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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
03 August 2023 (03.08.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/146688 A1

(51) International Patent Classification:
H04N 19/105 (2014.01) *H04N 19/117* (2014.01)
H04N 19/157 (2014.01) *H04N 19/44* (2014.01)
H04N 19/186 (2014.01) *H04N 19/593* (2014.01)
H04N 19/11 (2014.01)

(21) International Application Number:
PCT/US2022/079154

(22) International Filing Date:
02 November 2022 (02.11.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/305,168 31 January 2022 (31.01.2022) US
17/976,470 28 October 2022 (28.10.2022) US

(71) Applicant: **TENCENT AMERICA LLC** [US/US]; 2747 Park Boulevard, Palo Alto, California 94306 (US).

(72) Inventors: **XU, Xiaozhong**; c/o TENCENT AMERICA LLC, 2747 Park Boulevard, Palo Alto, California 94306 (US). **LIU, Shan**; c/o TENCENT AMERICA LLC, 2747 Park Boulevard, Palo Alto, California 94306 (US).

(74) Agent: **MA, Johnny**; ArentFox Schiff LLP, 1717 K Street NW, Washington, District of Columbia 20006 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG,

(54) Title: CROSS-COMPONENT PLANAR PREDICTION IN IMAGE AND VIDEO COMPRESSION

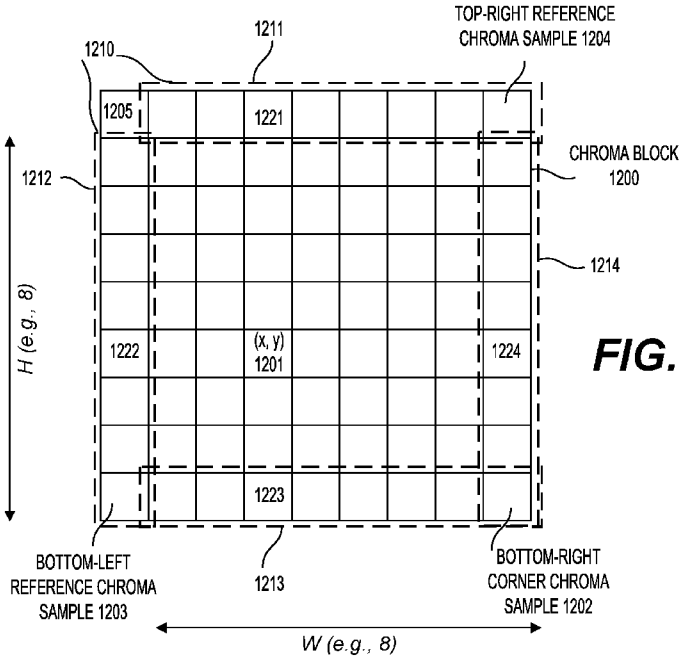


FIG. 12

(57) Abstract: Aspects of the disclosure provide a method and an apparatus including processing circuitry that decode prediction information of a chroma block to be reconstructed in a current picture. The prediction information indicates that the chroma block is predicted based at least on a luma block in the current picture using a planar mode. The luma block is reconstructed and collocated with the chroma block. The processing circuitry predicts one or more first chroma samples in the chroma block based on one or more luma samples in the luma block. The processing circuitry predicts a second chroma sample in the chroma block based on the one or more first chroma samples and at least one reference chroma sample of the chroma block. The second chroma sample is different from the one or more first chroma samples. The at least one reference chroma sample is adjacent to the chroma block.

WO 2023/146688 A1

[Continued on next page]

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, CV, 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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

CROSS-COMPONENT PLANAR PREDICTION IN IMAGE AND VIDEO COMPRESSION

INCORPORATION BY REFERENCE

[0001] The present application claims the benefit of priority to U.S. Patent Application No. 17/976,470, “CROSS-COMPONENT PLANAR PREDICTION IN IMAGE AND VIDEO COMPRESSION” filed on October 28, 2022, which claims the benefit of priority to U.S. Provisional Application No. 63/305,168, "Cross-Component Planar Prediction in Image and Video Compression" filed on January 31, 2022. The disclosures of the prior applications are hereby incorporated by reference in their entirety.

TECHNICAL FIELD

[0002] The present disclosure describes embodiments generally related to video coding.

BACKGROUND

[0003] The background description provided herein is for the purpose of generally presenting the context of the disclosure. Work of the presently named inventors, to the extent the work is described in this background section, as well as aspects of the description that may not otherwise qualify as prior art at the time of filing, are neither expressly nor impliedly admitted as prior art against the present disclosure.

[0004] Uncompressed digital images and/or video can include a series of pictures, each picture having a spatial dimension of, for example, 1920 x 1080 luminance samples and associated chrominance samples. The series of pictures can have a fixed or variable picture rate (informally also known as frame rate), of, for example 60 pictures per second or 60 Hz. Uncompressed image and/or video has specific bitrate requirements. For example, 1080p60 4:2:0 video at 8 bit per sample (1920x1080 luminance sample resolution at 60 Hz frame rate) requires close to 1.5 Gbit/s bandwidth. An hour of such video requires more than 600 GBytes of storage space.

[0005] One purpose of image and/or video coding and decoding can be the reduction of redundancy in the input image and/or video signal, through compression. Compression can help reduce the aforementioned bandwidth and/or storage space requirements, in some cases by two orders of magnitude or more. Although the descriptions herein use video encoding/decoding as illustrative examples, the same techniques can be applied to image encoding/decoding in similar fashion without departing from the spirit of the present disclosure. Both lossless compression and lossy compression, as well as a combination thereof can be employed. Lossless compression

refers to techniques where an exact copy of the original signal can be reconstructed from the compressed original signal. When using lossy compression, the reconstructed signal may not be identical to the original signal, but the distortion between original and reconstructed signals is small enough to make the reconstructed signal useful for the intended application. In the case of video, lossy compression is widely employed. The amount of distortion tolerated depends on the application; for example, users of certain consumer streaming applications may tolerate higher distortion than users of television distribution applications. The compression ratio achievable can reflect that: higher allowable/tolerable distortion can yield higher compression ratios.

[0006] A video encoder and decoder can utilize techniques from several broad categories, including, for example, motion compensation, transform processing, quantization, and entropy coding.

[0007] Video codec technologies can include techniques known as intra coding. In intra coding, sample values are represented without reference to samples or other data from previously reconstructed reference pictures. In some video codecs, the picture is spatially subdivided into blocks of samples. When all blocks of samples are coded in intra mode, that picture can be an intra picture. Intra pictures and their derivations such as independent decoder refresh pictures, can be used to reset the decoder state and can, therefore, be used as the first picture in a coded video bitstream and a video session, or as a still image. The samples of an intra block can be exposed to a transform, and the transform coefficients can be quantized before entropy coding. Intra prediction can be a technique that minimizes sample values in the pre-transform domain. In some cases, the smaller the DC value after a transform is, and the smaller the AC coefficients are, the fewer the bits that are required at a given quantization step size to represent the block after entropy coding.

[0008] Traditional intra coding used in, for example, MPEG-2 generation coding technologies, does not use intra prediction. However, some newer video compression technologies include techniques that attempt to perform prediction based on, for example, surrounding sample data and/or metadata obtained during the encoding and/or decoding of blocks of data. Such techniques are henceforth called “intra prediction” techniques. Note that in at least some cases, intra prediction is using reference data only from the current picture under reconstruction and not from reference pictures.

[0009] There can be many different forms of intra prediction. When more than one of such techniques can be used in a given video coding technology, a specific technique in use can

be coded as a specific intra prediction mode that uses the specific technique. In certain cases, intra prediction modes can have submodes and/or parameters, where the submodes and/or parameters can be coded individually or included in a mode codeword, which defines the prediction mode being used. Which codeword to use for a given mode, submode, and/or parameter combination can have an impact in the coding efficiency gain through intra prediction, and so can the entropy coding technology used to translate the codewords into a bitstream.

[0010] A certain mode of intra prediction was introduced with H.264, refined in H.265, and further refined in newer coding technologies such as joint exploration model (JEM), versatile video coding (VVC), and benchmark set (BMS). A predictor block can be formed using neighboring sample values of already available samples. Sample values of neighboring samples are copied into the predictor block according to a direction. A reference to the direction in use can be coded in the bitstream or may itself be predicted.

[0011] Referring to FIG. 1A, depicted in the lower right is a subset of nine predictor directions known from the 33 possible predictor directions (corresponding to the 33 angular modes of the 35 intra modes) defined in H.265. The point where the arrows converge (101) represents the sample being predicted. The arrows represent the direction from which the sample is being predicted. For example, arrow (102) indicates that sample (101) is predicted from a sample or samples to the upper right, at a 45 degree angle from the horizontal. Similarly, arrow (103) indicates that sample (101) is predicted from a sample or samples to the lower left of sample (101), in a 22.5 degree angle from the horizontal.

[0012] Still referring to FIG. 1A, on the top left there is depicted a square block (104) of 4 x 4 samples (indicated by a dashed, boldface line). The square block (104) includes 16 samples, each labelled with an "S", its position in the Y dimension (e.g., row index) and its position in the X dimension (e.g., column index). For example, sample S21 is the second sample in the Y dimension (from the top) and the first (from the left) sample in the X dimension. Similarly, sample S44 is the fourth sample in block (104) in both the Y and X dimensions. As the block is 4 x 4 samples in size, S44 is at the bottom right. Further shown are reference samples that follow a similar numbering scheme. A reference sample is labelled with an R, its Y position (e.g., row index) and X position (column index) relative to block (104). In both H.264 and H.265, prediction samples neighbor the block under reconstruction; therefore, no negative values need to be used.

[0013] Intra picture prediction can work by copying reference sample values from the neighboring samples indicated by the signaled prediction direction. For example, assume the coded video bitstream includes signaling that, for this block, indicates a prediction direction consistent with arrow (102)—that is, samples are predicted from samples to the upper right, at a 45 degree angle from the horizontal. In that case, samples S41, S32, S23, and S14 are predicted from the same reference sample R05. Sample S44 is then predicted from reference sample R08.

[0014] In certain cases, the values of multiple reference samples may be combined, for example through interpolation, in order to calculate a reference sample; especially when the directions are not evenly divisible by 45 degrees.

[0015] The number of possible directions has increased as video coding technology has developed. In H.264 (year 2003), nine different direction could be represented. That increased to 33 in H.265 (year 2013). Currently, JEM/VVC/BMS can support up to 65 directions. Experiments have been conducted to identify the most likely directions, and certain techniques in the entropy coding are used to represent those likely directions in a small number of bits, accepting a certain penalty for less likely directions. Further, the directions themselves can sometimes be predicted from neighboring directions used in neighboring, already decoded, blocks.

[0016] FIG. 1B shows a schematic (110) that depicts 65 intra prediction directions according to JEM to illustrate the increasing number of prediction directions over time.

[0017] The mapping of intra prediction direction bits that represent the direction in the coded video bitstream can be different from video coding technology to video coding technology. Such mapping can range, for example, from simple direct mappings, to codewords, to complex adaptive schemes involving most probable modes, and similar techniques. In most cases, however, there can be certain directions that are statistically less likely to occur in video content than certain other directions. As the goal of video compression is the reduction of redundancy, those less likely directions will, in a well working video coding technology, be represented by a larger number of bits than more likely directions.

[0018] Image and/or video coding and decoding can be performed using inter-picture prediction with motion compensation. Motion compensation can be a lossy compression technique and can relate to techniques where a block of sample data from a previously reconstructed picture or part thereof (reference picture), after being spatially shifted in a direction indicated by a motion vector (MV henceforth), is used for the prediction of a newly

reconstructed picture or picture part. In some cases, the reference picture can be the same as the picture currently under reconstruction. MVs can have two dimensions X and Y, or three dimensions, the third being an indication of the reference picture in use (the latter, indirectly, can be a time dimension).

[0019] In some video compression techniques, an MV applicable to a certain area of sample data can be predicted from other MVs, for example from those related to another area of sample data spatially adjacent to the area under reconstruction, and preceding that MV in decoding order. Doing so can substantially reduce the amount of data required for coding the MV, thereby removing redundancy and increasing compression. MV prediction can work effectively, for example, because when coding an input video signal derived from a camera (known as natural video) there is a statistical likelihood that areas larger than the area to which a single MV is applicable move in a similar direction and, therefore, can in some cases be predicted using a similar motion vector derived from MVs of neighboring area. That results in the MV found for a given area to be similar or the same as the MV predicted from the surrounding MVs, and that in turn can be represented, after entropy coding, in a smaller number of bits than what would be used if coding the MV directly. In some cases, MV prediction can be an example of lossless compression of a signal (namely: the MVs) derived from the original signal (namely: the sample stream). In other cases, MV prediction itself can be lossy, for example because of rounding errors when calculating a predictor from several surrounding MVs.

[0020] Various MV prediction mechanisms are described in H.265/HEVC (ITU-T Rec. H.265, "High Efficiency Video Coding", December 2016). Out of the many MV prediction mechanisms that H.265 offers, described with reference to FIG. 2 is a technique henceforth referred to as "spatial merge".

[0021] Referring to FIG. 2, a current block (201) comprises samples that have been found by the encoder during the motion search process to be predictable from a previous block of the same size that has been spatially shifted. Instead of coding that MV directly, the MV can be derived from metadata associated with one or more reference pictures, for example from the most recent (in decoding order) reference picture, using the MV associated with either one of five surrounding samples, denoted A0, A1, and B0, B1, B2 (202 through 206, respectively). In H.265, the MV prediction can use predictors from the same reference picture that the neighboring block is using.

SUMMARY

[0022] Aspects of the disclosure provide methods and apparatuses for video encoding/decoding. In some examples, an apparatus for video decoding includes processing circuitry. The processing circuitry decode prediction information of a chroma block to be reconstructed in a current picture. The prediction information indicates that the chroma block is predicted based at least on a luma block in the current picture using a planar mode. The luma block is already reconstructed and collocated with the chroma block. The processing circuitry predicts one or more first chroma samples in the chroma block based on one or more luma samples in the luma block. The processing circuitry predicts a second chroma sample in the chroma block based on the one or more first chroma samples and at least one reference chroma sample of the chroma block. The second chroma sample is different from the one or more first chroma samples. The at least one reference chroma sample is adjacent to the chroma block.

[0023] In an embodiment, the one or more first chroma samples is a bottom-right corner chroma sample in the chroma block. The processing circuitry predicts, based on the bottom-right corner chroma sample and the at least one reference chroma sample, at least one of (i) chroma samples in a bottom row of the chroma block or (ii) chroma samples in a right most column of the chroma block. In response to the chroma samples in the bottom row of the chroma block being predicted, the processing circuitry generates a vertical predictor of the second chroma sample based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row. The second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample are in a same column in the chroma block. In response to the chroma samples in the right most column of the chroma block being predicted, the processing circuitry generates a horizontal predictor of the second chroma sample based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column. The second chroma sample, the left reference chroma sample, and the predicted right chroma sample are in a same row of the chroma block.

[0024] In an example, the processing circuitry predicts the chroma samples in the bottom row of the chroma block and the chroma samples in the right most column of the chroma block. The processing circuitry predicts the second chroma sample based on the vertical predictor and the horizontal predictor.

[0025] In an example, a width of the chroma block is larger than a height of the chroma block or the height of the chroma block is less than a threshold. The processing circuitry predicts the chroma samples in the bottom row of the chroma block and predicts the second chroma sample based on the vertical predictor.

[0026] In an example, a width of the chroma block is less than a height of the chroma block or a threshold. The processing circuitry predicts the chroma samples in the right most column of the chroma block and predicts the second chroma sample based on the horizontal predictor.

[0027] In an example, the processing circuitry predicts the bottom-right corner chroma sample in the chroma block based on at least one of a bottom-right corner luma sample in the luma block or luma samples in the luma block that are collocated with the bottom-right corner chroma sample.

[0028] In an example, the one or more first chroma samples include (i) chroma samples in a bottom row of the chroma block and (ii) chroma samples in a right most column of the chroma block. The processing circuitry generates a vertical predictor of the second chroma sample based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row. The second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample are in a same column in the chroma block. The processing circuitry generates a horizontal predictor of the second chroma sample based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column. The second chroma sample, the left reference chroma sample, and the predicted right chroma sample are in a same row in the chroma block. The processing circuitry predicts the second chroma sample based on the vertical predictor and the horizontal predictor.

[0029] In an example, a width of the chroma block is larger than a height of the chroma block or the height of the chroma block is less than a threshold. The one or more first chroma samples include chroma samples in a bottom row of the chroma block. The processing circuitry generates a vertical predictor of the second chroma sample based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row. The second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample are in a same column in the chroma

block. The processing circuitry predicts the second chroma sample based on the vertical predictor.

[0030] In an example, a width of the chroma block is less than a height of the chroma block or a threshold. The one or more first chroma samples include chroma samples in a right most column of the chroma block. The processing circuitry generates a horizontal predictor of the second chroma sample based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column. The second chroma sample, the left reference chroma sample, and the predicted right chroma sample are in a same row in the chroma block. The processing circuitry predicts the second chroma sample based on the horizontal predictor.

[0031] In an example, the processing circuitry predicts the one or more first chroma samples in the chroma block based on the one or more luma samples in the luma block that are collocated with the one or more first chroma samples using a cross-component linear model prediction (CCLM) mode.

[0032] In an example, the at least one reference chroma sample includes at least one of a top reference chroma sample that is a top neighbor of the chroma block or a left reference chroma sample that is a left neighbor of the chroma block.

[0033] Aspects of the disclosure also provide a non-transitory computer-readable medium storing instructions which when executed by a computer for video decoding cause the computer to perform the method for video decoding.

BRIEF DESCRIPTION OF THE DRAWINGS

[0034] Further features, the nature, and various advantages of the disclosed subject matter will be more apparent from the following detailed description and the accompanying drawings in which:

[0035] FIG. 1A is a schematic illustration of an exemplary subset of intra prediction modes.

[0036] FIG. 1B is an illustration of exemplary intra prediction directions.

[0037] FIG. 2 shows an example of a current block (201) and surrounding samples.

[0038] FIG. 3 is a schematic illustration of an exemplary block diagram of a communication system (300).

[0039] FIG. 4 is a schematic illustration of an exemplary block diagram of a communication system (400).

[0040] FIG. 5 is a schematic illustration of an exemplary block diagram of a decoder.

[0041] FIG. 6 is a schematic illustration of an exemplary block diagram of an encoder.

[0042] FIG. 7 shows a block diagram of an exemplary encoder.

[0043] FIG. 8 shows a block diagram of an exemplary decoder.

[0044] FIG. 9 shows exemplary reference samples of a current block.

[0045] FIG. 10 shows examples of reconstructed neighboring luma samples and reconstructed neighboring chroma samples used in a cross-component linear model (CCLM) mode.

[0046] FIG. 11 shows examples of reconstructed neighboring luma samples and reconstructed neighboring chroma samples used in the CCLM mode.

[0047] FIG. 12 shows exemplary samples or sample positions in a chroma block.

[0048] FIG. 13 shows a flow chart outlining an encoding process according to some embodiments of the disclosure.

[0049] FIG. 14 shows a flow chart outlining a decoding process according to some embodiments of the disclosure.

[0050] FIG. 15 is a schematic illustration of a computer system in accordance with an embodiment.

DETAILED DESCRIPTION OF EMBODIMENTS

[0051] FIG. 3 illustrates an exemplary block diagram of a communication system (300). The communication system (300) includes a plurality of terminal devices that can communicate with each other, via, for example, a network (350). For example, the communication system (300) includes a first pair of terminal devices (310) and (320) interconnected via the network (350). In the FIG. 3 example, the first pair of terminal devices (310) and (320) performs unidirectional transmission of data. For example, the terminal device (310) may code video data (e.g., a stream of video pictures that are captured by the terminal device (310)) for transmission to the other terminal device (320) via the network (350). The encoded video data can be transmitted in the form of one or more coded video bitstreams. The terminal device (320) may receive the coded video data from the network (350), decode the coded video data to recover the video pictures and display video pictures according to the recovered video data. Unidirectional data transmission may be common in media serving applications and the like.

