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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 January 2025 (16.01.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/011995 A1

(51) International Patent Classification:
H04N 19/103 (2014.01) *H04N 19/82* (2014.01)
H04N 19/186 (2014.01)

(21) International Application Number: PCT/EP2024/068438

(22) International Filing Date: 01 July 2024 (01.07.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
23306151.4 07 July 2023 (07.07.2023) EP

(71) Applicant: **INTERDIGITAL CE PATENT HOLDINGS, SAS** [FR/FR]; 3 rue du Colonel Moll, 75017 PARIS (FR).

(72) Inventors: **GALPIN, Franck**; c/o InterDigital, Immeuble ZEN 2, 845A Avenue des Champs Blancs, 35510 CESSON-SEVIGNE (FR). **BORDES, Philippe**; c/o InterDigital, Immeuble ZEN 2, 845A Avenue des Champs Blancs, 35510 CESSON-SEVIGNE (FR). **NASER, Karam**; c/o InterDigital, Immeuble ZEN 2, 845A Avenue des Champs Blancs, 35510 CESSON-SEVIGNE (FR). **LO BIANCO, Federico**; c/o InterDigital, Immeuble ZEN 2, 845A Avenue des Champs Blancs, 35510 CESSON-SEVIGNE (FR).

(74) Agent: **INTERDIGITAL**; Immeuble Zen 2, 845A Avenue des Champs Blancs, 35510 CESSON-SEVIGNE (FR).

(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, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,

(54) Title: CROSS COMPONENT RESIDUAL PREDICTION

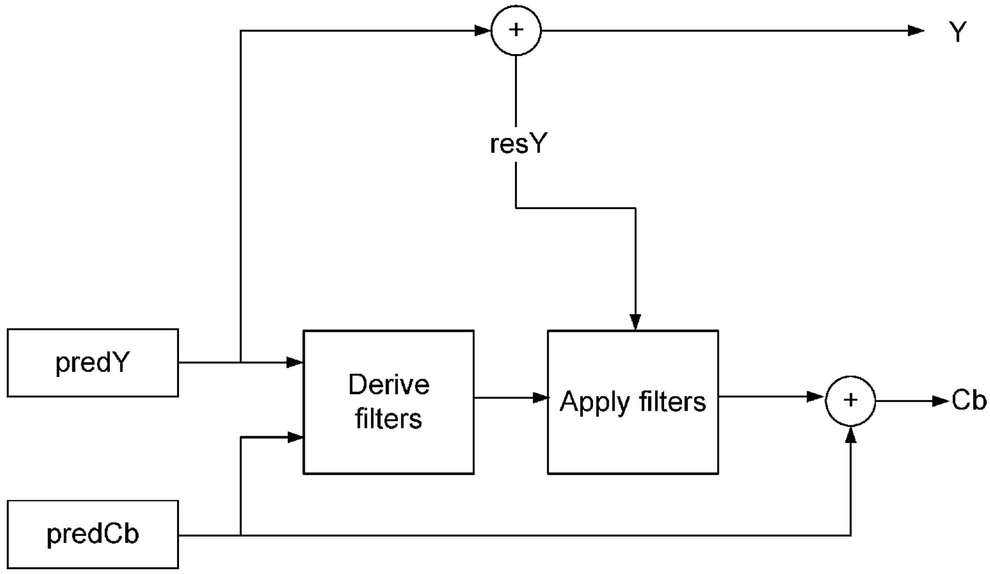


FIG. 16

(57) **Abstract:** Systems, devices, and methods are described herein for cross component residual prediction. An example method may include estimating a relationship between residual of a first chrominance component of a block and residual of a second chrominance component of the block based on a prediction of the first chrominance component and a prediction of the second chrominance component. The method may include predicting the residual of the first chrominance component based on the estimated relationship and the residual of the second chrominance component. The method may include decoding the block based on the predicted residual of the first chrominance component.

