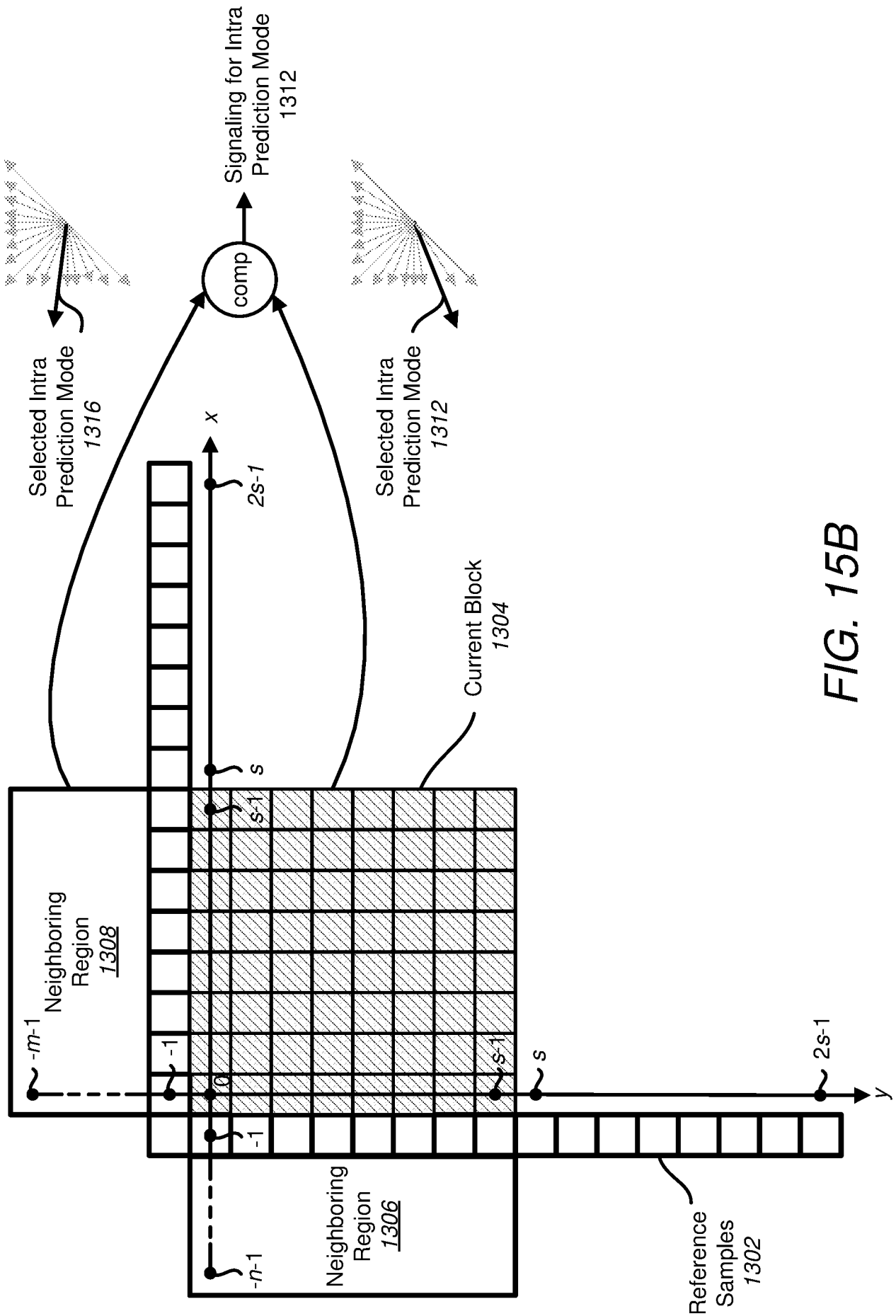

Signaling for Intra Prediction Mode 1312
comp
Selected Intra Prediction Mode 1316
Selected Intra Prediction Mode 1312
x
2s-1
s
s-1
Current Block 1304
Neighboring Region 1308
-m-1
-1
0
-1
s-1
s
2s-1
y
-n-1
Neighboring Region 1306
Reference Samples 1302