[0052] In another example, the communication system (300) includes a second pair of terminal devices (330) and (340) that perform bidirectional transmission of coded video data, for

example, during videoconferencing. For bidirectional transmission of data, in an example, each terminal device of the terminal devices (330) and (340) may code video data (e.g., a stream of video pictures that are captured by the terminal device) for transmission to the other terminal device of the terminal devices (330) and (340) via the network (350). Each terminal device of the terminal devices (330) and (340) also may receive the coded video data transmitted by the other terminal device of the terminal devices (330) and (340), and may decode the coded video data to recover the video pictures and may display video pictures at an accessible display device according to the recovered video data.

[0053] In the example of FIG. 3, the terminal devices (310), (320), (330) and (340) are respectively illustrated as servers, personal computers and smart phones but the principles of the present disclosure may be not so limited. Embodiments of the present disclosure find application with laptop computers, tablet computers, media players, and/or dedicated video conferencing equipment. The network (350) represents any number of networks that convey coded video data among the terminal devices (310), (320), (330) and (340), including for example wireline (wired) and/or wireless communication networks. The communication network (350) may exchange data in circuit-switched and/or packet-switched channels. Representative networks include telecommunications networks, local area networks, wide area networks and/or the Internet. For the purposes of the present discussion, the architecture and topology of the network (350) may be immaterial to the operation of the present disclosure unless explained herein below.

[0054] FIG. 4 illustrates, as an example of an application for the disclosed subject matter, a video encoder and a video decoder in a streaming environment. The disclosed subject matter can be equally applicable to other video enabled applications, including, for example, video conferencing, digital TV, streaming services, storing of compressed video on digital media including CD, DVD, memory stick and the like, and so on.

[0055] A streaming system may include a capture subsystem (413), that can include a video source (401), for example a digital camera, creating for example a stream of video pictures (402) that are uncompressed. In an example, the stream of video pictures (402) includes samples that are taken by the digital camera. The stream of video pictures (402), depicted as a bold line to emphasize a high data volume when compared to encoded video data (404) (or coded video bitstreams), can be processed by an electronic device (420) that includes a video encoder (403) coupled to the video source (401). The video encoder (403) can include hardware, software, or a combination thereof to enable or implement aspects of the disclosed subject matter as described

in more detail below. The encoded video data (404) (or encoded video bitstream), depicted as a thin line to emphasize the lower data volume when compared to the stream of video pictures (402), can be stored on a streaming server (405) for future use. One or more streaming client subsystems, such as client subsystems (406) and (408) in FIG. 4 can access the streaming server (405) to retrieve copies (407) and (409) of the encoded video data (404). A client subsystem (406) can include a video decoder (410), for example, in an electronic device (430). The video decoder (410) decodes the incoming copy (407) of the encoded video data and creates an outgoing stream of video pictures (411) that can be rendered on a display (412) (e.g., display screen) or other rendering device (not depicted). In some streaming systems, the encoded video data (404), (407), and (409) (e.g., video bitstreams) can be encoded according to certain video coding/compression standards. Examples of those standards include ITU-T Recommendation H.265. In an example, a video coding standard under development is informally known as Versatile Video Coding (VVC). The disclosed subject matter may be used in the context of VVC.

[0056] It is noted that the electronic devices (420) and (430) can include other components (not shown). For example, the electronic device (420) can include a video decoder (not shown) and the electronic device (430) can include a video encoder (not shown) as well.

[0057] FIG. 5 shows an exemplary block diagram of a video decoder (510). The video decoder (510) can be included in an electronic device (530). The electronic device (530) can include a receiver (531) (e.g., receiving circuitry). The video decoder (510) can be used in the place of the video decoder (410) in the FIG. 4 example.

[0058] The receiver (531) may receive one or more coded video sequences to be decoded by the video decoder (510). In an embodiment, one coded video sequence is received at a time, where the decoding of each coded video sequence is independent from the decoding of other coded video sequences. The coded video sequence may be received from a channel (501), which may be a hardware/software link to a storage device which stores the encoded video data. The receiver (531) may receive the encoded video data with other data, for example, coded audio data and/or ancillary data streams, that may be forwarded to their respective using entities (not depicted). The receiver (531) may separate the coded video sequence from the other data. To combat network jitter, a buffer memory (515) may be coupled in between the receiver (531) and an entropy decoder / parser (520) ("parser (520)" henceforth). In certain applications, the buffer memory (515) is part of the video decoder (510). In others, it can be outside of the video decoder

(510) (not depicted). In still others, there can be a buffer memory (not depicted) outside of the video decoder (510), for example to combat network jitter, and in addition another buffer memory (515) inside the video decoder (510), for example to handle playout timing. When the receiver (531) is receiving data from a store/forward device of sufficient bandwidth and controllability, or from an isosynchronous network, the buffer memory (515) may not be needed, or can be small. For use on best effort packet networks such as the Internet, the buffer memory (515) may be required, can be comparatively large and can be advantageously of adaptive size, and may at least partially be implemented in an operating system or similar elements (not depicted) outside of the video decoder (510).

[0059] The video decoder (510) may include the parser (520) to reconstruct symbols (521) from the coded video sequence. Categories of those symbols include information used to manage operation of the video decoder (510), and potentially information to control a rendering device such as a render device (512) (e.g., a display screen) that is not an integral part of the electronic device (530) but can be coupled to the electronic device (530), as shown in FIG. 5. The control information for the rendering device(s) may be in the form of Supplemental Enhancement Information (SEI) messages or Video Usability Information (VUI) parameter set fragments (not depicted). The parser (520) may parse / entropy-decode the coded video sequence that is received. The coding of the coded video sequence can be in accordance with a video coding technology or standard, and can follow various principles, including variable length coding, Huffman coding, arithmetic coding with or without context sensitivity, and so forth. The parser (520) may extract from the coded video sequence, a set of subgroup parameters for at least one of the subgroups of pixels in the video decoder, based upon at least one parameter corresponding to the group. Subgroups can include Groups of Pictures (GOPs), pictures, tiles, slices, macroblocks, Coding Units (CUs), blocks, Transform Units (TUs), Prediction Units (PUs) and so forth. The parser (520) may also extract from the coded video sequence information such as transform coefficients, quantizer parameter values, motion vectors, and so forth.

[0060] The parser (520) may perform an entropy decoding / parsing operation on the video sequence received from the buffer memory (515), so as to create symbols (521).

[0061] Reconstruction of the symbols (521) can involve multiple different units depending on the type of the coded video picture or parts thereof (such as: inter and intra picture, inter and intra block), and other factors. Which units are involved, and how, can be controlled by subgroup control information parsed from the coded video sequence by the parser (520). The

flow of such subgroup control information between the parser (520) and the multiple units below is not depicted for clarity.

[0062] Beyond the functional blocks already mentioned, the video decoder (510) can be conceptually subdivided into a number of functional units as described below. In a practical implementation operating under commercial constraints, many of these units interact closely with each other and can, at least partly, be integrated into each other. However, for the purpose of describing the disclosed subject matter, the conceptual subdivision into the functional units below is appropriate.

[0063] A first unit is the scaler / inverse transform unit (551). The scaler / inverse transform unit (551) receives a quantized transform coefficient as well as control information, including which transform to use, block size, quantization factor, quantization scaling matrices, etc. as symbol(s) (521) from the parser (520). The scaler / inverse transform unit (551) can output blocks comprising sample values, that can be input into aggregator (555).

[0064] In some cases, the output samples of the scaler / inverse transform unit (551) can pertain to an intra coded block. The intra coded block is a block that is not using predictive information from previously reconstructed pictures, but can use predictive information from previously reconstructed parts of the current picture. Such predictive information can be provided by an intra picture prediction unit (552). In some cases, the intra picture prediction unit (552) generates a block of the same size and shape of the block under reconstruction, using surrounding already reconstructed information fetched from the current picture buffer (558). The current picture buffer (558) buffers, for example, partly reconstructed current picture and/or fully reconstructed current picture. The aggregator (555), in some cases, adds, on a per sample basis, the prediction information the intra prediction unit (552) has generated to the output sample information as provided by the scaler / inverse transform unit (551).

[0065] In other cases, the output samples of the scaler / inverse transform unit (551) can pertain to an inter coded, and potentially motion compensated, block. In such a case, a motion compensation prediction unit (553) can access reference picture memory (557) to fetch samples used for prediction. After motion compensating the fetched samples in accordance with the symbols (521) pertaining to the block, these samples can be added by the aggregator (555) to the output of the scaler / inverse transform unit (551) (in this case called the residual samples or residual signal) so as to generate output sample information. The addresses within the reference picture memory (557) from where the motion compensation prediction unit (553) fetches

prediction samples can be controlled by motion vectors, available to the motion compensation prediction unit (553) in the form of symbols (521) that can have, for example X, Y, and reference picture components. Motion compensation also can include interpolation of sample values as fetched from the reference picture memory (557) when sub-sample exact motion vectors are in use, motion vector prediction mechanisms, and so forth.

[0066] The output samples of the aggregator (555) can be subject to various loop filtering techniques in the loop filter unit (556). Video compression technologies can include in-loop filter technologies that are controlled by parameters included in the coded video sequence (also referred to as coded video bitstream) and made available to the loop filter unit (556) as symbols (521) from the parser (520). Video compression can also be responsive to meta-information obtained during the decoding of previous (in decoding order) parts of the coded picture or coded video sequence, as well as responsive to previously reconstructed and loop-filtered sample values.

[0067] The output of the loop filter unit (556) can be a sample stream that can be output to the render device (512) as well as stored in the reference picture memory (557) for use in future inter-picture prediction.

[0068] Certain coded pictures, once fully reconstructed, can be used as reference pictures for future prediction. For example, once a coded picture corresponding to a current picture is fully reconstructed and the coded picture has been identified as a reference picture (by, for example, the parser (520)), the current picture buffer (558) can become a part of the reference picture memory (557), and a fresh current picture buffer can be reallocated before commencing the reconstruction of the following coded picture.

[0069] The video decoder (510) may perform decoding operations according to a predetermined video compression technology or a standard, such as ITU-T Rec. H.265. The coded video sequence may conform to a syntax specified by the video compression technology or standard being used, in the sense that the coded video sequence adheres to both the syntax of the video compression technology or standard and the profiles as documented in the video compression technology or standard. Specifically, a profile can select certain tools as the only tools available for use under that profile from all the tools available in the video compression technology or standard. Also necessary for compliance can be that the complexity of the coded video sequence is within bounds as defined by the level of the video compression technology or standard. In some cases, levels restrict the maximum picture size, maximum frame rate,

maximum reconstruction sample rate (measured in, for example megasamples per second), maximum reference picture size, and so on. Limits set by levels can, in some cases, be further restricted through Hypothetical Reference Decoder (HRD) specifications and metadata for HRD buffer management signaled in the coded video sequence.

[0070] In an embodiment, the receiver (531) may receive additional (redundant) data with the encoded video. The additional data may be included as part of the coded video sequence(s). The additional data may be used by the video decoder (510) to properly decode the data and/or to more accurately reconstruct the original video data. Additional data can be in the form of, for example, temporal, spatial, or signal noise ratio (SNR) enhancement layers, redundant slices, redundant pictures, forward error correction codes, and so on.

[0071] FIG. 6 shows an exemplary block diagram of a video encoder (603). The video encoder (603) is included in an electronic device (620). The electronic device (620) includes a transmitter (640) (e.g., transmitting circuitry). The video encoder (603) can be used in the place of the video encoder (403) in the FIG. 4 example.

[0072] The video encoder (603) may receive video samples from a video source (601) (that is not part of the electronic device (620) in the FIG. 6 example) that may capture video image(s) to be coded by the video encoder (603). In another example, the video source (601) is a part of the electronic device (620).

[0073] The video source (601) may provide the source video sequence to be coded by the video encoder (603) in the form of a digital video sample stream that can be of any suitable bit depth (for example: 8 bit, 10 bit, 12 bit, ...), any colorspace (for example, BT.601 Y CrCB, RGB, ...), and any suitable sampling structure (for example Y CrCb 4:2:0, Y CrCb 4:4:4). In a media serving system, the video source (601) may be a storage device storing previously prepared video. In a videoconferencing system, the video source (601) may be a camera that captures local image information as a video sequence. Video data may be provided as a plurality of individual pictures that impart motion when viewed in sequence. The pictures themselves may be organized as a spatial array of pixels, wherein each pixel can comprise one or more samples depending on the sampling structure, color space, etc. in use. A person skilled in the art can readily understand the relationship between pixels and samples. The description below focuses on samples.

[0074] According to an embodiment, the video encoder (603) may code and compress the pictures of the source video sequence into a coded video sequence (643) in real time or under

any other time constraints as required. Enforcing appropriate coding speed is one function of a controller (650). In some embodiments, the controller (650) controls other functional units as described below and is functionally coupled to the other functional units. The coupling is not depicted for clarity. Parameters set by the controller (650) can include rate control related parameters (picture skip, quantizer, lambda value of rate-distortion optimization techniques, ...), picture size, group of pictures (GOP) layout, maximum motion vector search range, and so forth. The controller (650) can be configured to have other suitable functions that pertain to the video encoder (603) optimized for a certain system design.

[0075] In some embodiments, the video encoder (603) is configured to operate in a coding loop. As an oversimplified description, in an example, the coding loop can include a source coder (630) (e.g., responsible for creating symbols, such as a symbol stream, based on an input picture to be coded, and a reference picture(s)), and a (local) decoder (633) embedded in the video encoder (603). The decoder (633) reconstructs the symbols to create the sample data in a similar manner as a (remote) decoder also would create. The reconstructed sample stream (sample data) is input to the reference picture memory (634). As the decoding of a symbol stream leads to bit-exact results independent of decoder location (local or remote), the content in the reference picture memory (634) is also bit exact between the local encoder and remote encoder. In other words, the prediction part of an encoder "sees" as reference picture samples exactly the same sample values as a decoder would "see" when using prediction during decoding. This fundamental principle of reference picture synchronicity (and resulting drift, if synchronicity cannot be maintained, for example because of channel errors) is used in some related arts as well.

[0076] The operation of the "local" decoder (633) can be the same as of a "remote" decoder, such as the video decoder (510), which has already been described in detail above in conjunction with FIG. 5. Briefly referring also to FIG. 5, however, as symbols are available and encoding/decoding of symbols to a coded video sequence by an entropy coder (645) and the parser (520) can be lossless, the entropy decoding parts of the video decoder (510), including the buffer memory (515), and parser (520) may not be fully implemented in the local decoder (633).

[0077] In an embodiment, a decoder technology except the parsing/entropy decoding that is present in a decoder is present, in an identical or a substantially identical functional form, in a corresponding encoder. Accordingly, the disclosed subject matter focuses on decoder operation. The description of encoder technologies can be abbreviated as they are the inverse of the

comprehensively described decoder technologies. In certain areas a more detail description is provided below.

[0078] During operation, in some examples, the source coder (630) may perform motion compensated predictive coding, which codes an input picture predictively with reference to one or more previously coded picture from the video sequence that were designated as "reference pictures." In this manner, the coding engine (632) codes differences between pixel blocks of an input picture and pixel blocks of reference picture(s) that may be selected as prediction reference(s) to the input picture.

[0079] The local video decoder (633) may decode coded video data of pictures that may be designated as reference pictures, based on symbols created by the source coder (630). Operations of the coding engine (632) may advantageously be lossy processes. When the coded video data may be decoded at a video decoder (not shown in FIG. 6), the reconstructed video sequence typically may be a replica of the source video sequence with some errors. The local video decoder (633) replicates decoding processes that may be performed by the video decoder on reference pictures and may cause reconstructed reference pictures to be stored in the reference picture memory (634). In this manner, the video encoder (603) may store copies of reconstructed reference pictures locally that have common content as the reconstructed reference pictures that will be obtained by a far-end video decoder (absent transmission errors).

[0080] The predictor (635) may perform prediction searches for the coding engine (632). That is, for a new picture to be coded, the predictor (635) may search the reference picture memory (634) for sample data (as candidate reference pixel blocks) or certain metadata such as reference picture motion vectors, block shapes, and so on, that may serve as an appropriate prediction reference for the new pictures. The predictor (635) may operate on a sample block-by-pixel block basis to find appropriate prediction references. In some cases, as determined by search results obtained by the predictor (635), an input picture may have prediction references drawn from multiple reference pictures stored in the reference picture memory (634).

[0081] The controller (650) may manage coding operations of the source coder (630), including, for example, setting of parameters and subgroup parameters used for encoding the video data.

[0082] Output of all aforementioned functional units may be subjected to entropy coding in the entropy coder (645). The entropy coder (645) translates the symbols as generated by the various functional units into a coded video sequence, by applying lossless compression to the

symbols according to technologies such as Huffman coding, variable length coding, arithmetic coding, and so forth.

[0083] The transmitter (640) may buffer the coded video sequence(s) as created by the entropy coder (645) to prepare for transmission via a communication channel (660), which may be a hardware/software link to a storage device which would store the encoded video data. The transmitter (640) may merge coded video data from the video encoder (603) with other data to be transmitted, for example, coded audio data and/or ancillary data streams (sources not shown).

[0084] The controller (650) may manage operation of the video encoder (603). During coding, the controller (650) may assign to each coded picture a certain coded picture type, which may affect the coding techniques that may be applied to the respective picture. For example, pictures often may be assigned as one of the following picture types:

[0085] An Intra Picture (I picture) may be one that may be coded and decoded without using any other picture in the sequence as a source of prediction. Some video codecs allow for different types of intra pictures, including, for example Independent Decoder Refresh ("IDR") Pictures. A person skilled in the art is aware of those variants of I pictures and their respective applications and features.

[0086] A predictive picture (P picture) may be one that may be coded and decoded using intra prediction or inter prediction using at most one motion vector and reference index to predict the sample values of each block.

[0087] A bi-directionally predictive picture (B Picture) may be one that may be coded and decoded using intra prediction or inter prediction using at most two motion vectors and reference indices to predict the sample values of each block. Similarly, multiple-predictive pictures can use more than two reference pictures and associated metadata for the reconstruction of a single block.

[0088] Source pictures commonly may be subdivided spatially into a plurality of sample blocks (for example, blocks of 4x4, 8x8, 4x8, or 16x16 samples each) and coded on a block-by-block basis. Blocks may be coded predictively with reference to other (already coded) blocks as determined by the coding assignment applied to the blocks' respective pictures. For example, blocks of I pictures may be coded non-predictively or they may be coded predictively with reference to already coded blocks of the same picture (spatial prediction or intra prediction). Pixel blocks of P pictures may be coded predictively, via spatial prediction or via temporal prediction with reference to one previously coded reference picture. Blocks of B pictures may be

coded predictively, via spatial prediction or via temporal prediction with reference to one or two previously coded reference pictures.

[0089] The video encoder (603) may perform coding operations according to a predetermined video coding technology or standard, such as ITU-T Rec. H.265. In its operation, the video encoder (603) may perform various compression operations, including predictive coding operations that exploit temporal and spatial redundancies in the input video sequence. The coded video data, therefore, may conform to a syntax specified by the video coding technology or standard being used.

[0090] In an embodiment, the transmitter (640) may transmit additional data with the encoded video. The source coder (630) may include such data as part of the coded video sequence. Additional data may comprise temporal/spatial/SNR enhancement layers, other forms of redundant data such as redundant pictures and slices, SEI messages, VUI parameter set fragments, and so on.

[0091] A video may be captured as a plurality of source pictures (video pictures) in a temporal sequence. Intra-picture prediction (often abbreviated to intra prediction) makes use of spatial correlation in a given picture, and inter-picture prediction makes uses of the (temporal or other) correlation between the pictures. In an example, a specific picture under encoding/decoding, which is referred to as a current picture, is partitioned into blocks. When a block in the current picture is similar to a reference block in a previously coded and still buffered reference picture in the video, the block in the current picture can be coded by a vector that is referred to as a motion vector. The motion vector points to the reference block in the reference picture, and can have a third dimension identifying the reference picture, in case multiple reference pictures are in use.

[0092] In some embodiments, a bi-prediction technique can be used in the inter-picture prediction. According to the bi-prediction technique, two reference pictures, such as a first reference picture and a second reference picture that are both prior in decoding order to the current picture in the video (but may be in the past and future, respectively, in display order) are used. A block in the current picture can be coded by a first motion vector that points to a first reference block in the first reference picture, and a second motion vector that points to a second reference block in the second reference picture. The block can be predicted by a combination of the first reference block and the second reference block.