WO 2025/011995 A1

[Continued on next page]

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, SC, 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 RESIDUAL PREDICTION

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of European Provisional Patent Application No. EP23306151.4, filed July 7, 2023, the contents of which are hereby incorporated by reference herein.

BACKGROUND

[0002] Video coding systems may be used to compress digital video signals, e.g., to reduce the storage and/or transmission bandwidth needed for such signals. Video coding systems may include, for example, block-based, wavelet-based, and/or object-based systems.

SUMMARY

[0003] Systems, devices, and methods are described herein for cross component residual prediction. An example device (e.g., for video decoding (or video encoding)) may estimate a relationship between residual of a first chrominance component of a block and residual of a second chrominance component of the block based on a prediction of the first chrominance component and a prediction of the second chrominance component. The device may predict the residual of the first chrominance component based on the estimated relationship and the residual of the second chrominance component. The device may decode (or encode) the block based on the predicted residual of the first chrominance component.

[0004] The block may be a first block. The estimated relationship may be a first estimated relationship. The device may estimate a second relationship between residual of the first chrominance component of a second block and residual of a luminance component of the second block based on a prediction of the first chrominance component of the second block and a prediction of the luminance component of the second block. The device may predict the residual of the first chrominance component of the second block based on the second estimated relationship and the residual of the luminance component of the second block. The device may decode (or encode) the second block based, at least in part, on the predicted residual of the first chrominance component of the second block.

[0005] The device may predict residual of the second chrominance component of the second block based the residual of the first chrominance component of the second block and the residual of the luminance component of the second block. The device may decode (or encode) the second block based on the predicted residual of the first chrominance component of the second block and the predicted residual of the second chrominance component of the second block.

[0006] The device may compute residual of the predicted residual of the first chrominance component of the second block based on a difference between an original sample of the first chrominance component of the second block and a sum of the prediction of the first chrominance component of the second block and the predicted residual of the first chrominance component of the second block. The device may decode (or encode) the second block based on the predicted residual of the first chrominance component of the second block and the residual of the predicted residual of the first chrominance component of the second block.

[0007] The device may compute residual of the predicted residual of the first chrominance component based on a difference between an original sample of the first chrominance component and a sum of the prediction of the first chrominance component and the predicted residual of the first chrominance component. The device may decode (or encode) the second block based on the predicted residual of the first chrominance component of the second block and the residual of the predicted residual of the first chrominance component.

[0008] The first chrominance component may be a red-difference chrominance component. The second chrominance component may be a blue-difference chrominance component. The device may send a flag to an encoder (or decoder). The flag may indicate the predicted residual of the first chrominance component. An example method may involve performing one or more of the actions described above.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] FIG. 1A is a system diagram illustrating an example communications system in which one or more disclosed embodiments may be implemented.

[0010] FIG. 1B is a system diagram illustrating an example wireless transmit/receive unit (WTRU) that may be used within the communications system illustrated in FIG. 1A according to an embodiment.

[0011] FIG. 1C is a system diagram illustrating an example radio access network (RAN) and an example core network (CN) that may be used within the communications system illustrated in FIG. 1A according to an embodiment.

[0012] FIG. 1D is a system diagram illustrating a further example RAN and a further example CN that may be used within the communications system illustrated in FIG. 1A according to an embodiment.

[0013] FIG. 2 is a diagram showing an example block-based video encoder.

[0014] FIG. 3 is a diagram showing an example video decoder.

[0015] FIG. 4 is a diagram showing an example of a system in which various aspects and examples may be implemented.

[0016] FIG. 5 illustrates example locations of samples used to derive prediction model(s) (e.g., model parameters α and β).

[0017] FIG. 6 illustrates an example effect of slope adjustment.

[0018] FIG. 7 illustrates example spatial part of a convolutional filter.

[0019] FIG. 8 illustrates an example reference area used to derive filter coefficients.

[0020] FIG. 9 illustrates an example spatial part of a gradient and location based convolutional cross-component model (GL-CCCM) convolutional filter.

[0021] FIG. 10 illustrates an example spatial part of non-down sampled convolutional cross-component model (CCCM) luma terms used for a convolutional filter.