FIG. 15B

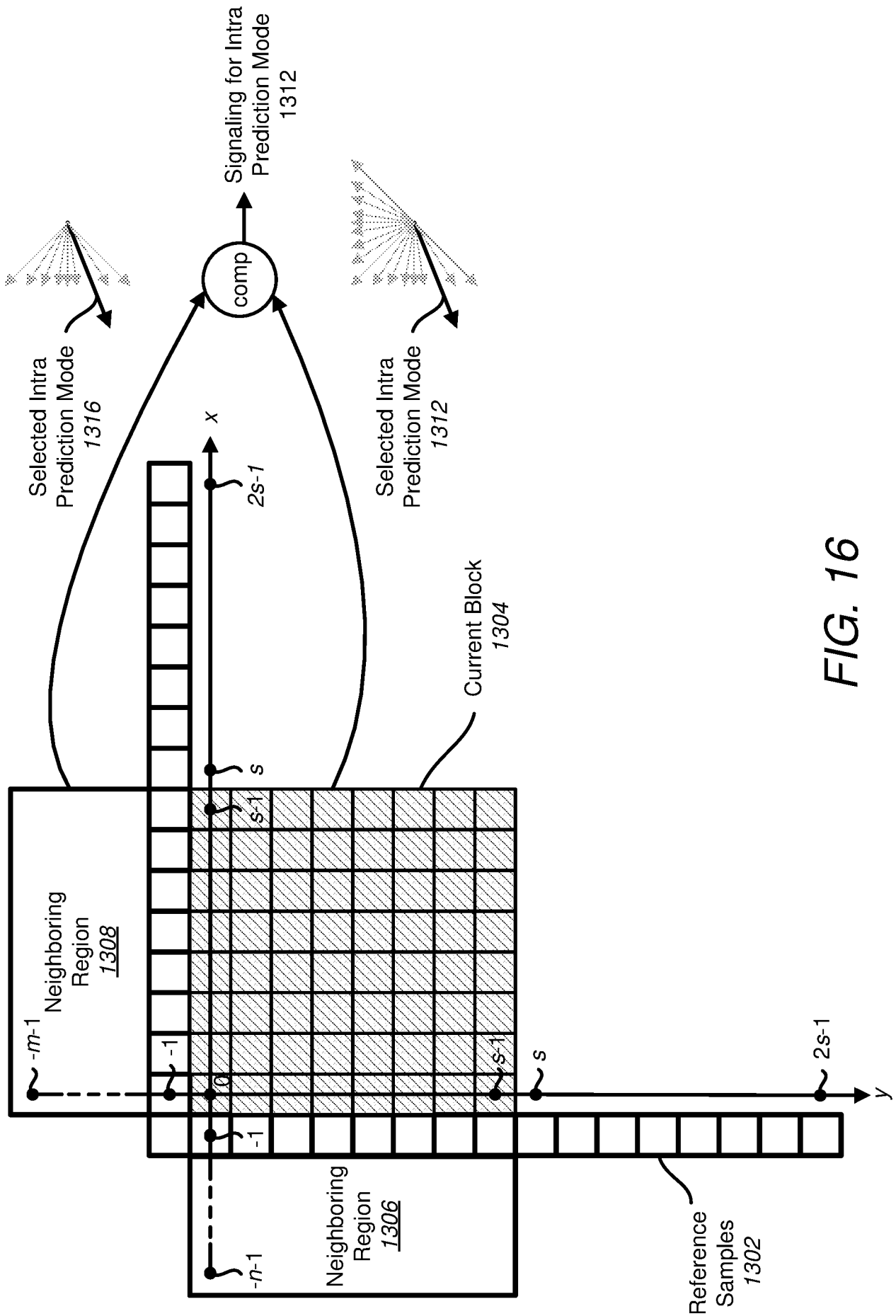

Signaling for Intra Prediction Mode 1312
comp
Selected Intra Prediction Mode 1316
Selected Intra Prediction Mode 1312
x
2s-1
s
s-1
Current Block 1304
Neighboring Region 1308
-m-1
-1
0
-1
s-1
s
2s-1
y
-n-1
Neighboring Region 1306
Reference Samples 1302