[0093] Further, a merge mode technique can be used in the inter-picture prediction to improve coding efficiency.

[0094] According to some embodiments of the disclosure, predictions, such as inter-picture predictions and intra-picture predictions, are performed in the unit of blocks. For example, according to the HEVC standard, a picture in a sequence of video pictures is partitioned into coding tree units (CTU) for compression, the CTUs in a picture have the same size, such as 64x64 pixels, 32x32 pixels, or 16x16 pixels. In general, a CTU includes three coding tree blocks (CTBs), which are one luma CTB and two chroma CTBs. Each CTU can be recursively quadtree split into one or multiple coding units (CUs). For example, a CTU of 64x64 pixels can be split into one CU of 64x64 pixels, or 4 CUs of 32x32 pixels, or 16 CUs of 16x16 pixels. In an example, each CU is analyzed to determine a prediction type for the CU, such as an inter prediction type or an intra prediction type. The CU is split into one or more prediction units (PUs) depending on the temporal and/or spatial predictability. Generally, each PU includes a luma prediction block (PB), and two chroma PBs. In an embodiment, a prediction operation in coding (encoding/decoding) is performed in the unit of a prediction block. Using a luma prediction block as an example of a prediction block, the prediction block includes a matrix of values (e.g., luma values) for pixels, such as 8x8 pixels, 16x16 pixels, 8x16 pixels, 16x8 pixels, and the like.

[0095] FIG. 7 shows an exemplary diagram of a video encoder (703). The video encoder (703) is configured to receive a processing block (e.g., a prediction block) of sample values within a current video picture in a sequence of video pictures, and encode the processing block into a coded picture that is part of a coded video sequence. In an example, the video encoder (703) is used in the place of the video encoder (403) in the FIG. 4 example.

[0096] In an HEVC example, the video encoder (703) receives a matrix of sample values for a processing block, such as a prediction block of 8x8 samples, and the like. The video encoder (703) determines whether the processing block is best coded using intra mode, inter mode, or bi-prediction mode using, for example, rate-distortion optimization. When the processing block is to be coded in intra mode, the video encoder (703) may use an intra prediction technique to encode the processing block into the coded picture; and when the processing block is to be coded in inter mode or bi-prediction mode, the video encoder (703) may use an inter prediction or bi-prediction technique, respectively, to encode the processing block into the coded picture. In certain video coding technologies, merge mode can be an inter

picture prediction submode where the motion vector is derived from one or more motion vector predictors without the benefit of a coded motion vector component outside the predictors. In certain other video coding technologies, a motion vector component applicable to the subject block may be present. In an example, the video encoder (703) includes other components, such as a mode decision module (not shown) to determine the mode of the processing blocks.

[0097] In the FIG. 7 example, the video encoder (703) includes an inter encoder (730), an intra encoder (722), a residue calculator (723), a switch (726), a residue encoder (724), a general controller (721), and an entropy encoder (725) coupled together as shown in FIG. 7.

[0098] The inter encoder (730) is configured to receive the samples of the current block (e.g., a processing block), compare the block to one or more reference blocks in reference pictures (e.g., blocks in previous pictures and later pictures), generate inter prediction information (e.g., description of redundant information according to inter encoding technique, motion vectors, merge mode information), and calculate inter prediction results (e.g., predicted block) based on the inter prediction information using any suitable technique. In some examples, the reference pictures are decoded reference pictures that are decoded based on the encoded video information.

[0099] The intra encoder (722) is configured to receive the samples of the current block (e.g., a processing block), in some cases compare the block to blocks already coded in the same picture, generate quantized coefficients after transform, and in some cases also generate intra prediction information (e.g., an intra prediction direction information according to one or more intra encoding techniques). In an example, the intra encoder (722) also calculates intra prediction results (e.g., predicted block) based on the intra prediction information and reference blocks in the same picture.

[0100] The general controller (721) is configured to determine general control data and control other components of the video encoder (703) based on the general control data. In an example, the general controller (721) determines the mode of the block, and provides a control signal to the switch (726) based on the mode. For example, when the mode is the intra mode, the general controller (721) controls the switch (726) to select the intra mode result for use by the residue calculator (723), and controls the entropy encoder (725) to select the intra prediction information and include the intra prediction information in the bitstream; and when the mode is the inter mode, the general controller (721) controls the switch (726) to select the inter prediction

result for use by the residue calculator (723), and controls the entropy encoder (725) to select the inter prediction information and include the inter prediction information in the bitstream.

[0101] The residue calculator (723) is configured to calculate a difference (residue data) between the received block and prediction results selected from the intra encoder (722) or the inter encoder (730). The residue encoder (724) is configured to operate based on the residue data to encode the residue data to generate the transform coefficients. In an example, the residue encoder (724) is configured to convert the residue data from a spatial domain to a frequency domain, and generate the transform coefficients. The transform coefficients are then subject to quantization processing to obtain quantized transform coefficients. In various embodiments, the video encoder (703) also includes a residue decoder (728). The residue decoder (728) is configured to perform inverse-transform, and generate the decoded residue data. The decoded residue data can be suitably used by the intra encoder (722) and the inter encoder (730). For example, the inter encoder (730) can generate decoded blocks based on the decoded residue data and inter prediction information, and the intra encoder (722) can generate decoded blocks based on the decoded residue data and the intra prediction information. The decoded blocks are suitably processed to generate decoded pictures and the decoded pictures can be buffered in a memory circuit (not shown) and used as reference pictures in some examples.

[0102] The entropy encoder (725) is configured to format the bitstream to include the encoded block. The entropy encoder (725) is configured to include various information in the bitstream according to a suitable standard, such as the HEVC standard. In an example, the entropy encoder (725) is configured to include the general control data, the selected prediction information (e.g., intra prediction information or inter prediction information), the residue information, and other suitable information in the bitstream. Note that, according to the disclosed subject matter, when coding a block in the merge submode of either inter mode or bi-prediction mode, there is no residue information.

[0103] FIG. 8 shows an exemplary diagram of a video decoder (810). The video decoder (810) is configured to receive coded pictures that are part of a coded video sequence, and decode the coded pictures to generate reconstructed pictures. In an example, the video decoder (810) is used in the place of the video decoder (410) in the FIG. 4 example.

[0104] In the FIG. 8 example, the video decoder (810) includes an entropy decoder (871), an inter decoder (880), a residue decoder (873), a reconstruction module (874), and an intra decoder (872) coupled together as shown in FIG. 8.

[0105] The entropy decoder (871) can be configured to reconstruct, from the coded picture, certain symbols that represent the syntax elements of which the coded picture is made up. Such symbols can include, for example, the mode in which a block is coded (such as, for example, intra mode, inter mode, bi-predicted mode, the latter two in merge submode or another submode) and prediction information (such as, for example, intra prediction information or inter prediction information) that can identify certain sample or metadata that is used for prediction by the intra decoder (872) or the inter decoder (880), respectively. The symbols can also include residual information in the form of, for example, quantized transform coefficients, and the like. In an example, when the prediction mode is inter or bi-predicted mode, the inter prediction information is provided to the inter decoder (880); and when the prediction type is the intra prediction type, the intra prediction information is provided to the intra decoder (872). The residual information can be subject to inverse quantization and is provided to the residue decoder (873).

[0106] The inter decoder (880) is configured to receive the inter prediction information, and generate inter prediction results based on the inter prediction information.

[0107] The intra decoder (872) is configured to receive the intra prediction information, and generate prediction results based on the intra prediction information.

[0108] The residue decoder (873) is configured to perform inverse quantization to extract de-quantized transform coefficients, and process the de-quantized transform coefficients to convert the residual information from the frequency domain to the spatial domain. The residue decoder (873) may also require certain control information (to include the Quantizer Parameter (QP)), and that information may be provided by the entropy decoder (871) (data path not depicted as this may be low volume control information only).

[0109] The reconstruction module (874) is configured to combine, in the spatial domain, the residual information as output by the residue decoder (873) and the prediction results (as output by the inter or intra prediction modules as the case may be) to form a reconstructed block, that may be part of the reconstructed picture, which in turn may be part of the reconstructed video. It is noted that other suitable operations, such as a deblocking operation and the like, can be performed to improve the visual quality.

[0110] It is noted that the video encoders (403), (603), and (703), and the video decoders (410), (510), and (810) can be implemented using any suitable technique. In an embodiment, the video encoders (403), (603), and (703), and the video decoders (410), (510), and (810) can be

implemented using one or more integrated circuits. In another embodiment, the video encoders (403), (603), and (603), and the video decoders (410), (510), and (810) can be implemented using one or more processors that execute software instructions.

[0111] In intra prediction or an intra prediction mode, sample values of a coding block can be predicted from neighboring samples that are already reconstructed or reconstructed neighboring samples (referred to as reference samples).

[0112] An example of intra prediction is directional intra prediction. In the directional intra prediction, a sample (e.g., a current sample) in a current block can be predicted using a reference sample (e.g., a prediction sample) or an interpolated reference sample. For example, a connected line between the current sample and the prediction sample forms a given angle direction, such as used in an angular mode.

[0113] In another example of intra prediction, a planar mode that is based on sample interpolation is used. In the planar mode, one or more key positions in or around the current block can be predicted using neighboring reference samples. Other positions in the current block can be predicted as a linear combination of sample(s) at the one or more key positions and the reference samples. Weights (e.g., combination weights) may be determined according to a location of the current sample in the current block.

[0114] An example of the planar mode, such as in VVC, is calculated below.

$$predV[x][y] = ((H - 1 - y) \times p[x][-1] + (y + 1) \times p[-1][H]) << Log2(W) \quad \text{Eq. 1}$$

$$predH[x][y] = ((W - 1 - x) \times p[-1][y] + (x + 1) \times p[W][-1]) << Log2(H) \quad \text{Eq. 2}$$

$$pred[x][y] = (predV[x][y] + predH[x][y] + W \times H) >> (Log2(W) + Log2(H) + 1) \quad \text{Eq. 3}$$

[0115] Referring to FIG. 9, a current block (900) includes samples at positions (0, 0) to (H-1, W-1) in the current block (900). W and H are a width and a height of the current block (900), respectively. A predicted sample value pred[x][y] of a current sample at a position (x, y) (x=0, 1, …, or W-1, and y=0, 1,…, or H-1) in the current block (900) can be obtained as a weighted average of reference sample values (e.g., p[−1][y], p[x][−1], p[−1][H], and p[W][-1]) of reference samples located at positions (−1, y), (x, −1), (−1, H), and (W, -1), respectively), as shown in Eqs. 1-3. The reference samples can include the reference sample located at the position (-1, y) in the same row as the current sample, the reference sample at the position (x, −1) in the same column as the current sample, the reference sample at the bottom-left position (-1, H) with respect to the current block, and the reference sample at the top-right position (W, -1) with respect to the current block.

[0116] As described above, the reference sample values of the reference samples can include p[−1][y] of the reference sample located at the position (-1, y), p[x][−1] of the reference sample located at the position (x, −1), p[−1][H] of the reference sample located at the bottom-left position (-1, H), and p[W][−1] of the reference sample located at the top-right position (W, -1). In an example shown in Eq. 1, a vertical predictor predV[x][y] is determined based on the reference samples located at the positions (x, −1) and (-1, H). In an example shown in Eq. 2, a horizontal predictor predH[x][y] is determined based on the reference samples located at the positions (-1, y) and (W, -1). In Eq. 3, the predicted sample value pred[x][y] is determined based on an average (e.g., a weighted average) of the horizontal predictor predH[x][y] and the vertical predictor predV[x][y].

[0117] A cross-component linear model prediction (CCLM) mode is a cross-component prediction method. In the CCLM, chroma samples can be predicted based on reconstructed luma samples using a linear model. The linear model can be built by neighboring already constructed samples of a current block (e.g., a chroma block to be coded). In some embodiments, the prediction performance is high when the luma and chroma channels are highly linearly correlated.

[0118] In an embodiment, a chroma block (e.g., a current Cb block or a current Cr block) is predicted based on a collocated luma block. A prediction block Pred_C of the chroma block can be derived as follows where a sample Pred_C(x, y) in the prediction block Pred_C of the chroma block can be determined based on sample(s) in the collocated luma block that is already reconstructed.

$$Pred_C(x, y) = a \times Rec_L'(x, y) + b \quad \text{Eq. 4}$$

[0119] Pred_C(x, y) represents a predicted chroma sample at a sample location (x, y) of the chroma block (e.g., in a chroma channel of a current picture or a chroma picture). Rec_L'(x, y) can be determined from reconstructed sample(s) in the collocated luma block (e.g., in a luma channel of the current picture or a luma picture) that is already reconstructed. Rec_L' (x, y) can represent a reconstructed luma sample in the collocated luma block or a downsampled luma sample of the collocated luma block.

[0120] In an example, Rec_L' is the collocated luma block that is already reconstructed, for example, when a color format is 4:4:4 and a size of the chroma block is identical to a size of the collocated luma block. Thus, Rec_L'(x, y) can represent a reconstructed sample in the

collocated luma block where the reconstructed sample corresponds to the sample location (x, y) of the chroma block.

[0121] In an example, Rec_L' is different from the collocated luma block that is already reconstructed. The current chroma block is collocated with the collocated luma block, and the luma channel and the chroma channel have different resolutions. In an example, when a color format is 4:2:0, the resolution of the luma block is two times that of the chroma block both vertically and horizontally. Therefore, when the color format is 4:2:0, Rec_L' can be a down-sampled block of the corresponding luma block to match the chroma block size during the derivation of linear model. In some embodiments, the collocated luma block is down-sampled when the color format is not 4:4:4.

[0122] Parameters (e.g., model parameters) a and b in Eq. 4 can represent a slope and an offset in the linear model shown in Eq. 4, and can be referred to as a slope parameter and an offset parameter, respectively. The parameters a and b in Eq. 4 can be derived from reconstructed neighboring samples (e.g., chroma samples and luma samples) around (i) the current chroma block in the chroma channel and (ii) the collocated luma block in the luma channel. The parameters a and b can be determined using any suitable method.

[0123] In an example, the parameters a and b are determined with the classical linear regression theory. A minimum linear least square solution between (i) reconstructed neighboring luma samples or down-sampled samples of the reconstructed neighboring luma samples and (ii) reconstructed neighboring chroma samples can be applied to derive the parameters a and b below.

$$a = \frac{N * \sum_{i=1}^{N} (Rec_C(i) * Rec_L'(i)) - \sum_{i=1}^{N} Rec_C(i) * \sum_{i=1}^{N} Rec_L'(i)}{N * \sum_{i=1}^{N} (Rec_L'(i) * Rec_L'(i)) - \sum_{i=1}^{N} Rec_L'(i) * \sum_{i=1}^{N} Rec_L'(i)} \quad \text{Eq. 5}$$

$$b = \frac{\sum_{i=1}^{N} (Rec_C(i) - a * \sum_{i=1}^{N} Rec_L'(i)}{N} \quad \text{Eq. 6}$$

[0124] In Eqs. 5-6, N reconstructed neighboring chroma samples $Rec_C(i)$ and N corresponding luma samples $Rec_L'(i)$ are used. In an example, such as when the color format is 4:4:4, the N corresponding luma samples $Rec_L'(i)$ includes N reconstructed neighboring luma samples of the luma block. In an example, such as when the color format is 4:2:0, the N corresponding luma samples $Rec_L'(i)$ includes N down-sampled samples of the reconstructed neighboring luma samples of the luma block. i can be an integer from 1 to N.

[0125] The N reconstructed neighboring chroma samples used to determine the parameters a and b can include any suitable neighboring samples that are already reconstructed of the chroma block. The reconstructed neighboring luma samples used to determine the parameters a and b can include any suitable neighboring samples that are already reconstructed of the collocated luma block.

[0126] The prediction process of the CCLM mode can include: (1) down-sampling the collocated luma block and the reconstructed neighboring luma samples of the collocated luma block to obtain Rec_L' and the down-sampled neighboring luma samples and thus match a size of the corresponding chroma block, (2) deriving the parameters a and b based on the down-sampled neighboring luma samples and the reconstructed neighboring chroma samples, for example, using Eqs. 5-6, and (3) applying the CCLM model (e.g., Eq. 4) to generate the chroma prediction block Pred_C. In some examples, step (1) is omitted when the spatial resolutions of the collocated luma block and the chroma block are identical and step (2) is based on the reconstructed neighboring luma samples.

[0127] FIG. 10 shows examples of reconstructed neighboring luma samples and reconstructed neighboring chroma samples used in the CCLM derivation. A chroma block (1000) is under reconstruction. A width and a height of the chroma block (1000) are M (e.g., 8) and N (e.g., 4), respectively. N and M can be positive integers. A luma block (e.g., a collocated luma block) (1001) that is collocated with the chroma block (1000) is used to predict the chroma block (1000). The luma block (1001) includes luma samples (1040). The luma block (1001) can have any suitable width and any suitable height. In the example shown in FIG. 10, the width and the height of the luma block (1001) are 2M (e.g., 16) and 2N (e.g., 8), respectively.

[0128] Neighboring chroma samples (1010) (e.g., shaded in gray) of the chroma block (1000) are already reconstructed. Neighboring luma samples (1020) of the luma block (1001) are already reconstructed. The neighboring chroma samples (1010) can include top neighboring chroma samples (1011) and left neighboring chroma samples (1012). The neighboring luma samples (1020) of the luma block (1001) can include top neighboring luma samples (1021) and left neighboring luma samples (1022).

[0129] In the example of FIG. 10, the neighboring luma samples (1020) are sub-sampled or down-sampled to generate down-sampled neighboring luma samples (1030) (e.g., shaded in gray) to match a number (e.g., 12) of the neighboring chroma samples. In the example shown in

FIG. 10, the neighboring chroma samples (1010) and the down-sampled neighboring luma samples (1030) can be used to determine parameters a and b, such as shown in Eqs. 5-6.

[0130] In an example, the neighboring chroma samples of the chroma block (1000) only include the top neighboring chroma samples (1011) and does not include the left neighboring chroma samples (1012). Accordingly, the neighboring luma samples of the luma block (1001) only include the top neighboring luma samples (1021) and does not include the left neighboring luma samples (1022). As described above, the top neighboring luma samples (1021) can be down-sampled (e.g., the down-sampled luma samples are shaded in gray) to match a number (e.g., 8) of the top neighboring chroma samples (1011).

[0131] In an example, the neighboring chroma samples of the chroma block (1000) only include the left neighboring chroma samples (1012) and does not include the top neighboring chroma samples (1011). Accordingly, the neighboring luma samples of the luma block (1001) only include the left neighboring luma samples (1022) and does not include the top neighboring luma samples (1021). As described above, the left neighboring luma samples (1022) can be down-sampled (e.g., the down-sampled luma samples are shaded in gray) to match a number (e.g., 4) of the left neighboring chroma samples (1012).

[0132] In some examples, chroma samples in neighboring reconstructed blocks of the chroma block (1000) and corresponding luma samples in neighboring reconstructed blocks of the luma block (1001) can be used to determine the parameters a and b.

[0133] The chroma block (1000) can be a Cb block in a Cb channel or a Cr block in a Cr channel. In an example, for each chroma channel (e.g., Cr or Cb), the parameters a and b can be determined separately. For example, the parameters a and b of the Cr block are determined based on chroma reconstructed neighboring samples of the Cr block, and the parameters a and b of the Cb block are determined based on chroma reconstructed neighboring samples of the Cb block.

[0134] FIG. 10 shows examples of the reconstructed neighboring luma samples and the reconstructed neighboring chroma samples used in the CCLM derivation for the chroma block (1000) that has a rectangular shape. The descriptions in FIG. 10 can be adapted to a chroma block and a collocated luma block that have any shape, such as a square shape. FIG. 11 shows examples of reconstructed neighboring luma samples (1120) and reconstructed neighboring chroma samples (1110) used in the CCLM derivation for a chroma block (1100). A width and a height of the chroma block (1100) are W1 (e.g., 8) and H1 (e.g., 8), respectively where W1 is equal to H1. A luma block (e.g., a collocated luma block) (1101) that is collocated with the

chroma block (1100) is used to predict the chroma block (1100). A width and a height of the luma block (1101) are 2W1 (e.g., 16) and 2H1 (e.g., 16), respectively.