[0022] FIG. 11 illustrates example down-sampling filters.

[0023] FIG. 12 illustrates an example of inter prediction using a cross-component residual model (CCRM).

[0024] FIG. 13 illustrates example luma samples in relation to a chroma sample.

[0025] FIG. 14 illustrates an example of Cr residual prediction.

[0026] FIG. 15 illustrates an example of function support to predict Cr.

[0027] FIG. 16 illustrates an example prediction of Cb residual from Y.

[0028] FIG. 17 illustrates an example of Cr residual derivation from Y and Cb residual.

[0029] FIG. 18 illustrates example luma/Cb coefficient locations.

[0030] FIG. 19 illustrates an example of adding residual on the residual prediction.

DETAILED DESCRIPTION

[0031] A detailed description of illustrative embodiments will now be described with reference to the various Figures. Although this description provides a detailed example of possible implementations, it should be noted that the details are intended to be exemplary and in no way limit the scope of the application.

[0032] FIG. 1A is a diagram illustrating an example communications system 100 in which one or more disclosed embodiments may be implemented. The communications system 100 may be a multiple access system that provides content, such as voice, data, video, messaging, broadcast, etc., to multiple wireless users. The communications system 100 may enable multiple wireless users to access such content through the sharing of system resources, including wireless bandwidth. For example, the communications systems 100 may employ one or more channel access methods, such as code division multiple access (CDMA), time division multiple access (TDMA), frequency division multiple access (FDMA), orthogonal FDMA (OFDMA), single-carrier FDMA (SC-FDMA), zero-tail unique-word DFT-Spread OFDM (ZT UW DTS-s OFDM), unique word OFDM (UW-OFDM), resource block-filtered OFDM, filter bank multicarrier (FBMC), and the like.

[0033] As shown in FIG. 1A, the communications system 100 may include wireless transmit/receive units (WTRUs) 102a, 102b, 102c, 102d, a RAN 104/113, a CN 106/115, a public switched telephone network (PSTN) 108, the Internet 110, and other networks 112, though it will be appreciated that the disclosed embodiments contemplate any number of WTRUs, base stations, networks, and/or network elements. Each of the WTRUs 102a, 102b, 102c, 102d may be any type of device configured to operate and/or communicate in a wireless environment. By way of example, the WTRUs 102a, 102b, 102c, 102d, any of which may be referred to as a “station” and/or a “STA”, may be configured to transmit and/or receive wireless signals and may include a user equipment (UE), a mobile station, a fixed or mobile subscriber unit, a subscription-based unit, a pager, a cellular telephone, a personal digital assistant (PDA), a smartphone, a laptop, a netbook, a personal computer, a wireless sensor, a hotspot or Mi-Fi device, an Internet of Things (IoT) device, a watch or other wearable, a head-mounted display (HMD), a vehicle, a drone, a medical device and applications (e.g., remote surgery), an industrial device and applications (e.g., a robot and/or other wireless devices operating in an industrial and/or an automated processing chain contexts), a consumer electronics device, a device operating on commercial and/or industrial wireless networks, and the like. Any of the WTRUs 102a, 102b, 102c and 102d may be interchangeably referred to as a UE.

[0034] The communications systems 100 may also include a base station 114a and/or a base station 114b. Each of the base stations 114a, 114b may be any type of device configured to wirelessly interface with at least one of the WTRUs 102a, 102b, 102c, 102d to facilitate access to one or more communication networks, such as the CN 106/115, the Internet 110, and/or the other networks 112. By way of example, the base stations 114a, 114b may be a base transceiver station (BTS), a Node-B, an eNode B, a Home Node B, a Home eNode B, a gNB, a NR NodeB, a site controller, an access point (AP), a wireless router, and the like. While the base stations 114a, 114b are each depicted as a single element, it will be appreciated that the base stations 114a, 114b may include any number of interconnected base stations and/or network elements.