FIG. 16

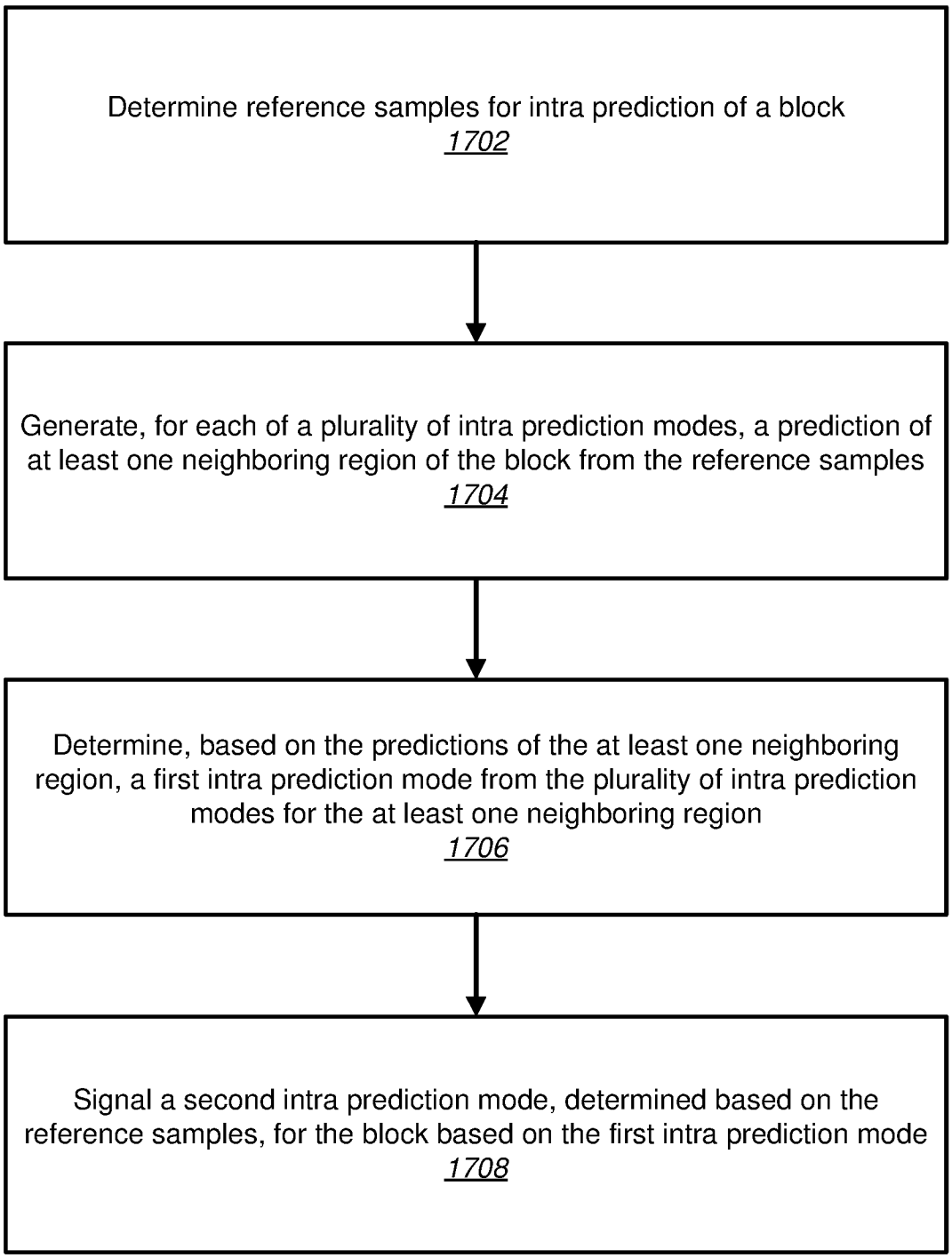

Determine reference samples for intra prediction of a block
1702
Generate, for each of a plurality of intra prediction modes, a prediction of at least one neighboring region of the block from the reference samples
1704
Determine, based on the predictions of the at least one neighboring region, a first intra prediction mode from the plurality of intra prediction modes for the at least one neighboring region
1706
Signal a second intra prediction mode, determined based on the reference samples, for the block based on the first intra prediction mode
1708
1700