[0135] In some embodiments, the reference samples used to generate the linear model parameters a and b are noisy and/or less representative to a content inside an actual prediction block. Thus, the prediction may be suboptimal with respect to coding efficiency.

[0136] The disclosure describes a cross-component planar prediction method in image and video compression. In an embodiment, a cross-component prediction (e.g., a cross-component linear model (CCLM) mode) and a planar mode (e.g., a planar intra prediction) are combined in image and video compression.

[0137] According to an embodiment of the disclosure, a cross-component planar prediction can be performed on a chroma block to be coded (e.g., to be reconstructed or to be encoded) as follows. One or more chroma samples in the chroma block can be predicted based on one or more luma samples in a luma block (also referred to as a collocated luma block) that is collocated with the chroma block. A remaining chroma sample in the chroma block can be predicted based on the predicted one or more chroma samples in the chroma block and reference chroma sample(s) outside the chroma block, for example, using a planar mode. The remaining chroma sample is different from the one or more chroma samples. The chroma block and the reference chroma sample(s) can be in a chroma channel (e.g., a Cb channel or a Cr channel) of a current picture. The luma block is in the current picture. The reference chroma sample(s) can be adjacent to the chroma block. The reference chroma sample(s) can include chroma sample(s) in neighboring chroma block(s) of the chroma block to be coded. For example, the reference chroma sample(s) include top reference chroma sample(s) that are above the chroma block and/or left reference chroma sample(s) that are to the left of the chroma block. The chroma block and the luma block can have an identical resolution or have different resolutions.

[0138] The cross-component planar prediction can be more accurate than the planar mode without the cross-component prediction because the one or more chroma samples in the cross-component planar prediction are predicted from luma sample(s) in the collocated luma block in the same picture (e.g., the current picture).

[0139] In an example, key position(s) (e.g., the one or more chroma samples at the key position(s)) of the chroma block are predicted using luma information (e.g., luma sample(s) in the collocated luma block) using cross-component prediction, such as the CCLM mode. Subsequently, remaining positions (e.g., all other positions) in the chroma block can be predicted

using sample value(s) at the key positions and the reference sample(s) (e.g., the top reference sample(s) and the left reference sample(s)) of the chroma block. In an example, a planar-like slope effect can be expected for the prediction block of the current block. A planar slope in the cross-component planar prediction can be more accurate than the planar mode without the cross-component prediction, and thus the cross-component planar prediction can be more accurate.

[0140] FIG. 12 shows exemplary samples or sample positions in a chroma block (1200). The chroma block (1200) is to be coded. In an example, the chroma block (1200) is under reconstruction. A luma block (also referred to as a collocated luma block) (not shown) is collocated with the chroma block (1200). The chroma block (1200) and the luma block are in a current picture. The chroma block (1200) has a width W (e.g., 8) and a height H (e.g., 8). W and H can be any suitable positive integers. W can be identical to or different from H.

[0141] Reference chroma samples (1210) (shaded by gray) can include top reference chroma samples (1211) and left reference chroma samples (1212). The top reference chroma samples (1211) include a top-right reference chroma sample (1204) (referred to as p[W-1][-1]). The left reference chroma samples (1212) include a bottom-left reference chroma sample (1203) (referred to as p[-1][H-1]).

[0142] In an example, the reference chroma samples (1210) further includes a top-left reference chroma sample (1205). In an example, the reference chroma samples (1210) do not include the top-left reference chroma sample (1205).

[0143] The cross-component planar prediction can be used to predict the chroma block (1200). A chroma sample in the chroma block (1200), such as a bottom-right corner chroma sample (1202) at a bottom-right corner position of the chroma block (1200), can be predicted based on luma sample(s) using a cross-component prediction technology (e.g., the CCLM mode). Subsequently, the predicted chroma sample, such as the predicted bottom-right corner chroma sample (1202) referred to as pred[W-1][H-1], can be used to generate (i) prediction samples (e.g., pred[x][H-1] where x is an integer from 0 to W-1) of chroma samples in a bottom row (1213) of the chroma block (1200) and/or (ii) prediction samples (e.g., pred[W-1][y] where y is an integer from 0 to H-1) of chroma samples in a right most column (1214) of the chroma block (1200).

[0144] For example, the predicted chroma sample pred[x][H-1] in the bottom row (1213) is generated based on an average (e.g., a linear interpolation) of the predicted bottom-right corner chroma sample pred[W-1][H-1] (1202) and the bottom-left reference chroma sample

p[-1][H-1] (1203), such as shown in Eq. 7. Similarly, the predicted chroma sample pred[W-1][y] in the right most column (1214) can be generated using an average (e.g., a linear interpolation) of the predicted bottom-right corner chroma sample pred[W-1][H-1] (1202) and the top-right reference chroma sample p[W-1][-1] (1204), such as shown in Eq. 8.

$$pred[x][H-1] = ((W-1-x) \times p[-1][H-1] + (x+1) \times p[W-1][H-1]) << Log2(W) \quad \text{Eq. 7}$$

$$pred[W-1][y] = ((H-1-y) \times p[W-1][-1] + (y+1) \times p[W-1][H-1]) << Log2(H) \quad \text{Eq. 8}$$

[0145] Weight(s) used in Eqs. 7-8 can be determined by a location of the predicted chroma sample pred[x][H-1] in the bottom row (1213) or the predicted chroma sample pred[W-1][y] in the right most column (1214). The weight (e.g., (W − 1 − x)) of the predictor (or the predictor sample) (e.g., p[−1][H−1]) can increase when a distance between the predicted sample (e.g., pred[x][H-1]) and the predictor decreases. In an example, such as shown in Eqs. 7-8, the relationship between (i) the weight of the predictor and (ii) the distance between the predicted sample and the predictor is linear.

[0146] For a chroma sample (1201) located at a position (x, y) in the chroma block (1200) (e.g., x=0,1,…, W-1, and y=0, 1, …, H-1), the prediction sample (also referred to as the predicted sample) pred[x][y] can be generated by averaging a horizontal predictor predH[x][y] and a vertical predictor predV[x][y], such as shown in Eq. 9.

$$pred[x][y] = (predV[x][y]+predH[x][y]+W \times H) >> (Log2(W) + Log2(H) + 1) \quad \text{Eq. 9}$$

[0147] The horizontal predictor predH[x][y] of the chroma sample (1201) can be generated based on a left reference chroma sample (1222) (e.g., pred[-1][y]) in the left reference chroma samples (1212) and a predicted right chroma sample (1224) (e.g., p[W−1][y]) in the predicted chroma samples in the right most column (1214). For example, the horizontal predictor predH[x][y] is a weighted average or a linear interpolation of the left reference chroma sample (1222) (e.g., pred[-1][y]) and the predicted right chroma sample (1224) (e.g., p[W−1][y]), such as shown in Eq. 10. In an example, chroma sample (1201), the left reference chroma sample (1222), and the predicted right chroma sample (1224) are in a same row of the chroma block (1200).

[0148] The vertical predictor predV[x][y] of the chroma sample (1201) can be generated based on a top reference chroma sample (1221) (e.g., pred[x][-1]) in the top reference chroma samples (1211) and a predicted bottom chroma sample (1223) (e.g., p[x][H-1]) in the predicted chroma samples in the bottom row (1213). For example, the vertical predictor predV[x][y] is a

weighted average or a linear interpolation of the top reference chroma sample (1221) (e.g., pred[x][-1]) and the predicted bottom chroma sample (1223) (e.g., p[x][H-1]), such as shown in Eq. 11. In an example, chroma sample (1201), the top reference chroma sample (1221), and the predicted bottom chroma sample (1223) are in a same column of the chroma block (1200).

[0149] Weight(s) (e.g., ((W − 1 − x), (x + 1), (H − 1 − y), and (y + 1)) used to generate the horizontal predictor predH[x][y] and the vertical predictor predV[x][y] in Eqs. 10-11 can be determined by a location of the chroma sample (1201). The weight (e.g., (W − 1 − x)) of the predictor (or the predictor sample) (e.g., p[−1][y]) can increase when a distance between the chroma sample (1201) and the predictor decreases. In the examples shown in Eqs. 10-11, the relationship between (i) the weight of the respective predictor and (ii) the distance between the chroma sample (1201) and the predictor predictor is linear.

$$predH[x][y] = ((W - 1 - x) \times p[-1][y] + (x + 1) \times p[W-1][y]\ ) << Log2\ (W) \quad \text{Eq. 10}$$

$$predV[x][y] = ((H - 1 - y) \times p[x][-1] + (y + 1) \times p[x][H-1]) << Log2\ (H) \quad \text{Eq. 11}$$

[0150] In an embodiment, the predicted chroma samples pred[W-1][y] in the right most column (1214) (e.g., all the predicted chroma samples pred[W-1][y] in the entire right most column (1214)) and the predicted chroma samples pred[x][H-1] in the bottom row (1213) (e.g., all the predicted chroma samples pred[x][H-1] in the entire bottom row (1213)) can be predicted from luma samples in the collocated luma block using a cross-component prediction technology (e.g., the CCLM mode). For example, the predicted chroma samples pred[W-1][y] in the entire right most column (1214)) and the predicted chroma samples pred[x][H-1] in the entire bottom row (1213) are from the cross-component prediction directly and are directly predicted from the luma samples in the collocated luma block without using Eqs. 7-8. Subsequently, the predicted sample pred[x][y] of the chroma sample (1201) can be generated by averaging the horizontal predictor predH[x][y] and the vertical predictor predV[x][y], such as described above in Eqs. 9-11.

[0151] In some examples, a predicted chroma sample (e.g., pred[W-1][H-1])) is used directly to predict another chroma sample (e.g., pred[x][H-1]) or pred[W-1][y])), such as described above in Eqs. 7-11. In an example, a respective residue is added to the predicted chroma sample (e.g., pred[W-1][H-1])) to obtain an updated predicted chroma sample (e.g., a reconstructed sample) where the updated predicted chroma sample is used to predict the other chroma sample (e.g., pred[x][H-1]) or pred[W-1][y])).

[0152] In an embodiment, the horizontal predictor predH[x][y] or the vertical predictor predV[x][y] can be used separately, for example, when a chroma block is rectangular (non-square). For example, when the width W of the chroma block is larger than the height H of the chroma block, only the vertical predictor predV[x][y] is used. For example, when the width W is less than the height H, only the horizontal predictor predH[x][y] is used.

[0153] In an embodiment, the horizontal predictor predH[x][y] or the vertical predictor predV[x][y] can be used when a block size condition is met. For example, when the width W is less than a width threshold (e.g., a preset constant W0), the horizontal predictor predH[x][y] is allowed. When the height H is less than a height threshold (e.g., a preset constant H0), the vertical predictor predV[x][y] is allowed. When the horizontal predictor predH[x][y] and the vertical predictor predV[x][y] are allowed, the final prediction pred[x][y] can be an average of the horizontal predictor predH[x][y] and the vertical predictor predV[x][y], such as described above in Eq. 9.

[0154] In an example, no prediction is performed for chroma samples in the right most column (1214) except the bottom-right corner chroma sample (1202) when only the vertical predictor predV[x][y] is used. In an example, no prediction is performed for chroma samples in the bottom row (1213) except the bottom-right corner chroma sample (1202) when only the horizontal predictor predH[x][y] is used.

[0155] FIG. 13 shows a flow chart outlining a process (e.g., an encoding process) (1300) according to an embodiment of the disclosure. The process (1300) can be executed by an apparatus for video coding that can include processing circuitry. The processing circuitry in the apparatus, such as the processing circuitry in the terminal devices (310), (320), (330) and (340), processing circuitry that performs functions of a video encoder (e.g., (403), (603), (703)), or the like can be configured to execute the process (1300). In some embodiments, the process (1300) is implemented in software instructions, thus when the processing circuitry executes the software instructions, the processing circuitry performs the process (1300). The process starts at (S1301), and proceeds to (S1310).

[0156] At (S1310), one or more first chroma samples in a chroma block in a current picture can be predicted based on one or more luma samples in a luma block in the current picture using a planar mode. The luma block is already reconstructed and is collocated with the chroma block.

[0157] In an embodiment, the one or more first chroma samples in the chroma block are predicted based on the one or more luma samples in the luma block that are collocated with the one or more first chroma samples using a cross-component linear model prediction (CCLM) mode.

[0158] At (S1320), a second chroma sample in the chroma block can be predicted based on the one or more first chroma samples and at least one reference chroma sample of the chroma block. The second chroma sample can be different from the one or more first chroma samples. The at least one reference chroma sample can be adjacent to the chroma block.

[0159] In an example, the at least one reference chroma sample includes at least one of a top reference chroma sample that is a top neighbor of the chroma block or a left reference chroma sample that is a left neighbor of the chroma block.

[0160] At (S1330), prediction information of the chroma block indicating that the chroma block is predicted based on the luma block using a planar mode can be encoded. The encoded prediction information can be included in a coded video bitstream.

[0161] The second chroma sample can be encoded based on the predicted second chroma sample. In an example, no residue is signaled for the second chroma sample in the coded video bitstream. In an example, a residue of the second chroma sample is signaled in the coded video bitstream.

[0162] Then, the process (1300) proceeds to (S1399) and terminates.

[0163] The process (1300) can be suitably adapted to various scenarios and steps in the process (1300) can be adjusted accordingly. One or more of the steps in the process (1300) can be adapted, omitted, repeated, and/or combined. Any suitable order can be used to implement the process (1300). Additional step(s) can be added.

[0164] In an embodiment, the one or more first chroma samples is a bottom-right corner chroma sample in the chroma block. Based on the bottom-right corner chroma sample and the at least one reference chroma sample, at least one of (i) chroma samples in a bottom row of the chroma block or (ii) chroma samples in a right most column of the chroma block can be predicted. In an example, the chroma samples in the bottom row of the chroma block are predicted, a vertical predictor of the second chroma sample is generated based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row. The second chroma sample, the top

reference chroma sample, and the predicted bottom chroma sample can be in a same column in the chroma block.

[0165] In an example, the chroma samples in the right most column of the chroma block are predicted, and a horizontal predictor of the second chroma sample is generated based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column. The second chroma sample, the left reference chroma sample, and the predicted right chroma sample can be in a same row of the chroma block.

[0166] In an example, the chroma samples in the bottom row of the chroma block and the chroma samples in the right most column of the chroma block are predicted. The second chroma sample can be predicted based on the vertical predictor and the horizontal predictor.

[0167] In an example, a width of the chroma block is larger than a height of the chroma block or the height of the chroma block is less than a threshold. The chroma samples in the bottom row of the chroma block are predicted and the second chroma sample is predicted based on the vertical predictor.

[0168] In an example, a width of the chroma block is less than a height of the chroma block or a threshold. The chroma samples in the right most column of the chroma block are predicted and the second chroma sample is predicted based on the horizontal predictor.

[0169] In an example, the bottom-right corner chroma sample in the chroma block is predicted based on at least one of a bottom-right corner luma sample in the luma block or luma samples in the luma block that are collocated with the bottom-right corner chroma sample.

[0170] In an embodiment, the one or more first chroma samples include (i) chroma samples in a bottom row of the chroma block and (ii) chroma samples in a right most column of the chroma block. A vertical predictor of the second chroma sample can be generated based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row. The second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample can be in a same column in the chroma block. A horizontal predictor of the second chroma sample can be generated based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column. The second chroma sample, the left reference chroma sample, and the predicted right chroma

sample can be in a same row in the chroma block. The second chroma sample can be predicted based on the vertical predictor and the horizontal predictor.

[0171] In an embodiment, a width of the chroma block is larger than a height of the chroma block or the height of the chroma block is less than a threshold. The one or more first chroma samples include chroma samples in a bottom row of the chroma block. A vertical predictor of the second chroma sample is generated based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row. The second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample can be in a same column in the chroma block. The second chroma sample is predicted based on the vertical predictor.

[0172] In an embodiment, a width of the chroma block is less than a height of the chroma block or a threshold. The one or more first chroma samples include chroma samples in a right most column of the chroma block. A horizontal predictor of the second chroma sample can be generated based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column. The second chroma sample, the left reference chroma sample, and the predicted right chroma sample can be in a same row in the chroma block. The second chroma sample is predicted based on the horizontal predictor.

[0173] FIG. 14 shows a flow chart outlining a process (e.g., a decoding process) (1400) according to an embodiment of the disclosure. The process (1400) can be used in a video decoder. The process (1400) can be executed by an apparatus for video coding that can include receiving circuitry and processing circuitry. The processing circuitry in the apparatus, such as the processing circuitry in the terminal devices (310), (320), (330) and (340), the processing circuitry that performs functions of the video decoder (410), the processing circuitry that performs functions of the video decoder (510), and the like can be configured to execute the process (1400). In some examples, the process (1400) is used in a video encoder (e.g., the video encoder (403), the video encoder (603)). In an example, the process (1400) is executed by processing circuitry that performs functions of a video encoder (e.g., the video encoder (403), the video encoder (603)). In some embodiments, the process (1400) is implemented in software instructions, thus when the processing circuitry executes the software instructions, the processing circuitry performs the process (1400). The process starts at (S1401) and proceeds to (S1410).

[0174] At (S1410), prediction information of a chroma block to be reconstructed in a current picture can be decoded. The prediction information indicates that the chroma block is predicted based at least on a luma block in the current picture using a planar mode. The luma block is already reconstructed and collocated with the chroma block.

[0175] At (S1420), one or more first chroma samples in the chroma block can be predicted based on one or more luma samples in the luma block.

[0176] In an example, the one or more first chroma samples in the chroma block are identified as located at one or more respective key positions before being predicted.

[0177] In an embodiment, the one or more first chroma samples in the chroma block are predicted based on the one or more luma samples in the luma block that are collocated with the one or more first chroma samples using a cross-component linear model prediction (CCLM) mode.

[0178] At (S1430), a second chroma sample in the chroma block can be predicted based on the one or more first chroma samples and at least one reference chroma sample of the chroma block. The second chroma sample can be different from the one or more first chroma samples. In an example, the at least one reference chroma sample is adjacent to the chroma block.

[0179] The second chroma sample is reconstructed based on the predicted second chroma sample. In an example, the reconstructed second chroma sample is the predicted second chroma sample. In an example, the reconstructed second chroma sample is a sum of the predicted second chroma sample and a corresponding residue.

[0180] In an example, the at least one reference chroma sample includes at least one of a top reference chroma sample that is a top neighbor of the chroma block or a left reference chroma sample that is a left neighbor of the chroma block.

[0181] Then, the process (1400) proceeds to (S1499) and terminates.

[0182] The process (1400) can be suitably adapted to various scenarios and steps in the process (1400) can be adjusted accordingly. One or more of the steps in the process (1400) can be adapted, omitted, repeated, and/or combined. Any suitable order can be used to implement the process (1400). Additional step(s) can be added.

[0183] In an embodiment, the one or more first chroma samples is a bottom-right corner chroma sample in the chroma block. Based on the bottom-right corner chroma sample and the at least one reference chroma sample, at least one of (i) chroma samples in a bottom row of the chroma block or (ii) chroma samples in a right most column of the chroma block can be

predicted. In an example, the chroma samples in the bottom row of the chroma block are predicted, a vertical predictor of the second chroma sample is generated based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row. The second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample can be in a same column in the chroma block.

[0184] In an example, the chroma samples in the right most column of the chroma block are predicted, and a horizontal predictor of the second chroma sample is generated based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column. The second chroma sample, the left reference chroma sample, and the predicted right chroma sample can be in a same row of the chroma block.

[0185] In an example, the chroma samples in the bottom row of the chroma block and the chroma samples in the right most column of the chroma block are predicted. The second chroma sample can be predicted based on the vertical predictor and the horizontal predictor.

[0186] In an example, a width of the chroma block is larger than a height of the chroma block or the height of the chroma block is less than a threshold. The chroma samples in the bottom row of the chroma block are predicted and the second chroma sample is predicted based on the vertical predictor.

[0187] In an example, a width of the chroma block is less than a height of the chroma block or a threshold. The chroma samples in the right most column of the chroma block are predicted and the second chroma sample is predicted based on the horizontal predictor.

[0188] In an example, the bottom-right corner chroma sample in the chroma block is predicted based on at least one of a bottom-right corner luma sample in the luma block or luma samples in the luma block that are collocated with the bottom-right corner chroma sample.