[0035] The base station 114a may be part of the RAN 104/113, which may also include other base stations and/or network elements (not shown), such as a base station controller (BSC), a radio network controller (RNC), relay nodes, etc. The base station 114a and/or the base station 114b may be configured to transmit and/or receive wireless signals on one or more carrier frequencies, which may be referred to as a cell (not shown). These frequencies may be in licensed spectrum, unlicensed spectrum, or a combination of licensed and unlicensed spectrum. A cell may provide coverage for a wireless service to a specific geographical area that may be relatively fixed or that may change over time. The cell may further be divided into cell sectors. For example, the cell associated with the base station 114a may be divided into three sectors. Thus, in one embodiment, the base station 114a may include three transceivers, i.e., one for each sector of the cell. In an embodiment, the base station 114a may employ multiple-input multiple output (MIMO) technology and may utilize multiple transceivers for each sector of the cell. For example, beamforming may be used to transmit and/or receive signals in desired spatial directions.

[0036] The base stations 114a, 114b may communicate with one or more of the WTRUs 102a, 102b, 102c, 102d over an air interface 116, which may be any suitable wireless communication link (e.g., radio frequency (RF), microwave, centimeter wave, micrometer wave, infrared (IR), ultraviolet (UV), visible light, etc.). The air interface 116 may be established using any suitable radio access technology (RAT).

[0037] More specifically, as noted above, the communications system 100 may be a multiple access system and may employ one or more channel access schemes, such as CDMA, TDMA, FDMA, OFDMA, SC-FDMA, and the like. For example, the base station 114a in the RAN 104/113 and the WTRUs 102a, 102b, 102c may implement a radio technology such as Universal Mobile Telecommunications System (UMTS) Terrestrial Radio Access (UTRA), which may establish the air interface 115/116/117 using wideband CDMA (WCDMA). WCDMA may include communication protocols such as High-Speed Packet Access (HSPA) and/or Evolved HSPA (HSPA+). HSPA may include High-Speed Downlink (DL) Packet Access (HSDPA) and/or High-Speed UL Packet Access (HSUPA).

[0038] In an embodiment, the base station 114a and the WTRUs 102a, 102b, 102c may implement a radio technology such as Evolved UMTS Terrestrial Radio Access (E-UTRA), which may establish the air interface 116 using Long Term Evolution (LTE) and/or LTE-Advanced (LTE-A) and/or LTE-Advanced Pro (LTE-A Pro).

[0039] In an embodiment, the base station 114a and the WTRUs 102a, 102b, 102c may implement a radio technology such as NR Radio Access , which may establish the air interface 116 using New Radio (NR).

[0040] In an embodiment, the base station 114a and the WTRUs 102a, 102b, 102c may implement multiple radio access technologies. For example, the base station 114a and the WTRUs 102a, 102b, 102c may implement LTE radio access and NR radio access together, for instance using dual connectivity (DC) principles. Thus, the air interface utilized by WTRUs 102a, 102b, 102c may be characterized by multiple types of radio access technologies and/or transmissions sent to/from multiple types of base stations (e.g., a eNB and a gNB).

[0041] In other embodiments, the base station 114a and the WTRUs 102a, 102b, 102c may implement radio technologies such as IEEE 802.11 (i.e., Wireless Fidelity (WiFi), IEEE 802.16 (i.e., Worldwide Interoperability for Microwave Access (WiMAX)), CDMA2000, CDMA2000 1X, CDMA2000 EV-DO, Interim Standard 2000 (IS-2000), Interim Standard 95 (IS-95), Interim Standard 856 (IS-856), Global System for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE), GSM EDGE (GERAN), and the like.