FIG. 17

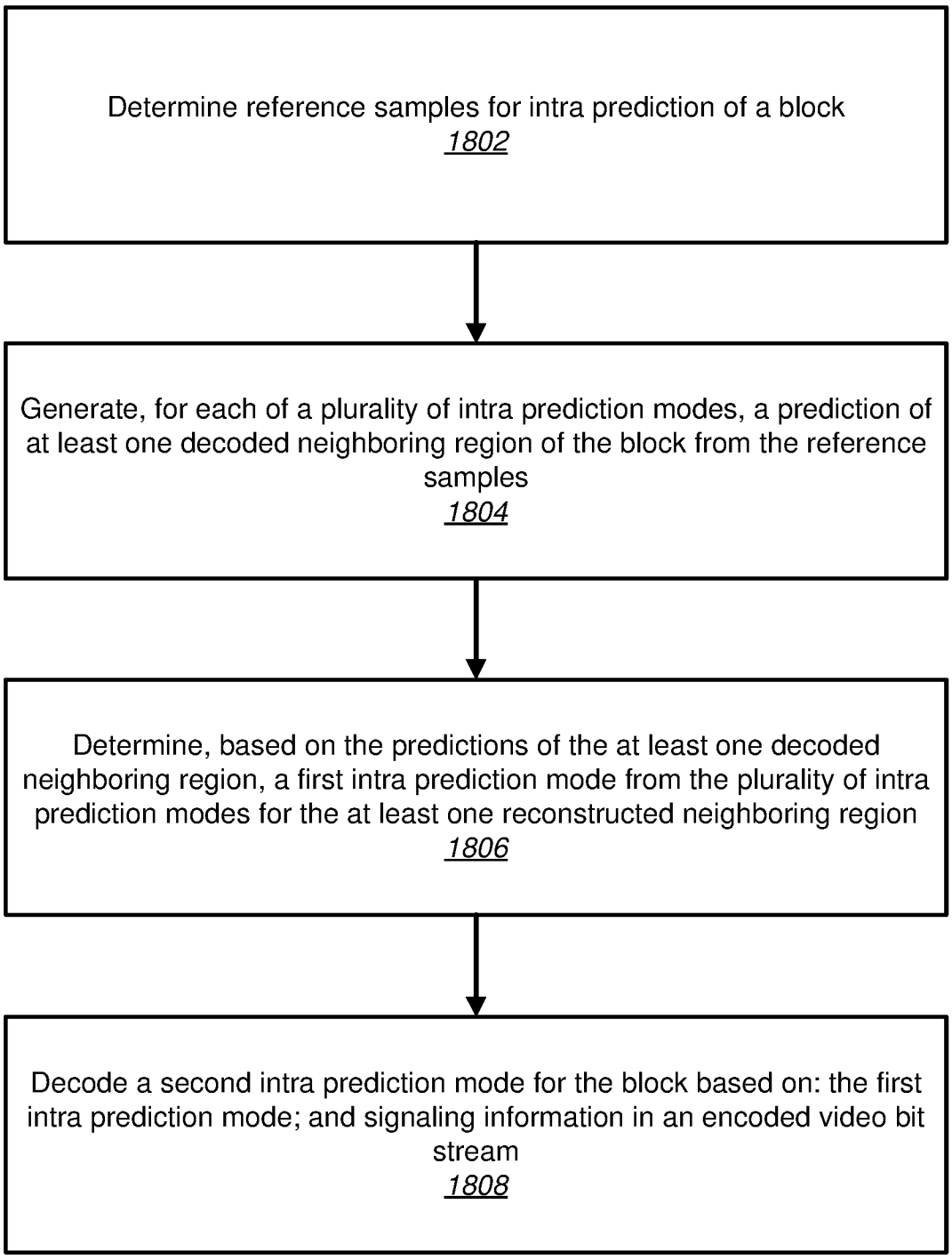

Determine reference samples for intra prediction of a block
1802
Generate, for each of a plurality of intra prediction modes, a prediction of at least one decoded neighboring region of the block from the reference samples
1804
Determine, based on the predictions of the at least one decoded neighboring region, a first intra prediction mode from the plurality of intra prediction modes for the at least one reconstructed neighboring region
1806
Decode a second intra prediction mode for the block based on: the first intra prediction mode; and signaling information in an encoded video bit stream
1808
1800