[0189] In an embodiment, the one or more first chroma samples include (i) chroma samples in a bottom row of the chroma block and (ii) chroma samples in a right most column of the chroma block. A vertical predictor of the second chroma sample can be generated based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row. The second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample can be in a same column in the chroma block. A horizontal predictor of the second chroma sample can be

generated based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column. The second chroma sample, the left reference chroma sample, and the predicted right chroma sample can be in a same row in the chroma block. The second chroma sample can be predicted based on the vertical predictor and the horizontal predictor.

[0190] In an embodiment, a width of the chroma block is larger than a height of the chroma block or the height of the chroma block is less than a threshold. The one or more first chroma samples include chroma samples in a bottom row of the chroma block. A vertical predictor of the second chroma sample is generated based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row. The second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample can be in a same column in the chroma block. The second chroma sample is predicted based on the vertical predictor.

[0191] In an embodiment, a width of the chroma block is less than a height of the chroma block or a threshold. The one or more first chroma samples include chroma samples in a right most column of the chroma block. A horizontal predictor of the second chroma sample can be generated based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column. The second chroma sample, the left reference chroma sample, and the predicted right chroma sample can be in a same row in the chroma block. The second chroma sample is predicted based on the horizontal predictor.

[0192] Embodiments in the disclosure may be used separately or combined in any order. Further, each of the methods (or embodiments), an encoder, and a decoder may be implemented by processing circuitry (e.g., one or more processors or one or more integrated circuits). In one example, the one or more processors execute a program that is stored in a non-transitory computer-readable medium.

[0193] The techniques described above, can be implemented as computer software using computer-readable instructions and physically stored in one or more computer-readable media. For example, FIG. 15 shows a computer system (1500) suitable for implementing certain embodiments of the disclosed subject matter.

[0194] The computer software can be coded using any suitable machine code or computer language, that may be subject to assembly, compilation, linking, or like mechanisms to

create code comprising instructions that can be executed directly, or through interpretation, micro-code execution, and the like, by one or more computer central processing units (CPUs), Graphics Processing Units (GPUs), and the like.

[0195] The instructions can be executed on various types of computers or components thereof, including, for example, personal computers, tablet computers, servers, smartphones, gaming devices, internet of things devices, and the like.

[0196] The components shown in FIG. 15 for computer system (1500) are exemplary in nature and are not intended to suggest any limitation as to the scope of use or functionality of the computer software implementing embodiments of the present disclosure. Neither should the configuration of components be interpreted as having any dependency or requirement relating to any one or combination of components illustrated in the exemplary embodiment of a computer system (1500).

[0197] Computer system (1500) may include certain human interface input devices. Such a human interface input device may be responsive to input by one or more human users through, for example, tactile input (such as: keystrokes, swipes, data glove movements), audio input (such as: voice, clapping), visual input (such as: gestures), olfactory input (not depicted). The human interface devices can also be used to capture certain media not necessarily directly related to conscious input by a human, such as audio (such as: speech, music, ambient sound), images (such as: scanned images, photographic images obtain from a still image camera), video (such as two-dimensional video, three-dimensional video including stereoscopic video).

[0198] Input human interface devices may include one or more of (only one of each depicted): keyboard (1501), mouse (1502), trackpad (1503), touch screen (1510), data-glove (not shown), joystick (1505), microphone (1506), scanner (1507), camera (1508).

[0199] Computer system (1500) may also include certain human interface output devices. Such human interface output devices may be stimulating the senses of one or more human users through, for example, tactile output, sound, light, and smell/taste. Such human interface output devices may include tactile output devices (for example tactile feedback by the touch-screen (1510), data-glove (not shown), or joystick (1505), but there can also be tactile feedback devices that do not serve as input devices), audio output devices (such as: speakers (1509), headphones (not depicted)), visual output devices (such as screens (1510) to include CRT screens, LCD screens, plasma screens, OLED screens, each with or without touch-screen input capability, each with or without tactile feedback capability—some of which may be capable to output two

dimensional visual output or more than three dimensional output through means such as stereographic output; virtual-reality glasses (not depicted), holographic displays and smoke tanks (not depicted)), and printers (not depicted).

[0200] Computer system (1500) can also include human accessible storage devices and their associated media such as optical media including CD/DVD ROM/RW (1520) with CD/DVD or the like media (1521), thumb-drive (1522), removable hard drive or solid state drive (1523), legacy magnetic media such as tape and floppy disc (not depicted), specialized ROM/ASIC/PLD based devices such as security dongles (not depicted), and the like.

[0201] Those skilled in the art should also understand that term "computer readable media" as used in connection with the presently disclosed subject matter does not encompass transmission media, carrier waves, or other transitory signals.

[0202] Computer system (1500) can also include an interface (1554) to one or more communication networks (1555). Networks can for example be wireless, wireline, optical. Networks can further be local, wide-area, metropolitan, vehicular and industrial, real-time, delay-tolerant, and so on. Examples of networks include local area networks such as Ethernet, wireless LANs, cellular networks to include GSM, 3G, 4G, 5G, LTE and the like, TV wireline or wireless wide area digital networks to include cable TV, satellite TV, and terrestrial broadcast TV, vehicular and industrial to include CANBus, and so forth. Certain networks commonly require external network interface adapters that attached to certain general purpose data ports or peripheral buses (1549) (such as, for example USB ports of the computer system (1500)); others are commonly integrated into the core of the computer system (1500) by attachment to a system bus as described below (for example Ethernet interface into a PC computer system or cellular network interface into a smartphone computer system). Using any of these networks, computer system (1500) can communicate with other entities. Such communication can be uni-directional, receive only (for example, broadcast TV), uni-directional send-only (for example CANbus to certain CANbus devices), or bi-directional, for example to other computer systems using local or wide area digital networks. Certain protocols and protocol stacks can be used on each of those networks and network interfaces as described above.

[0203] Aforementioned human interface devices, human-accessible storage devices, and network interfaces can be attached to a core (1540) of the computer system (1500).

[0204] The core (1540) can include one or more Central Processing Units (CPU) (1541), Graphics Processing Units (GPU) (1542), specialized programmable processing units in the form

of Field Programmable Gate Areas (FPGA) (1543), hardware accelerators for certain tasks (1544), graphics adapters (1550), and so forth. These devices, along with Read-only memory (ROM) (1545), Random-access memory (1546), internal mass storage such as internal non-user accessible hard drives, SSDs, and the like (1547), may be connected through a system bus (1548). In some computer systems, the system bus (1548) can be accessible in the form of one or more physical plugs to enable extensions by additional CPUs, GPU, and the like. The peripheral devices can be attached either directly to the core's system bus (1548), or through a peripheral bus (1549). In an example, the screen (1510) can be connected to the graphics adapter (1550). Architectures for a peripheral bus include PCI, USB, and the like.

[0205] CPUs (1541), GPUs (1542), FPGAs (1543), and accelerators (1544) can execute certain instructions that, in combination, can make up the aforementioned computer code. That computer code can be stored in ROM (1545) or RAM (1546). Transitional data can be also be stored in RAM (1546), whereas permanent data can be stored for example, in the internal mass storage (1547). Fast storage and retrieve to any of the memory devices can be enabled through the use of cache memory, that can be closely associated with one or more CPU (1541), GPU (1542), mass storage (1547), ROM (1545), RAM (1546), and the like.

[0206] The computer readable media can have computer code thereon for performing various computer-implemented operations. The media and computer code can be those specially designed and constructed for the purposes of the present disclosure, or they can be of the kind well known and available to those having skill in the computer software arts.

[0207] As an example and not by way of limitation, the computer system having architecture (1500), and specifically the core (1540) can provide functionality as a result of processor(s) (including CPUs, GPUs, FPGA, accelerators, and the like) executing software embodied in one or more tangible, computer-readable media. Such computer-readable media can be media associated with user-accessible mass storage as introduced above, as well as certain storage of the core (1540) that are of non-transitory nature, such as core-internal mass storage (1547) or ROM (1545). The software implementing various embodiments of the present disclosure can be stored in such devices and executed by core (1540). A computer-readable medium can include one or more memory devices or chips, according to particular needs. The software can cause the core (1540) and specifically the processors therein (including CPU, GPU, FPGA, and the like) to execute particular processes or particular parts of particular processes described herein, including defining data structures stored in RAM (1546) and modifying such

data structures according to the processes defined by the software. In addition or as an alternative, the computer system can provide functionality as a result of logic hardwired or otherwise embodied in a circuit (for example: accelerator (1544)), which can operate in place of or together with software to execute particular processes or particular parts of particular processes described herein. Reference to software can encompass logic, and vice versa, where appropriate. Reference to a computer-readable media can encompass a circuit (such as an integrated circuit (IC)) storing software for execution, a circuit embodying logic for execution, or both, where appropriate. The present disclosure encompasses any suitable combination of hardware and software.

Appendix A: Acronyms

JEM: joint exploration model

VVC: versatile video coding

BMS: benchmark set

MV: Motion Vector

HEVC: High Efficiency Video Coding

SEI: Supplementary Enhancement Information

VUI: Video Usability Information

GOPs: Groups of Pictures

TUs: Transform Units,

PUs: Prediction Units

CTUs: Coding Tree Units

CTBs: Coding Tree Blocks

PBs: Prediction Blocks

HRD: Hypothetical Reference Decoder

SNR: Signal Noise Ratio

CPUs: Central Processing Units

GPUs: Graphics Processing Units

CRT: Cathode Ray Tube

LCD: Liquid-Crystal Display

OLED: Organic Light-Emitting Diode

CD: Compact Disc

DVD: Digital Video Disc

ROM: Read-Only Memory

RAM: Random Access Memory

ASIC: Application-Specific Integrated Circuit

PLD: Programmable Logic Device

LAN: Local Area Network

GSM: Global System for Mobile communications

LTE: Long-Term Evolution

CANBus: Controller Area Network Bus

USB: Universal Serial Bus

PCI: Peripheral Component Interconnect

FPGA: Field Programmable Gate Areas

SSD: solid-state drive

IC: Integrated Circuit

CU: Coding Unit

CCLM: combine cross-component linear model

[0208] While this disclosure has described several exemplary embodiments, there are alterations, permutations, and various substitute equivalents, which fall within the scope of the disclosure. It will thus be appreciated that those skilled in the art will be able to devise numerous systems and methods which, although not explicitly shown or described herein, embody the principles of the disclosure and are thus within the spirit and scope thereof.

WHAT IS CLAIMED IS:

1. A method of video decoding in a decoder, comprising:

decoding prediction information of a chroma block to be reconstructed in a current picture, the prediction information indicating that the chroma block is predicted based at least on a luma block in the current picture using a planar mode, the luma block being reconstructed and collocated with the chroma block;

identifying one or more first chroma samples in the chroma block as located at one or more respective key positions;

predicting the one or more first chroma samples in the chroma block based on one or more luma samples in the luma block; and

predicting a second chroma sample in the chroma block based on the predicted one or more first chroma samples and at least one reference chroma sample of the chroma block, the second chroma sample being different from the one or more first chroma samples, the at least one reference chroma sample being adjacent to the chroma block.

2. The method of claim 1, wherein

the one or more first chroma samples is a bottom-right corner chroma sample in the chroma block; and

the predicting the second chroma sample includes:

predicting, based on the bottom-right corner chroma sample and the at least one reference chroma sample, at least one of (i) chroma samples in a bottom row of the chroma block or (ii) chroma samples in a right most column of the chroma block;

in response to the chroma samples in the bottom row of the chroma block being predicted, generating a vertical predictor of the second chroma sample based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row, the second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample being in a same column in the chroma block; and

in response to the chroma samples in the right most column of the chroma block being predicted, generating a horizontal predictor of the second chroma sample based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column, the second chroma

sample, the left reference chroma sample, and the predicted right chroma sample being in a same row of the chroma block.

3. The method of claim 2, wherein

the predicting the at least one of (i) the chroma samples in the bottom row of the chroma block or (ii) the chroma samples in the right most column of the chroma block includes predicting the chroma samples in the bottom row of the chroma block and the chroma samples in the right most column of the chroma block; and

the predicting the second chroma sample includes predicting the second chroma sample based on the vertical predictor and the horizontal predictor.

4. The method of claim 2, wherein

a width of the chroma block is larger than a height of the chroma block or the height of the chroma block is less than a threshold;

the predicting the at least one of (i) the chroma samples in the bottom row of the chroma block or (ii) the chroma samples in the right most column of the chroma block includes predicting the chroma samples in the bottom row of the chroma block; and

the predicting the second chroma sample includes predicting the second chroma sample based on the vertical predictor.

5. The method of claim 2, wherein

a width of the chroma block is less than a height of the chroma block or a threshold;

the predicting the at least one of (i) the chroma samples in the bottom row of the chroma block or (ii) the chroma samples in the right most column of the chroma block includes predicting the chroma samples in the right most column of the chroma block; and

the predicting the second chroma sample includes predicting the second chroma sample based on the horizontal predictor.

6. The method of claim 2, wherein the predicting the one or more first chroma samples comprises:

predicting the bottom-right corner chroma sample in the chroma block based on at least one of a bottom-right corner luma sample in the luma block or luma samples in the luma block that are collocated with the bottom-right corner chroma sample.

7. The method of claim 1, wherein

the one or more first chroma samples include (i) chroma samples in a bottom row of the chroma block and (ii) chroma samples in a right most column of the chroma block; and

the predicting the second chroma sample includes:

generating a vertical predictor of the second chroma sample based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row, the second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample being in a same column in the chroma block;

generating a horizontal predictor of the second chroma sample based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column, the second chroma sample, the left reference chroma sample, and the predicted right chroma sample being in a same row in the chroma block; and

predicting the second chroma sample based on the vertical predictor and the horizontal predictor.

8. The method of claim 1, wherein

a width of the chroma block is larger than a height of the chroma block or the height of the chroma block is less than a threshold;

the one or more first chroma samples include chroma samples in a bottom row of the chroma block; and

the predicting the second chroma sample includes:

generating a vertical predictor of the second chroma sample based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row, the second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample being in a same column in the chroma block; and

predicting the second chroma sample based on the vertical predictor.

9. The method of claim 1, wherein

a width of the chroma block is less than a height of the chroma block or a threshold;

the one or more first chroma samples include chroma samples in a right most column of the chroma block; and

the predicting the second chroma sample includes:

generating a horizontal predictor of the second chroma sample based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column, the second chroma sample, the left reference chroma sample, and the predicted right chroma sample being in a same row in the chroma block; and

predicting the second chroma sample based on the horizontal predictor.

10. The method of claim 1, wherein the predicting the one or more first chroma samples comprises:

predicting the one or more first chroma samples in the chroma block based on the one or more luma samples in the luma block that are collocated with the one or more first chroma samples using a cross-component linear model prediction (CCLM) mode.

11. The method of claim 1, wherein

the at least one reference chroma sample includes at least one of a top reference chroma sample that is a top neighbor of the chroma block or a left reference chroma sample that is a left neighbor of the chroma block.

12. An apparatus for video decoding, comprising:

processing circuitry configured to:

decode prediction information of a chroma block to be reconstructed in a current picture, the prediction information indicating that the chroma block is predicted based at least on a luma block in the current picture using a planar mode, the luma block being reconstructed and collocated with the chroma block;

identify one or more first chroma samples in the chroma block as located at one or more respective key positions;

predict the one or more first chroma samples in the chroma block based on one or more luma samples in the luma block; and

predict a second chroma sample in the chroma block based on the predicted one or more first chroma samples and at least one reference chroma sample of the chroma block, the second chroma sample being different from the one or more first chroma samples, the at least one reference chroma sample being adjacent to the chroma block.

13. The apparatus of claim 12, wherein

the one or more first chroma samples is a bottom-right corner chroma sample in the chroma block; and

the processing circuitry is configured to:

predict, based on the bottom-right corner chroma sample and the at least one reference chroma sample, at least one of (i) chroma samples in a bottom row of the chroma block or (ii) chroma samples in a right most column of the chroma block;

in response to the chroma samples in the bottom row of the chroma block being predicted, generate a vertical predictor of the second chroma sample based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row, the second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample being in a same column in the chroma block; and

in response to the chroma samples in the right most column of the chroma block being predicted, generate a horizontal predictor of the second chroma sample based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column, the second chroma sample, the left reference chroma sample, and the predicted right chroma sample being in a same row of the chroma block.

14. The apparatus of claim 13, wherein the processing circuitry is configured to:

predict the chroma samples in the bottom row of the chroma block and the chroma samples in the right most column of the chroma block; and

predict the second chroma sample based on the vertical predictor and the horizontal predictor.

15. The apparatus of claim 12, wherein

the one or more first chroma samples include (i) chroma samples in a bottom row of the chroma block and (ii) chroma samples in a right most column of the chroma block; and

the processing circuitry is configured to:

generate a vertical predictor of the second chroma sample based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row, the second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample being in a same column in the chroma block;

generate a horizontal predictor of the second chroma sample based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column, the second chroma sample, the left reference chroma sample, and the predicted right chroma sample being in a same row in the chroma block; and

predict the second chroma sample based on the vertical predictor and the horizontal predictor.

16. The apparatus of claim 12, wherein

a width of the chroma block is larger than a height of the chroma block or the height of the chroma block is less than a threshold;

the one or more first chroma samples include chroma samples in a bottom row of the chroma block; and

the processing circuitry is configured to:

generate a vertical predictor of the second chroma sample based on a top reference chroma sample in the at least one reference chroma sample and a predicted bottom chroma sample in the predicted chroma samples in the bottom row, the second chroma sample, the top reference chroma sample, and the predicted bottom chroma sample being in a same column in the chroma block; and

predict the second chroma sample based on the vertical predictor.

17. The apparatus of claim 12, wherein

a width of the chroma block is less than a height of the chroma block or a threshold;

the one or more first chroma samples include chroma samples in a right most column of the chroma block; and

the processing circuitry is configured to:

generate a horizontal predictor of the second chroma sample based on a left reference chroma sample in the at least one reference chroma sample and a predicted right chroma sample in the predicted chroma samples in the right most column, the second chroma sample, the left reference chroma sample, and the predicted right chroma sample being in a same row in the chroma block; and

predict the second chroma sample based on the horizontal predictor.

18. The apparatus of claim 12, wherein the predicting the one or more first chroma samples comprises:

predicting the one or more first chroma samples in the chroma block based on the one or more luma samples in the luma block that are collocated with the one or more first chroma samples using a cross-component linear model prediction (CCLM) mode.

19. The apparatus of claim 12, wherein

the at least one reference chroma sample includes at least one of a top reference chroma sample that is a top neighbor of the chroma block or a left reference chroma sample that is a left neighbor of the chroma block.

20. A non-transitory computer-readable storage medium storing instructions which when executed by at least one processor cause the at least one processor to perform:

decoding prediction information of a chroma block to be reconstructed in a current picture, the prediction information indicating that the chroma block is predicted based at least on a luma block in the current picture using a planar mode, the luma block being reconstructed and collocated with the chroma block;

identifying one or more first chroma samples in the chroma block as located at one or more respective key positions;

predicting the one or more first chroma samples in the chroma block based on one or more luma samples in the luma block; and

predicting a second chroma sample in the chroma block based on the predicted one or more first chroma samples and at least one reference chroma sample of the chroma block, the second chroma sample being different from the one or more first chroma samples, the at least one reference chroma sample being adjacent to the chroma block.