[0042] The base station 114b in FIG. 1A may be a wireless router, Home Node B, Home eNode B, or access point, for example, and may utilize any suitable RAT for facilitating wireless connectivity in a localized area, such as a place of business, a home, a vehicle, a campus, an industrial facility, an air corridor (e.g., for use by drones), a roadway, and the like. In one embodiment, the base station 114b and the WTRUs 102c, 102d may implement a radio technology such as IEEE 802.11 to establish a wireless local area network (WLAN). In an embodiment, the base station 114b and the WTRUs 102c, 102d may implement a radio technology such as IEEE 802.15 to establish a wireless personal area network (WPAN). In yet another embodiment, the base station 114b and the WTRUs 102c, 102d may utilize a cellular-based RAT (e.g., WCDMA, CDMA2000, GSM, LTE, LTE-A, LTE-A Pro, NR etc.) to establish a picocell or femtocell. As shown in FIG. 1A, the base station 114b may have a direct connection to the Internet 110. Thus, the base station 114b may not be required to access the Internet 110 via the CN 106/115.

[0043] The RAN 104/113 may be in communication with the CN 106/115, which may be any type of network configured to provide voice, data, applications, and/or voice over internet protocol (VoIP) services to one or more of the WTRUs 102a, 102b, 102c, 102d. The data may have varying quality of service (QoS) requirements, such as differing throughput requirements, latency requirements, error tolerance requirements, reliability requirements, data throughput requirements, mobility requirements, and the like. The CN 106/115 may provide call control, billing services, mobile location-based services, pre-paid calling, Internet connectivity, video distribution, etc., and/or perform high-level security functions, such as user authentication. Although not shown in FIG. 1A, it will be appreciated that the RAN 104/113 and/or the CN 106/115 may be in direct or indirect communication with other RANs that employ the same RAT as the RAN 104/113 or a different RAT. For example, in addition to being connected to the RAN 104/113, which may be utilizing a NR radio technology, the CN 106/115 may also be in communication with another RAN (not shown) employing a GSM, UMTS, CDMA 2000, WiMAX, E-UTRA, or WiFi radio technology.

[0044] The CN 106/115 may also serve as a gateway for the WTRUs 102a, 102b, 102c, 102d to access the PSTN 108, the Internet 110, and/or the other networks 112. The PSTN 108 may include circuit-switched telephone networks that provide plain old telephone service (POTS). The Internet 110 may include a global system of interconnected computer networks and devices that use common communication protocols, such as the transmission control protocol (TCP), user datagram protocol (UDP) and/or the internet protocol (IP) in the TCP/IP internet protocol suite. The networks 112 may include wired and/or wireless communications networks owned and/or operated by other service providers. For example, the networks 112 may include another CN connected to one or more RANs, which may employ the same RAT as the RAN 104/113 or a different RAT.

[0045] Some or all of the WTRUs 102a, 102b, 102c, 102d in the communications system 100 may include multi-mode capabilities (e.g., the WTRUs 102a, 102b, 102c, 102d may include multiple transceivers for communicating with different wireless networks over different wireless links). For example, the WTRU 102c shown in FIG. 1A may be configured to communicate with the base station 114a, which may employ a cellular-based radio technology, and with the base station 114b, which may employ an IEEE 802 radio technology.

[0046] FIG. 1B is a system diagram illustrating an example WTRU 102. As shown in FIG. 1B, the WTRU 102 may include a processor 118, a transceiver 120, a transmit/receive element 122, a speaker/microphone 124, a keypad 126, a display/touchpad 128, non-removable memory 130, removable memory 132, a power source 134, a global positioning system (GPS) chipset 136, and/or other peripherals 138, among others. It will be appreciated that the WTRU 102 may include any sub-combination of the foregoing elements while remaining consistent with an embodiment.

[0047] The processor 118 may be a general purpose processor, a special purpose processor, a conventional processor, a digital signal processor (DSP), a plurality of microprocessors, one or more microprocessors in association with a DSP core, a controller, a microcontroller, Application Specific Integrated Circuits (ASICs), Field Programmable Gate Arrays (FPGAs) circuits, any other type of integrated circuit (IC), a state machine, and the like. The processor 118 may perform signal coding, data processing, power control, input/output processing, and/or any other functionality that enables the WTRU 102 to operate in a wireless environment. The processor 118 may be coupled to the transceiver 120, which may be coupled to the transmit/receive element 122. While FIG. 1B depicts the processor

118 and the transceiver 120 as separate components, it will be appreciated that the processor 118 and the transceiver 120 may be integrated together in an electronic package or chip.