FIG. 18

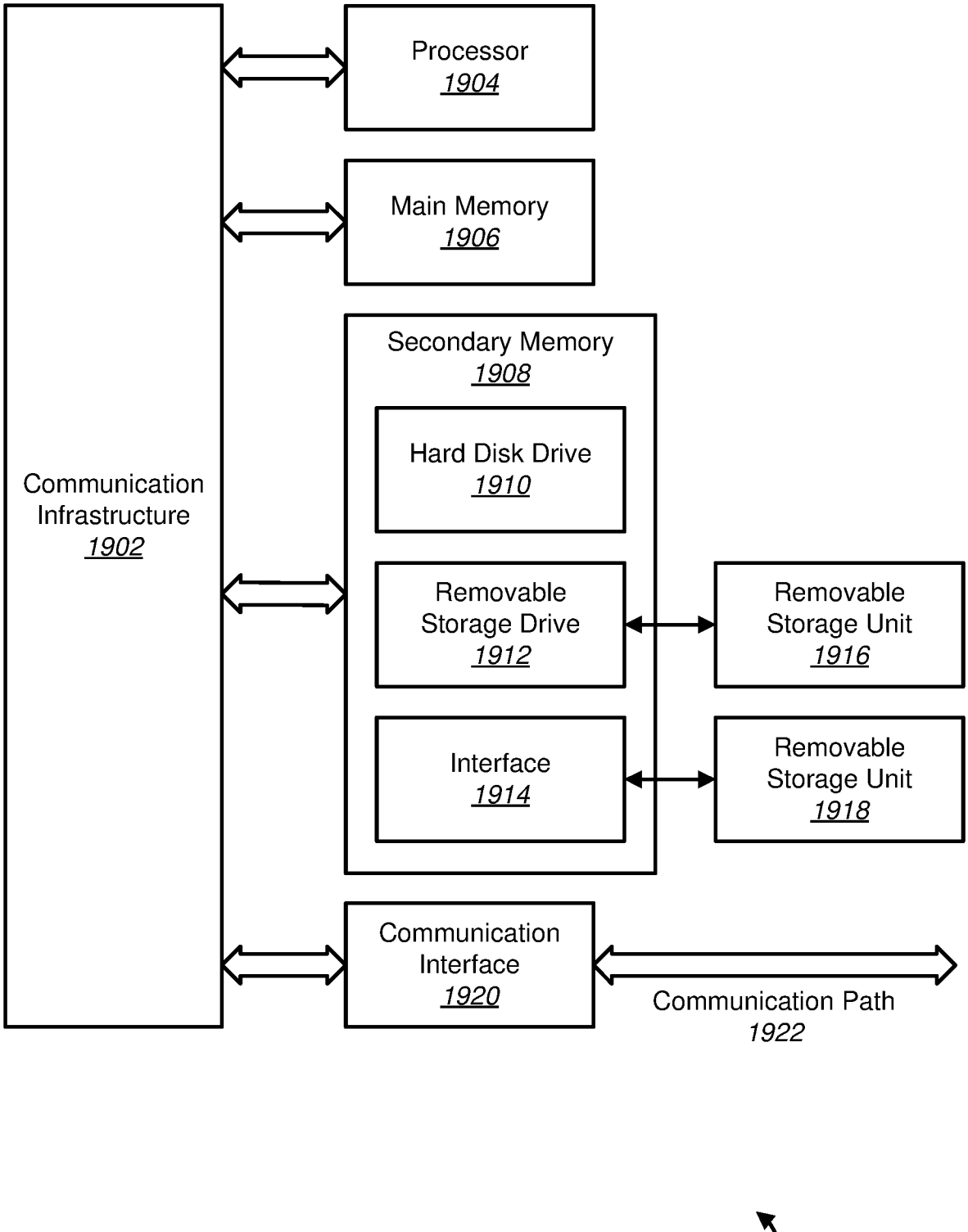

Processor
1904
Main Memory
1906
Secondary Memory
1908
Hard Disk Drive
1910
Removable
Storage Drive
1912
Removable
Storage Unit
1916
Interface
1914
Removable
Storage Unit
1918
Communication
Infrastructure
1902
Communication
Interface
1920
Communication Path
1922
1900

FIG. 19

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2021/034226

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N19/11 H04N19/154 H04N19/176 H04N19/463 H04N19/593
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04N

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2017/353719 A1 (LIU SHAN [US] ET AL) 7 December 2017 (2017-12-07) paragraphs [0022], [0064] - [0067], [0081] - [0085], [0093]	1-73
X	WO 2020/056779 A1 (HANGZHOU HIKVISION DIGITAL TEC [CN]) 26 March 2020 (2020-03-26) paragraphs [0134] - [0136], [0142] - [0144]	1-73
A	US 2017/374369 A1 (CHUANG TZU-DER [TW] ET AL) 28 December 2017 (2017-12-28) paragraphs [0010], [0015], [0023], [0065], [0085]	1-73

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 July 2021	09/08/2021

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

1

Form PCT/ISA/210 (second sheet) (April 2005)

International application No
PCT/US2021/034226

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	VAN DER AUWERA (QUALCOMM) G ET AL: "Description of Core Experiment 3 (CE3): Intra Prediction and Mode Coding", 12. JVET MEETING; 20181003 - 20181012; MACAO; (THE JOINT VIDEO EXPLORATION TEAM OF ISO/IEC JTC1/SC29/WG11 AND ITU-T SG.16), , no. JVET-L1023 4 November 2018 (2018-11-04), XP030198606, Retrieved from the Internet: URL:http://phenix.int-evry.fr/jvet/doc_end_user/documents/12_Macao/wg11/JVET-L1023-v2.zip JVET-L1023_v2.docx [retrieved on 2018-11-04] section 3.3.3.1 -----	1-73

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2021/034226

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2017353719 A1	07-12-2017	TW 201811031 A	16-03-2018
		US 2017353719 A1	07-12-2017
WO 2020056779 A1	26-03-2020	NONE	
US 2017374369 A1	28-12-2017	TW 201804799 A	01-02-2018
		US 2017374369 A1	28-12-2017