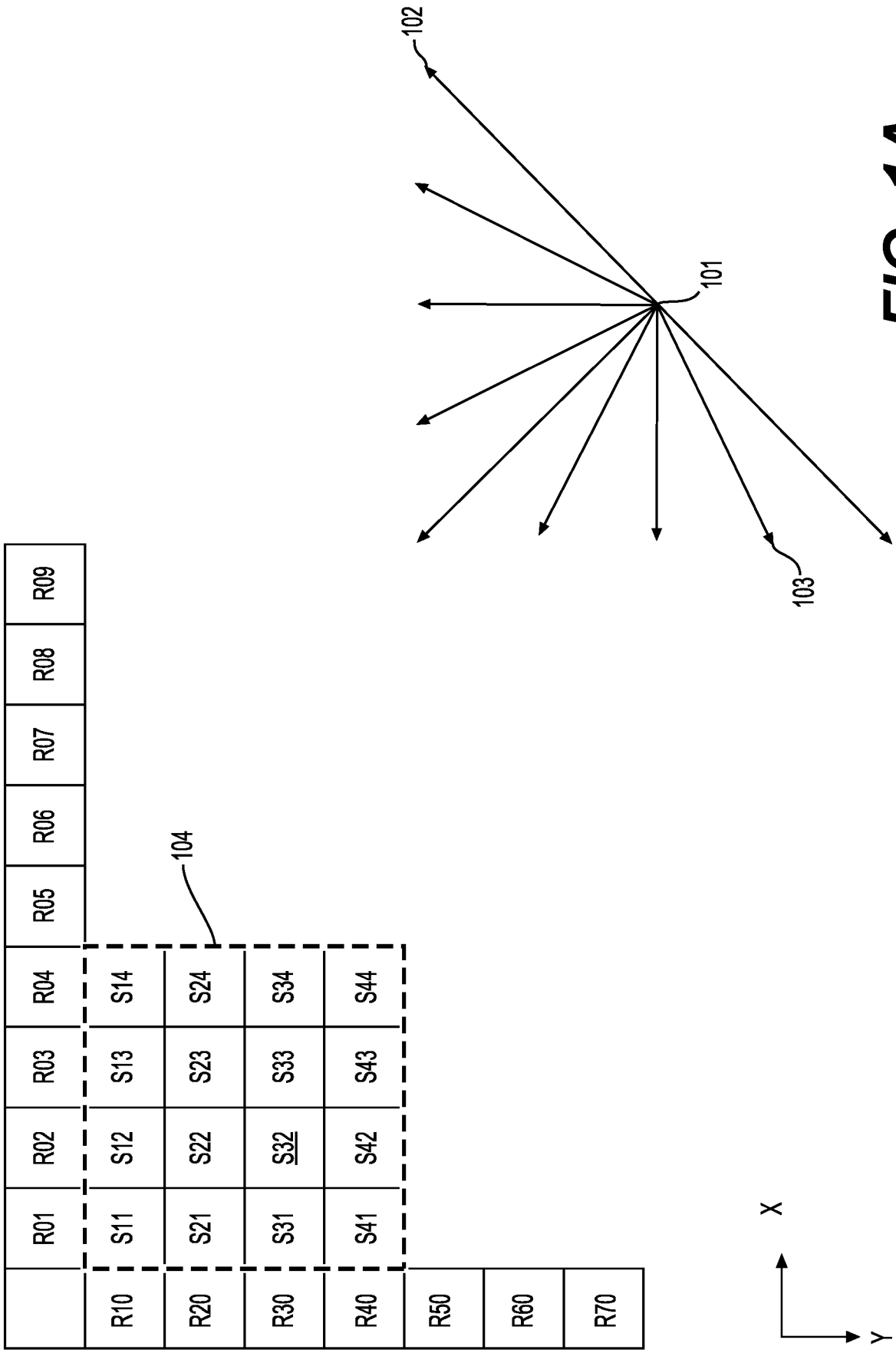

102
101
103
104
R01
R02
R03
R04
R05
R06
R07
R08
R09
R10
R20
R30
R40
R50
R60
R70
S11
S12
S13
S14
S21
S22
S23
S24
S31
S32
S33
S34
S41
S42
S43
S44
X
Y

FIG. 1A

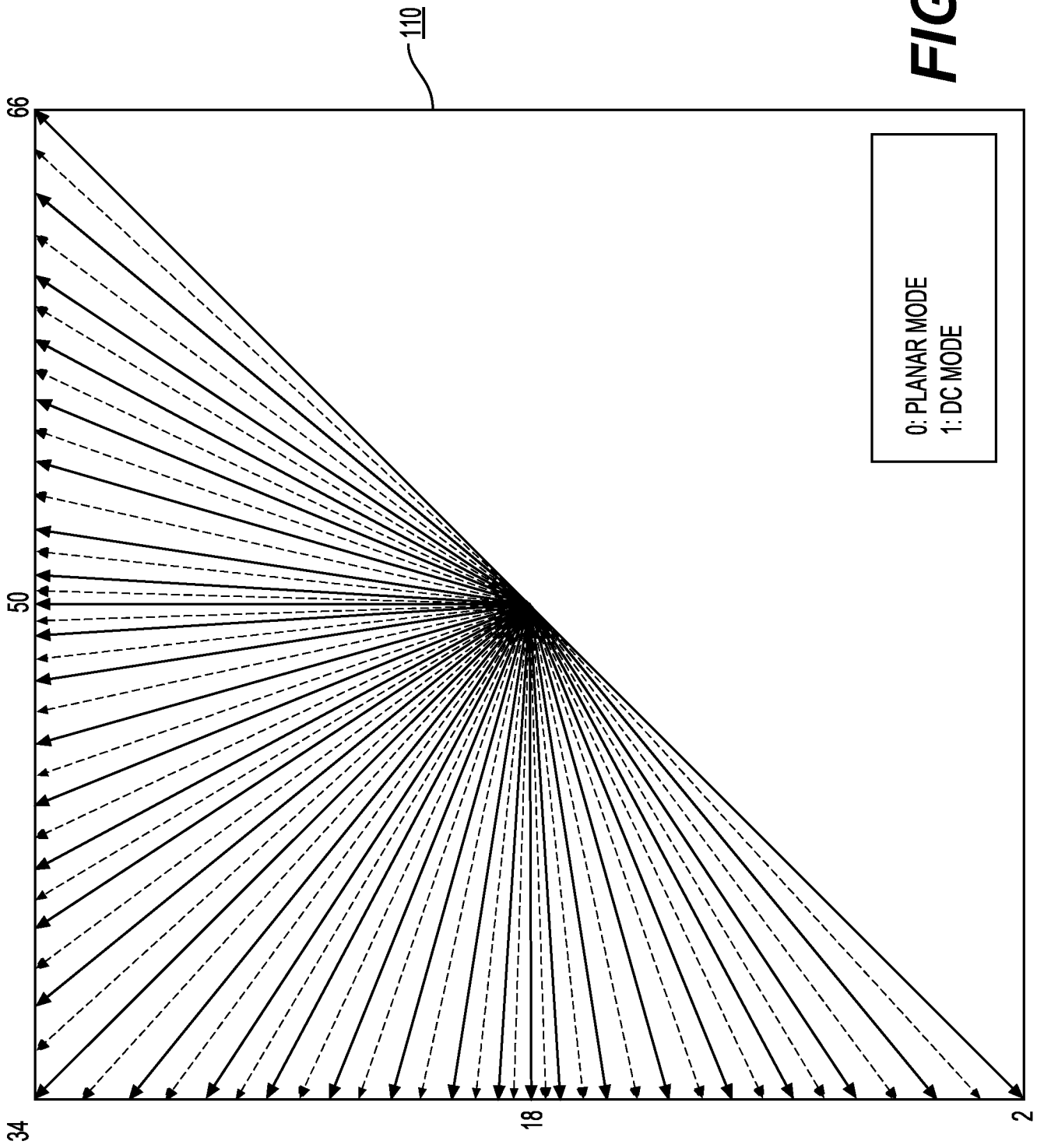


FIG. 1B

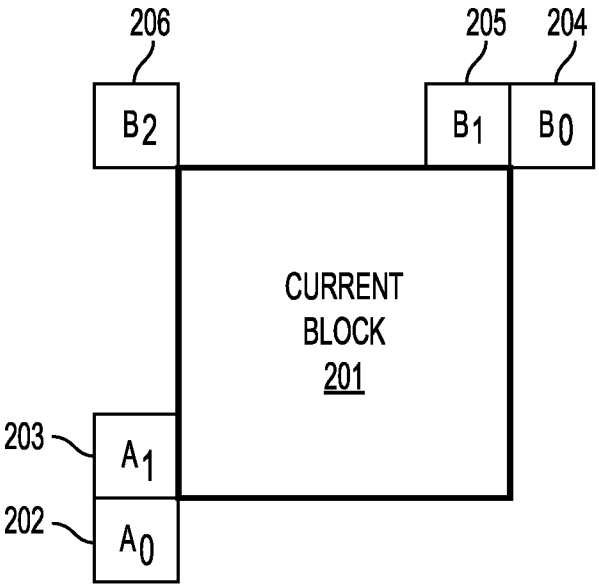


FIG. 2

(Related Art)

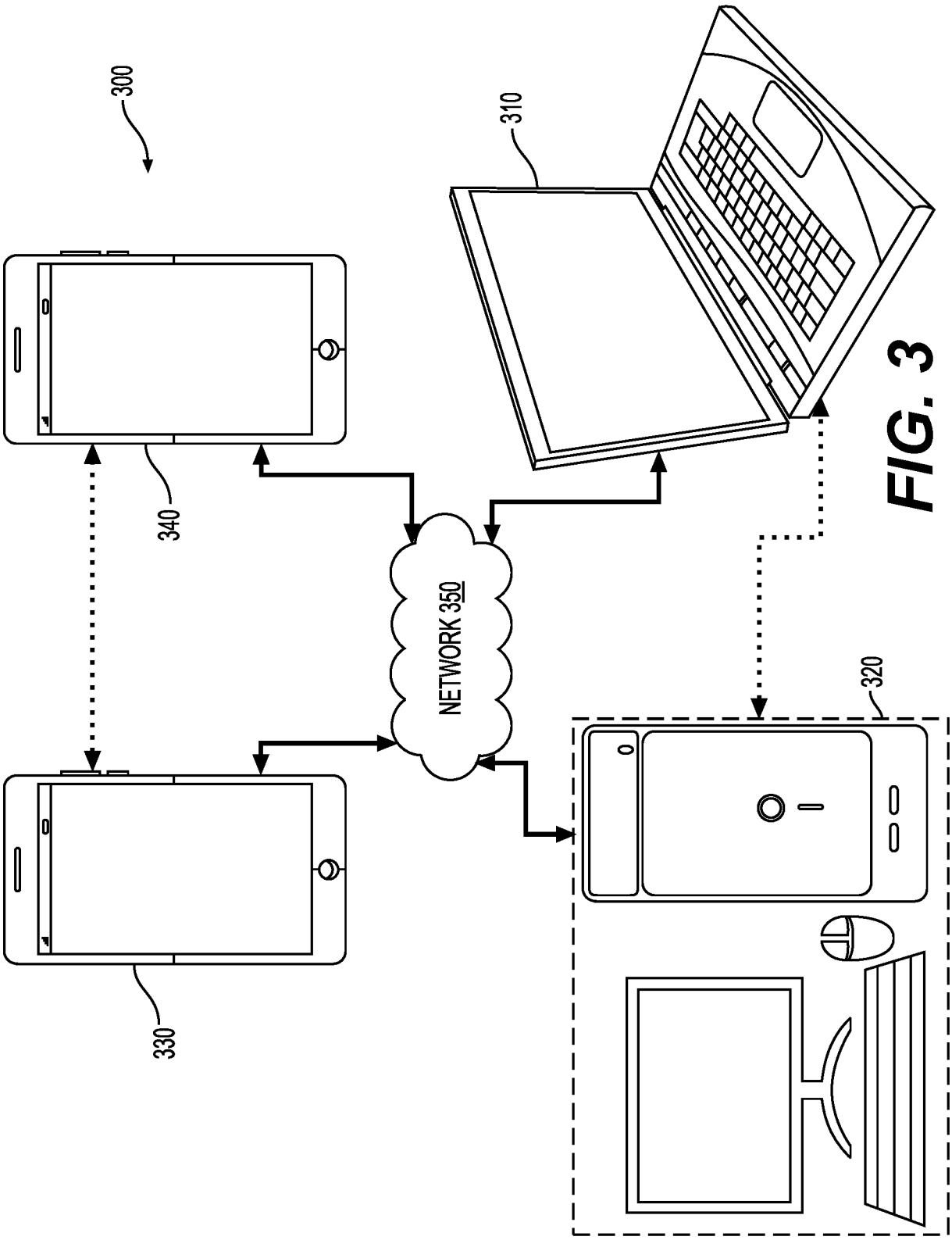


FIG. 3

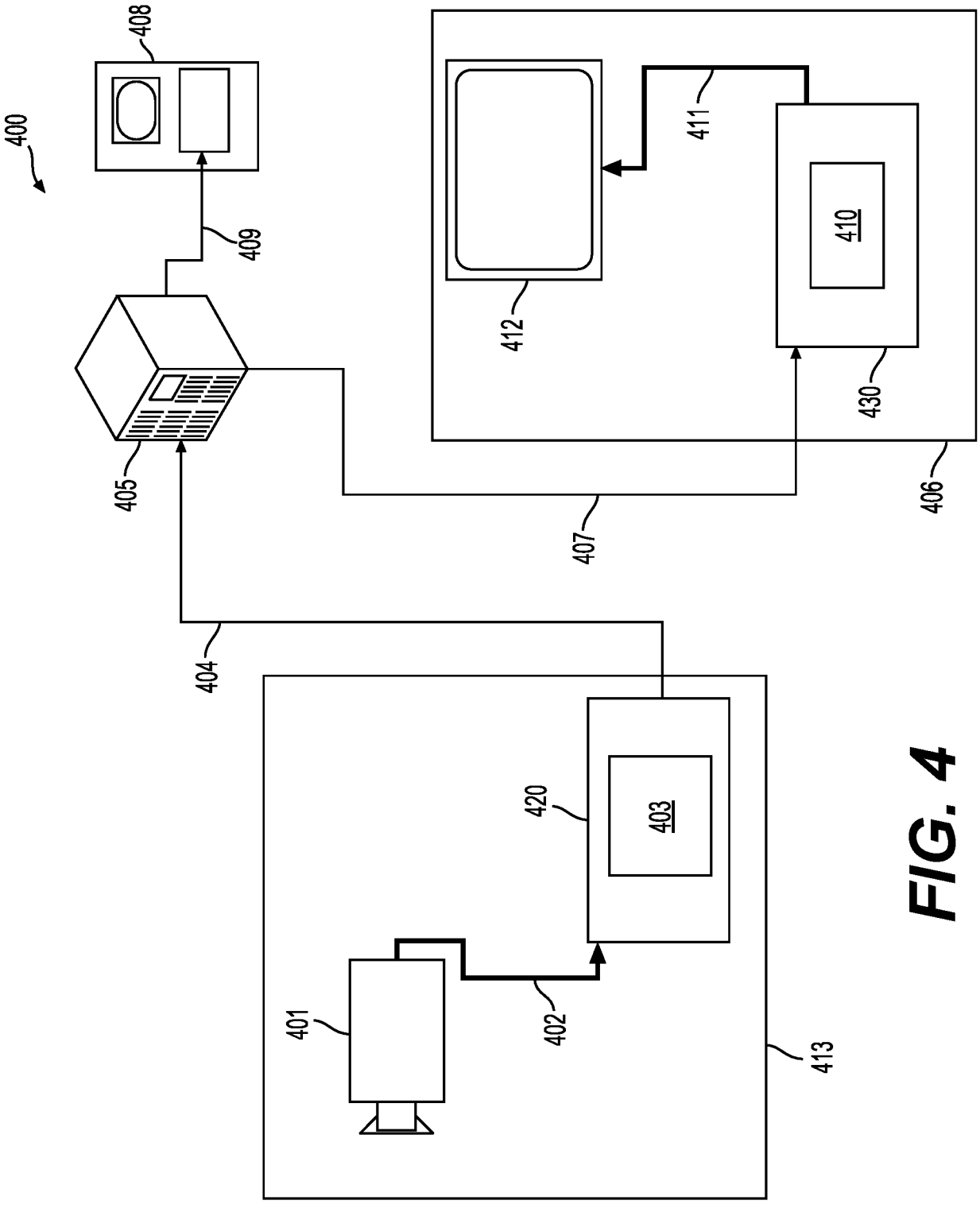


FIG. 4

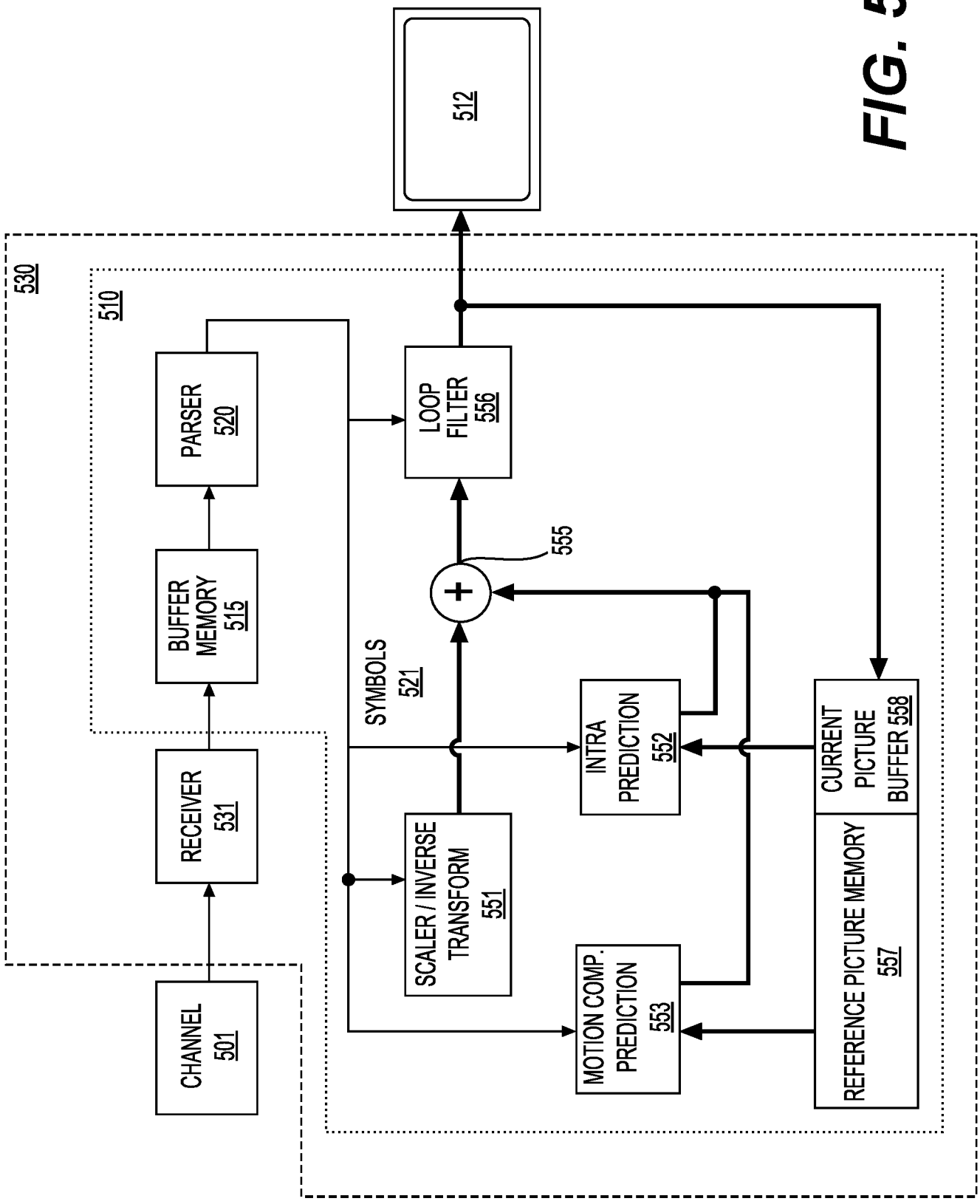

512
530
510
PARSER 520
BUFFER MEMORY 515
RECEIVER 531
CHANNEL 501
LOOP FILTER 556
555
SYMBOLS 521
SCALER / INVERSE TRANSFORM 551
INTRA PREDICTION 552
MOTION COMP. PREDICTION 553
CURRENT PICTURE BUFFER 558
REFERENCE PICTURE MEMORY 557