[0048] The transmit/receive element 122 may be configured to transmit signals to, or receive signals from, a base station (e.g., the base station 114a) over the air interface 116. For example, in one embodiment, the transmit/receive element 122 may be an antenna configured to transmit and/or receive RF signals. In an embodiment, the transmit/receive element 122 may be an emitter/detector configured to transmit and/or receive IR, UV, or visible light signals, for example. In yet another embodiment, the transmit/receive element 122 may be configured to transmit and/or receive both RF and light signals. It will be appreciated that the transmit/receive element 122 may be configured to transmit and/or receive any combination of wireless signals.

[0049] Although the transmit/receive element 122 is depicted in FIG. 1B as a single element, the WTRU 102 may include any number of transmit/receive elements 122. More specifically, the WTRU 102 may employ MIMO technology. Thus, in one embodiment, the WTRU 102 may include two or more transmit/receive elements 122 (e.g., multiple antennas) for transmitting and receiving wireless signals over the air interface 116.

[0050] The transceiver 120 may be configured to modulate the signals that are to be transmitted by the transmit/receive element 122 and to demodulate the signals that are received by the transmit/receive element 122. As noted above, the WTRU 102 may have multi-mode capabilities. Thus, the transceiver 120 may include multiple transceivers for enabling the WTRU 102 to communicate via multiple RATs, such as NR and IEEE 802.11, for example.

[0051] The processor 118 of the WTRU 102 may be coupled to, and may receive user input data from, the speaker/microphone 124, the keypad 126, and/or the display/touchpad 128 (e.g., a liquid crystal display (LCD) display unit or organic light-emitting diode (OLED) display unit). The processor 118 may also output user data to the speaker/microphone 124, the keypad 126, and/or the display/touchpad 128. In addition, the processor 118 may access information from, and store data in, any type of suitable memory, such as the non-removable memory 130 and/or the removable memory 132. The non-removable memory 130 may include random-access memory (RAM), read-only memory (ROM), a hard disk, or any other type of memory storage device. The removable memory 132 may include a subscriber identity module (SIM) card, a memory stick, a secure digital (SD) memory card, and the like. In other embodiments, the processor 118 may access information from, and store data in, memory that is not physically located on the WTRU 102, such as on a server or a home computer (not shown).

[0052] The processor 118 may receive power from the power source 134, and may be configured to distribute and/or control the power to the other components in the WTRU 102. The power source 134 may be any suitable device for powering the WTRU 102. For example, the power source 134 may include one or more dry cell batteries (e.g., nickel-cadmium (NiCd), nickel-zinc (NiZn), nickel metal hydride (NiMH), lithium-ion (Li-ion), etc.), solar cells, fuel cells, and the like.

[0053] The processor 118 may also be coupled to the GPS chipset 136, which may be configured to provide location information (e.g., longitude and latitude) regarding the current location of the WTRU 102. In addition to, or in lieu of, the information from the GPS chipset 136, the WTRU 102 may receive location information over the air interface 116 from a base station (e.g., base stations 114a, 114b) and/or determine its location based on the timing of the signals being received from two or more nearby base stations. It will be appreciated that the WTRU 102 may acquire location information by way of any suitable location-determination method while remaining consistent with an embodiment.

[0054] The processor 118 may further be coupled to other peripherals 138, which may include one or more software and/or hardware modules that provide additional features, functionality and/or wired or wireless connectivity. For example, the peripherals 138 may include an accelerometer, an e-compass, a satellite transceiver, a digital camera (for photographs and/or video), a universal serial bus (USB) port, a vibration device, a television transceiver, a hands free headset, a Bluetooth® module, a frequency modulated (FM) radio unit, a digital music player, a media player, a video game player module, an Internet browser, a Virtual Reality and/or Augmented Reality (VR/AR) device, an activity tracker, and the like. The peripherals 138 may include one or more sensors, the sensors may be one or more of a gyroscope, an accelerometer, a hall effect sensor, a magnetometer, an orientation sensor, a proximity sensor, a temperature sensor, a time sensor; a geolocation sensor; an altimeter, a light sensor, a touch sensor, a magnetometer, a barometer, a gesture sensor, a biometric sensor, and/or a humidity sensor.