FIG. 5

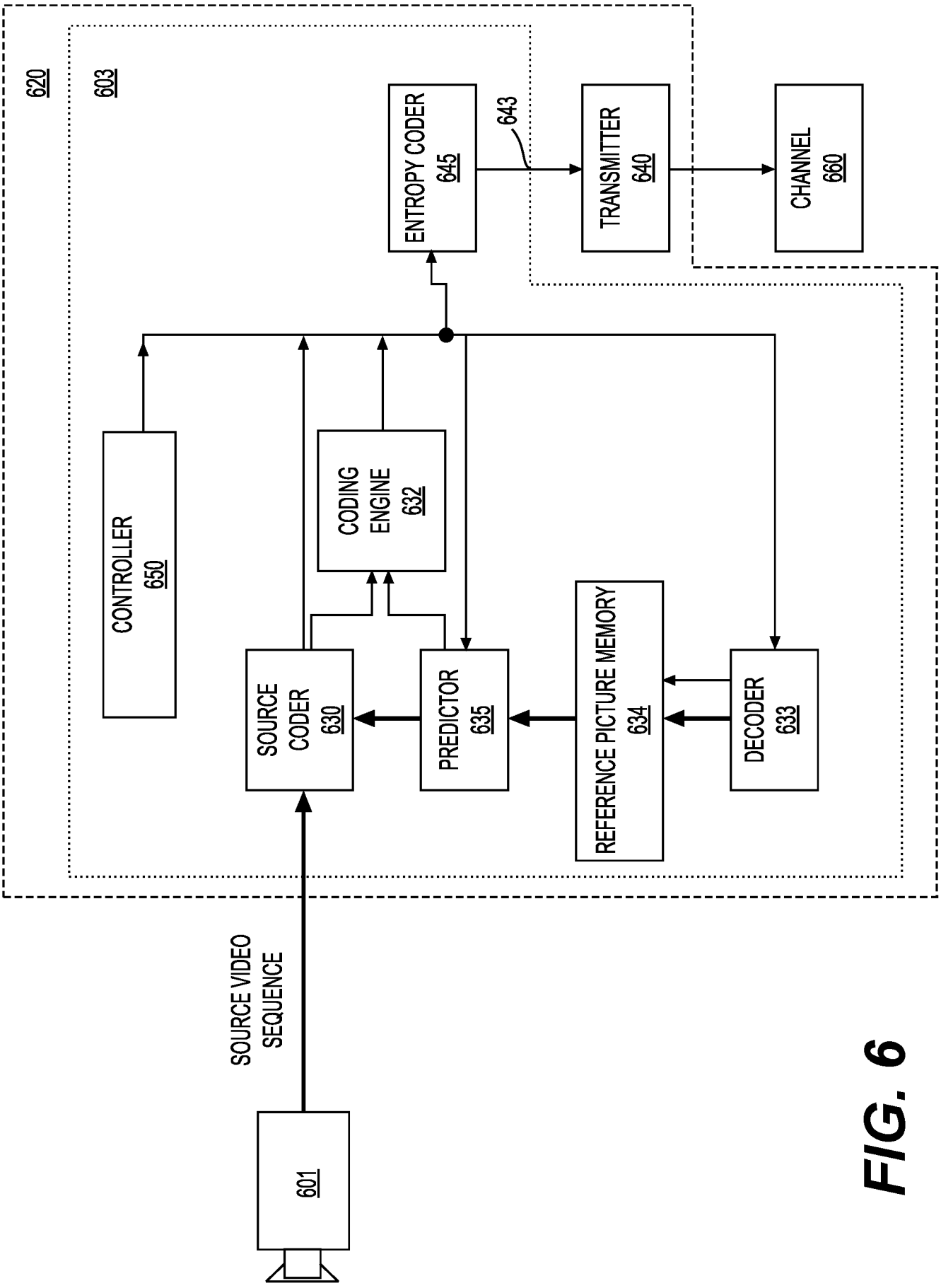

620
603
ENTROPY CODER 645
643
TRANSMITTER 640
CHANNEL 660
CONTROLLER 650
CODING ENGINE 632
SOURCE CODER 630
PREDICTOR 635
REFERENCE PICTURE MEMORY 634
DECODER 633
SOURCE VIDEO SEQUENCE
601

FIG. 6

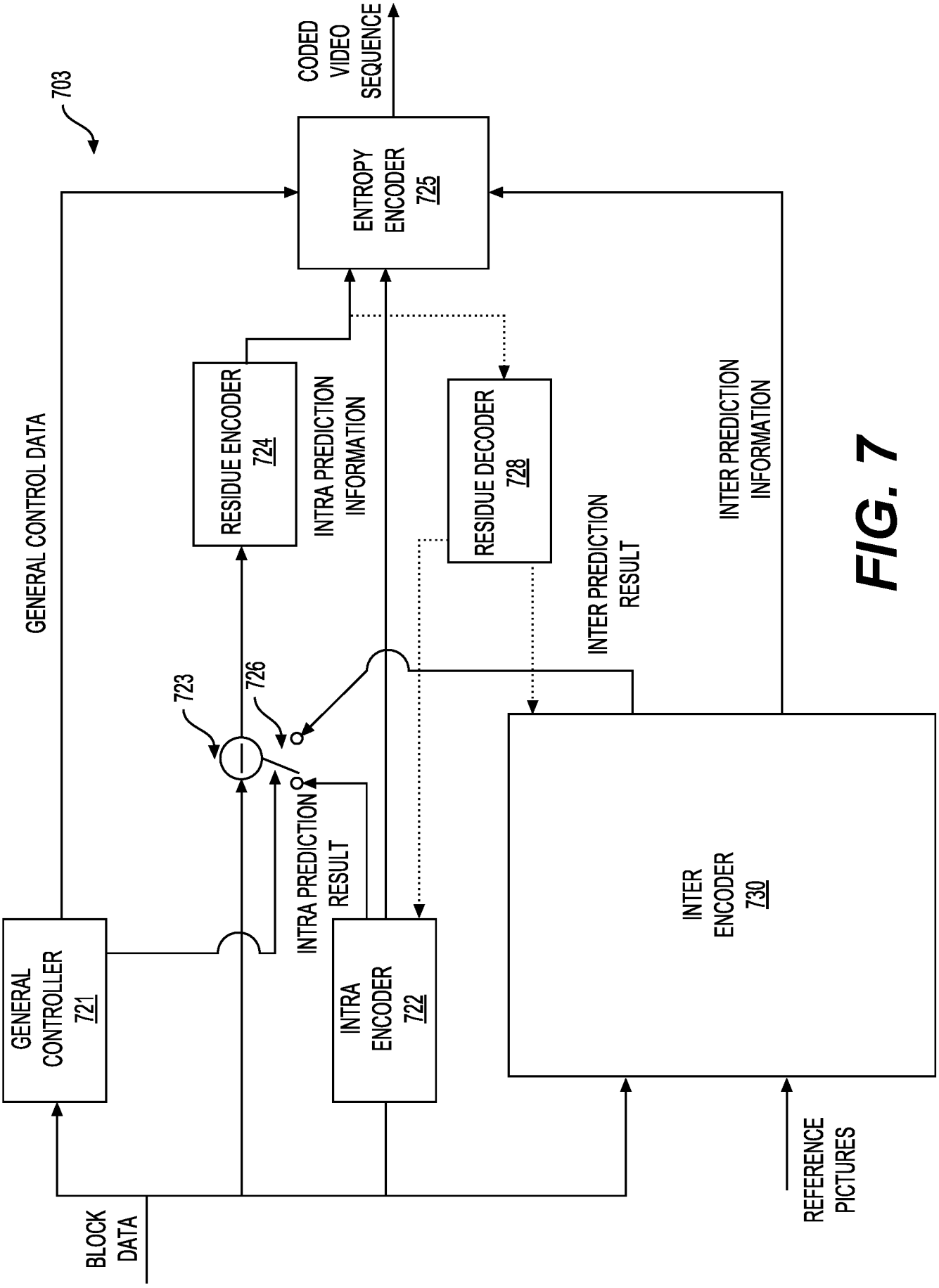

703
CODED VIDEO SEQUENCE
ENTROPY ENCODER 725
RESIDUE ENCODER 724
INTRA PREDICTION INFORMATION
RESIDUE DECODER 728
INTER PREDICTION INFORMATION
GENERAL CONTROL DATA
INTER PREDICTION RESULT
723
726
INTRA PREDICTION RESULT
GENERAL CONTROLLER 721
INTRA ENCODER 722
INTER ENCODER 730
BLOCK DATA
REFERENCE PICTURES

FIG. 7

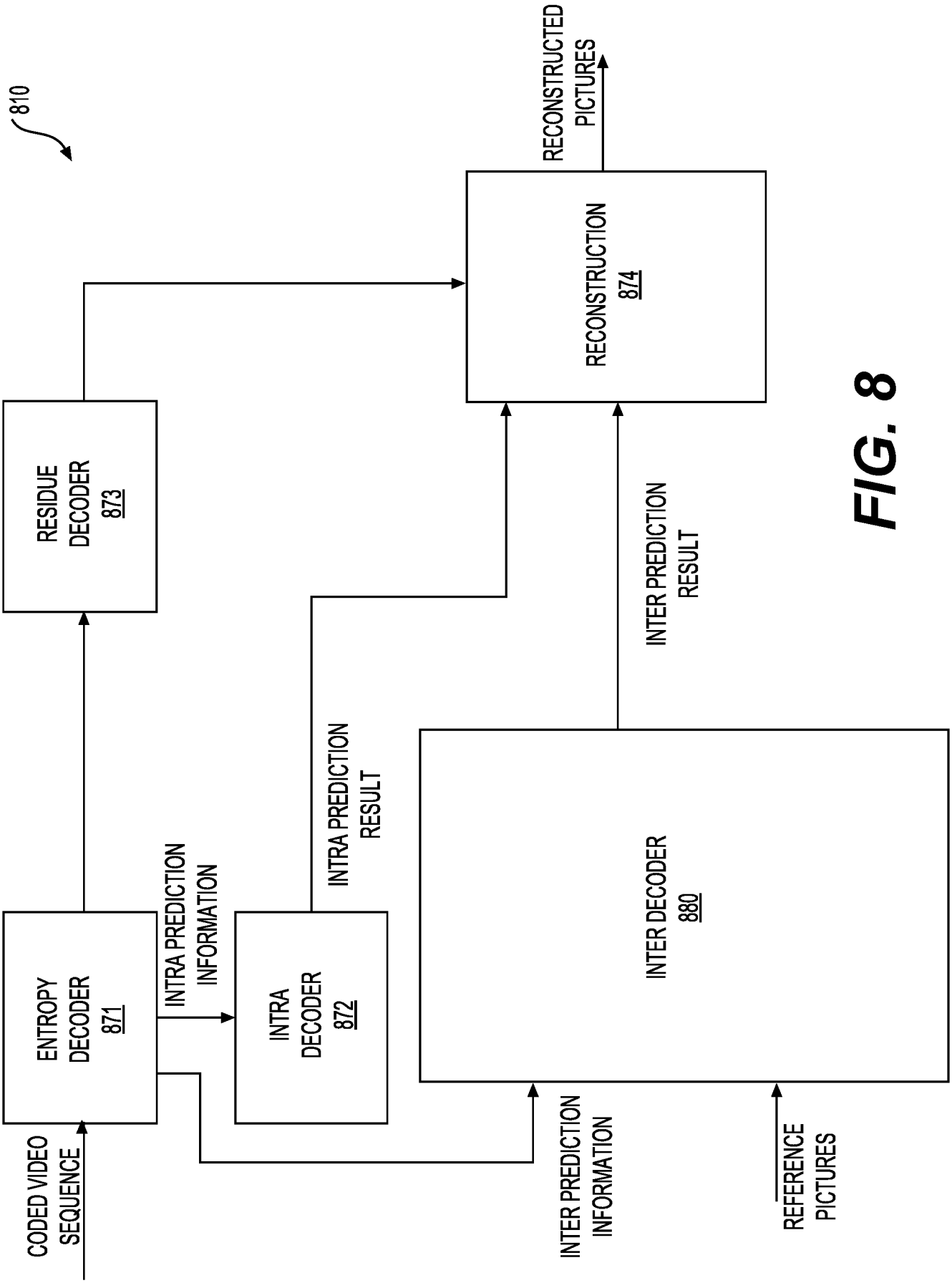

810
RECONSTRUCTED PICTURES
RECONSTRUCTION 874
RESIDUE DECODER 873
INTER PREDICTION RESULT
INTRA PREDICTION RESULT
INTER DECODER 880
ENTROPY DECODER 871
INTRA PREDICTION INFORMATION
INTRA DECODER 872
CODED VIDEO SEQUENCE
INTER PREDICTION INFORMATION
REFERENCE PICTURES

FIG. 8

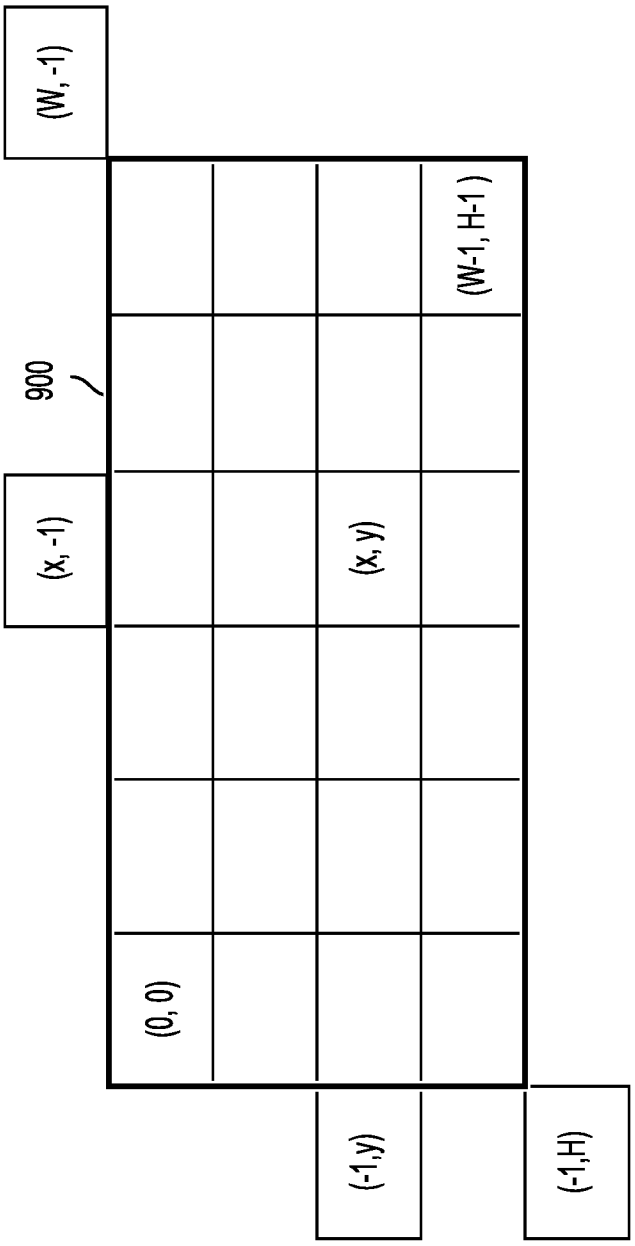


FIG. 9

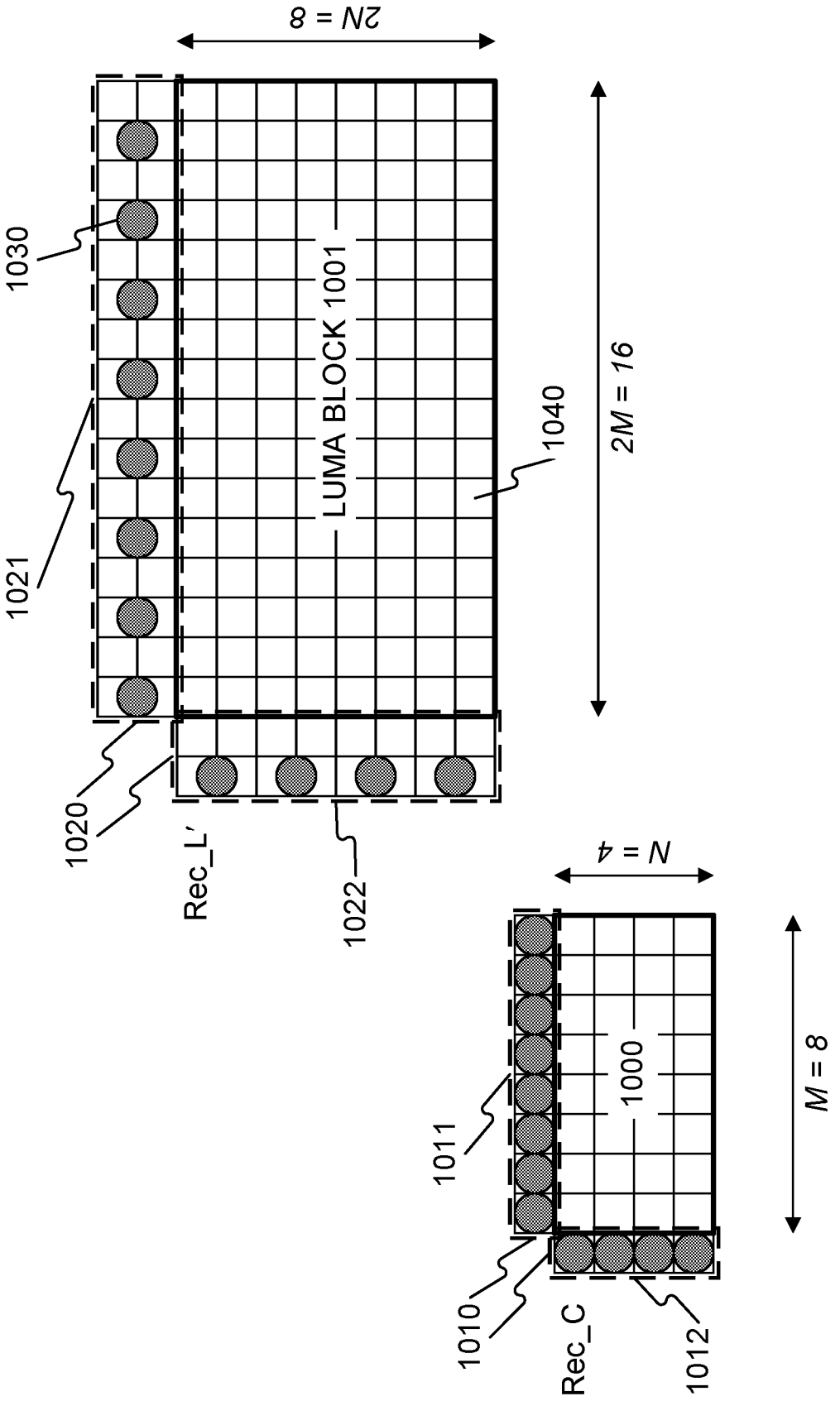

2N = 8
LUMA BLOCK 1001
1030
1040
2M = 16
1021
1020
Rec_L'
1022
N = 4
1000
1011
M = 8
1010
Rec_C
1012

FIG. 10

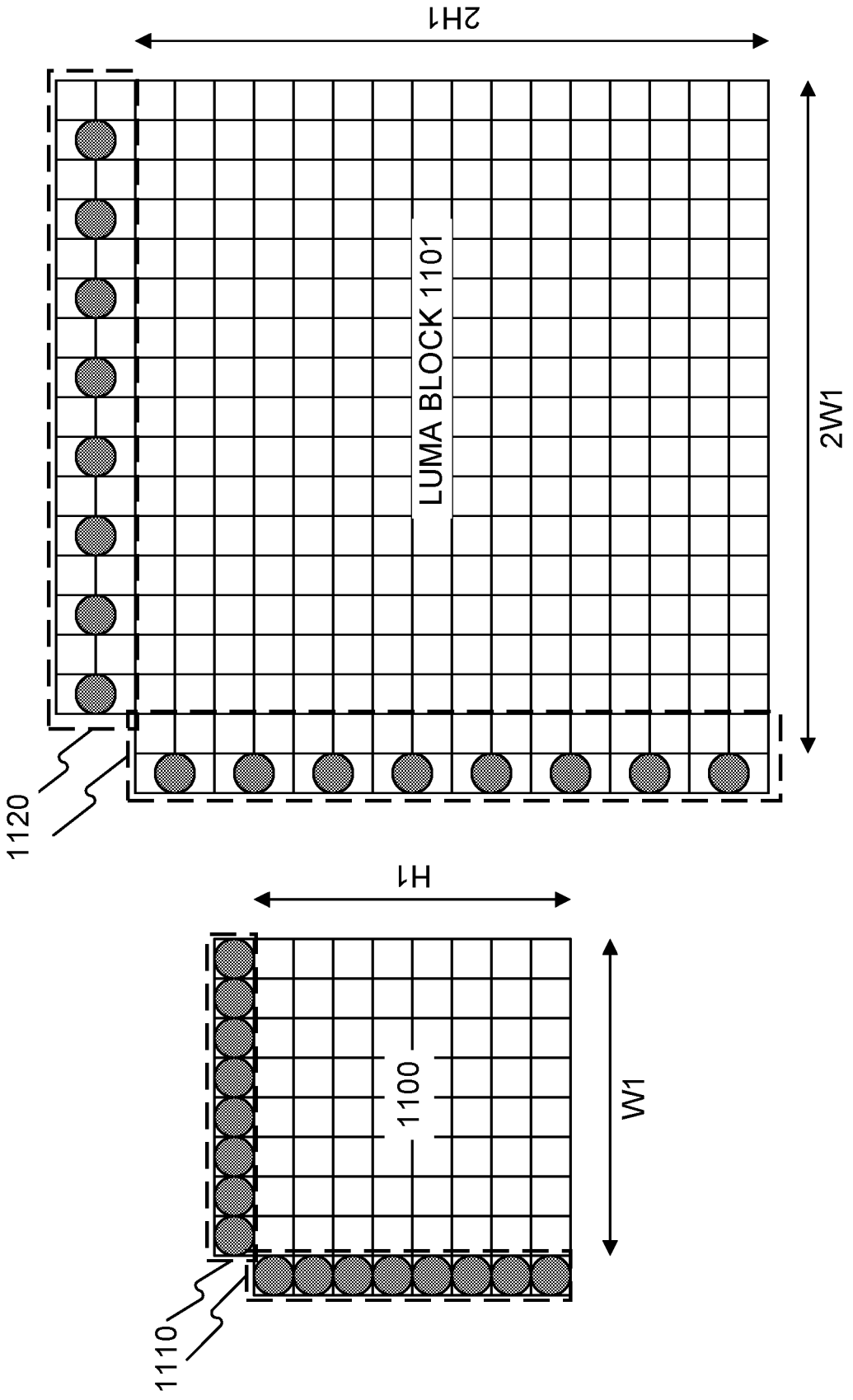


FIG. 11

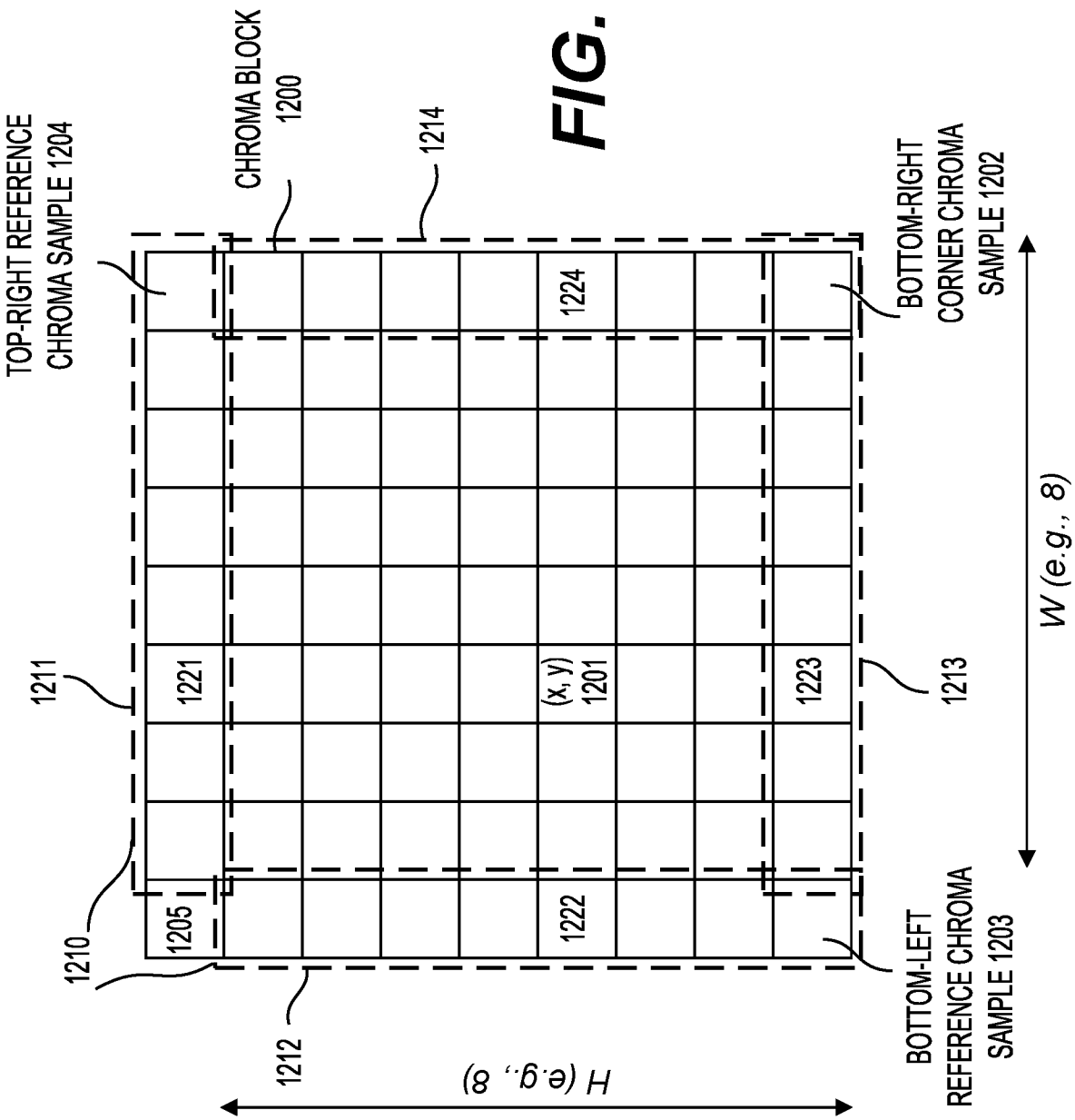

TOP-RIGHT REFERENCE CHROMA SAMPLE 1204
CHROMA BLOCK 1200
1214
BOTTOM-RIGHT CORNER CHROMA SAMPLE 1202
1224
1211
1221
(x, y) 1201
1223
1213
W (e.g., 8)
1205
1222
1210
1212
H (e.g., 8)
BOTTOM-LEFT REFERENCE CHROMA SAMPLE 1203

FIG. 12

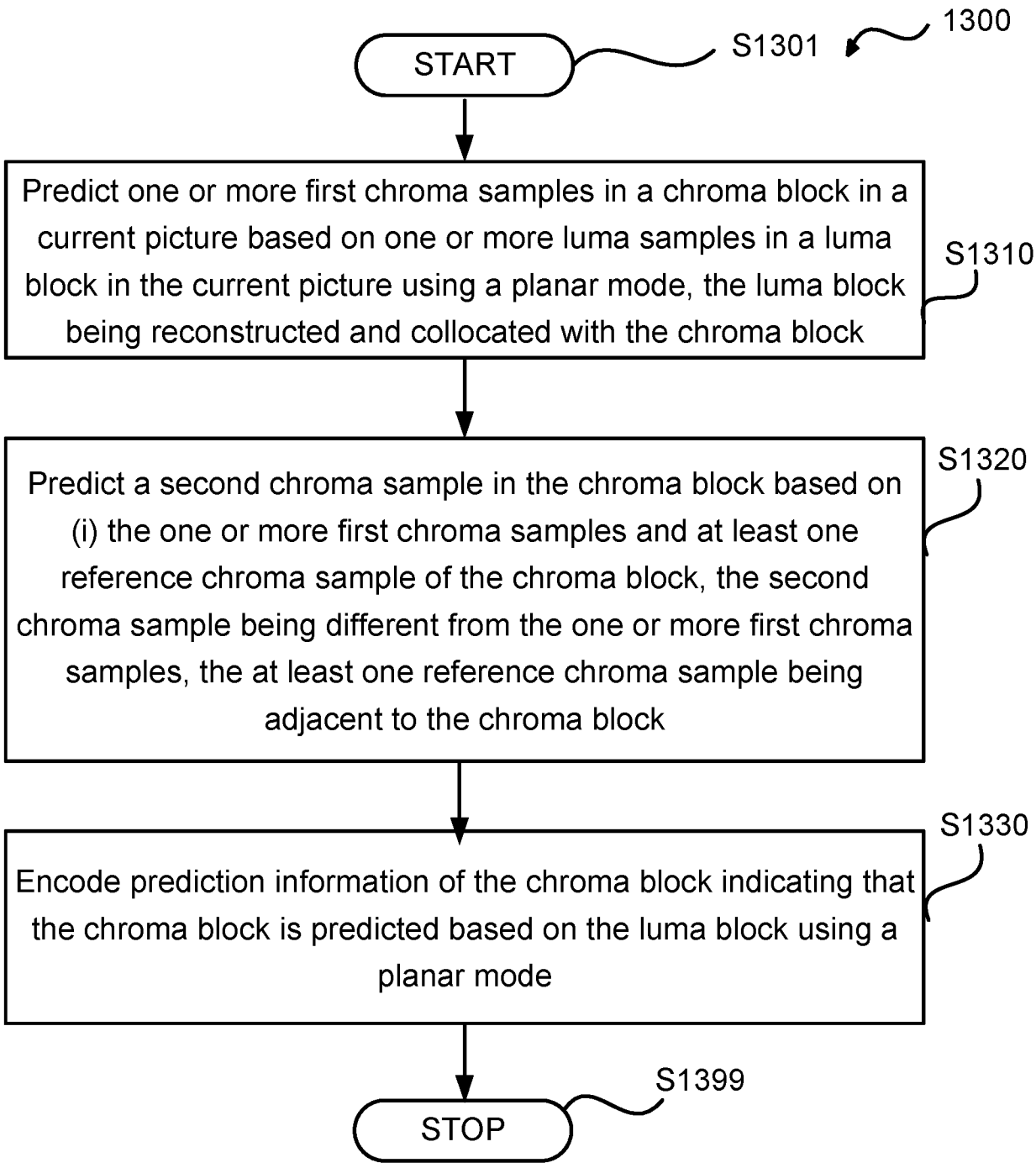

1300
START
S1301
Predict one or more first chroma samples in a chroma block in a current picture based on one or more luma samples in a luma block in the current picture using a planar mode, the luma block being reconstructed and collocated with the chroma block
S1310
Predict a second chroma sample in the chroma block based on (i) the one or more first chroma samples and at least one reference chroma sample of the chroma block, the second chroma sample being different from the one or more first chroma samples, the at least one reference chroma sample being adjacent to the chroma block
S1320
Encode prediction information of the chroma block indicating that the chroma block is predicted based on the luma block using a planar mode
S1330
STOP
S1399

FIG. 13

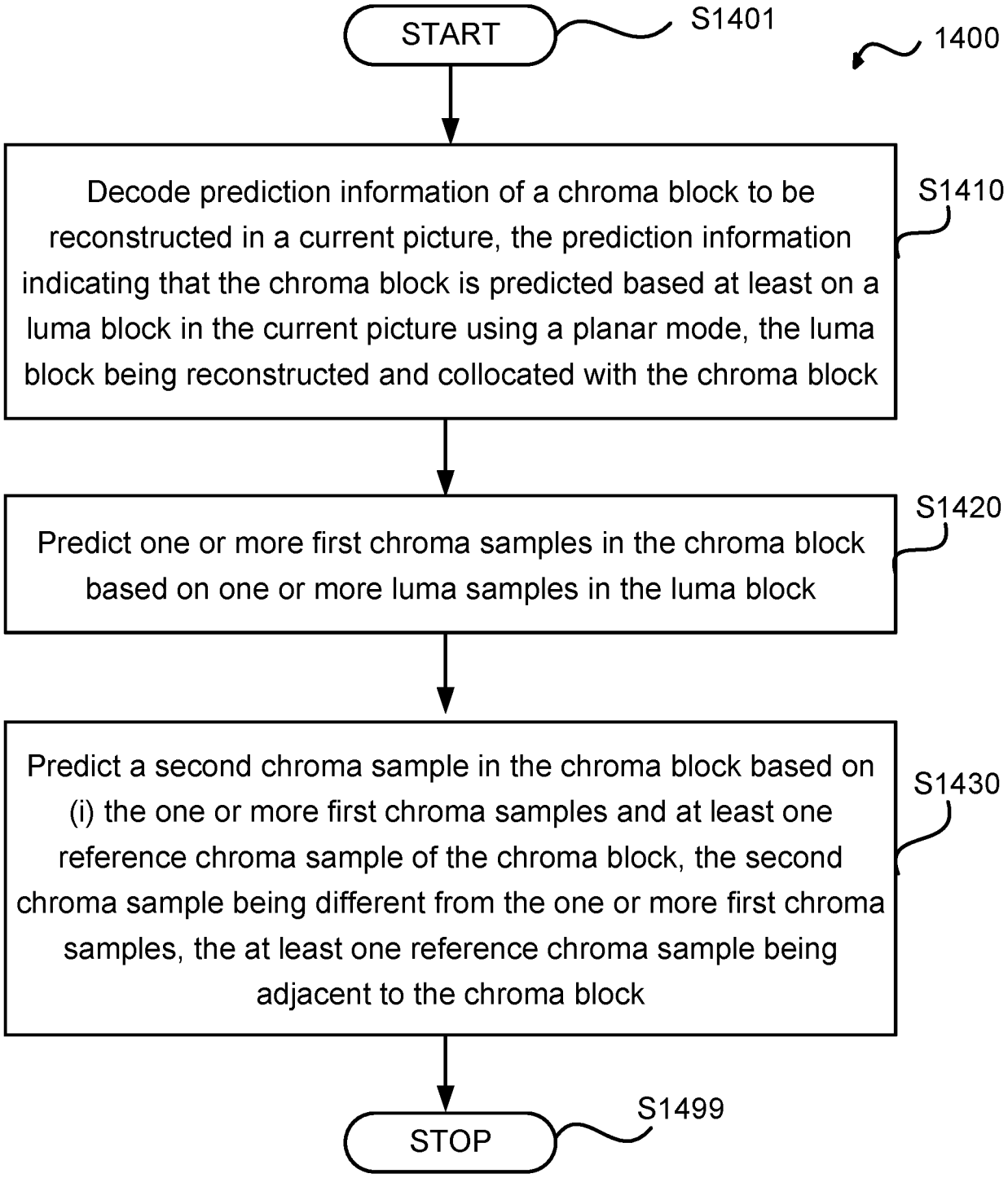

START
S1401
1400
Decode prediction information of a chroma block to be reconstructed in a current picture, the prediction information indicating that the chroma block is predicted based at least on a luma block in the current picture using a planar mode, the luma block being reconstructed and collocated with the chroma block
S1410
Predict one or more first chroma samples in the chroma block based on one or more luma samples in the luma block
S1420
Predict a second chroma sample in the chroma block based on (i) the one or more first chroma samples and at least one reference chroma sample of the chroma block, the second chroma sample being different from the one or more first chroma samples, the at least one reference chroma sample being adjacent to the chroma block
S1430
STOP
S1499

FIG. 14

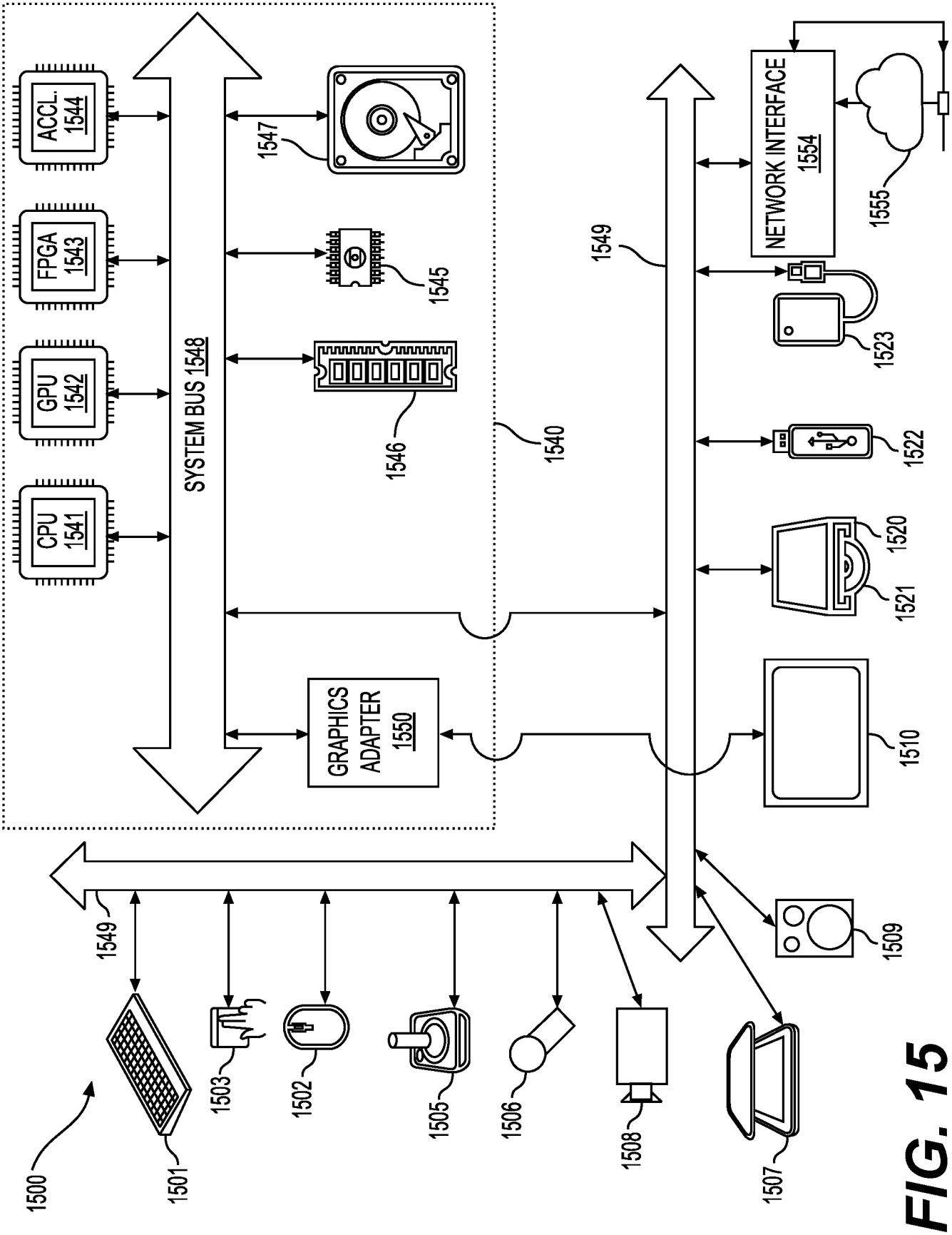

ACCL. 1544
FPGA 1543
GPU 1542
CPU 1541
SYSTEM BUS 1548
1547
1545
1546
1540
GRAPHICS ADAPTER 1550
1549
NETWORK INTERFACE 1554
1555
1523
1522
1520
1521
1510
1509
1500
1501
1503
1502
1505
1506
1508
1507

FIG. 15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US22/79154

A. CLASSIFICATION OF SUBJECT MATTER

IPC - INV. H04N 19/105; H04N 19/157; H04N 19/186 (2022.01)
ADD. H04N 19/11; H04N 19/117; H04N 19/44; H04N 19/593 (2022.01)

CPC - INV. H04N 19/105; H04N 19/157; H04N 19/186
ADD. H04N 19/11; H04N 19/117; H04N 19/44; H04N 19/593

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)
See Search History document

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

Electronic database consulted during the international search (name of database and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2021/0297656 A1 (HUAWEI TECHNOLOGIES CO., LTD.) 23 September 2021; Entire document	1-20
A	US 2021/0297694 A1 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 23 September 2021; Entire document	1-20
A	US 2017/0359595 A1 (QUALCOMM INCORPORATED) 14 December 2017; Entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☐ 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
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 January 2023 (12.01.2023)	FEB 03 2023
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Shane Thomas Telephone No. PCT Helpdesk: 571-272-4300