[0055] The WTRU 102 may include a full duplex radio for which transmission and reception of some or all of the signals (e.g., associated with particular subframes for both the UL (e.g., for transmission) and downlink (e.g., for reception) may be concurrent and/or simultaneous. The full duplex radio may include an interference management unit to reduce and or substantially eliminate self-interference via either hardware (e.g., a choke) or signal processing via a processor (e.g., a separate processor (not shown) or via processor 118). In an embodiment, the WRTU 102 may include a half-duplex radio for which transmission and reception of some or all of the signals (e.g., associated with particular subframes for either the UL (e.g., for transmission) or the downlink (e.g., for reception)).

[0056] FIG. 1C is a system diagram illustrating the RAN 104 and the CN 106 according to an embodiment. As noted above, the RAN 104 may employ an E-UTRA radio technology to communicate with the WTRUs 102a, 102b, 102c over the air interface 116. The RAN 104 may also be in communication with the CN 106.

[0057] The RAN 104 may include eNode-Bs 160a, 160b, 160c, though it will be appreciated that the RAN 104 may include any number of eNode-Bs while remaining consistent with an embodiment. The eNode-Bs 160a, 160b, 160c may each include one or more transceivers for communicating with the WTRUs 102a, 102b, 102c over the air interface 116. In one embodiment, the eNode-Bs 160a, 160b, 160c may implement MIMO technology. Thus, the eNode-B 160a, for example, may use multiple antennas to transmit wireless signals to, and/or receive wireless signals from, the WTRU 102a.

[0058] Each of the eNode-Bs 160a, 160b, 160c may be associated with a particular cell (not shown) and may be configured to handle radio resource management decisions, handover decisions, scheduling of users in the UL and/or DL, and the like. As shown in FIG. 1C, the eNode-Bs 160a, 160b, 160c may communicate with one another over an X2 interface.

[0059] The CN 106 shown in FIG. 1C may include a mobility management entity (MME) 162, a serving gateway (SGW) 164, and a packet data network (PDN) gateway (or PGW) 166. While each of the foregoing elements are depicted as part of the CN 106, it will be appreciated that any of these elements may be owned and/or operated by an entity other than the CN operator.

[0060] The MME 162 may be connected to each of the eNode-Bs 162a, 162b, 162c in the RAN 104 via an S1 interface and may serve as a control node. For example, the MME 162 may be responsible for authenticating users of the WTRUs 102a, 102b, 102c, bearer activation/deactivation, selecting a particular serving gateway during an initial attach of the WTRUs 102a, 102b, 102c, and the like. The MME 162 may provide a control plane function for switching between the RAN 104 and other RANs (not shown) that employ other radio technologies, such as GSM and/or WCDMA.

[0061] The SGW 164 may be connected to each of the eNode Bs 160a, 160b, 160c in the RAN 104 via the S1 interface. The SGW 164 may generally route and forward user data packets to/from the WTRUs 102a, 102b, 102c. The SGW 164 may perform other functions, such as anchoring user planes during inter-eNode B handovers, triggering paging when DL data is available for the WTRUs 102a, 102b, 102c, managing and storing contexts of the WTRUs 102a, 102b, 102c, and the like.

[0062] The SGW 164 may be connected to the PGW 166, which may provide the WTRUs 102a, 102b, 102c with access to packet-switched networks, such as the Internet 110, to facilitate communications between the WTRUs 102a, 102b, 102c and IP-enabled devices.

[0063] The CN 106 may facilitate communications with other networks. For example, the CN 106 may provide the WTRUs 102a, 102b, 102c with access to circuit-switched networks, such as the PSTN 108, to facilitate communications between the WTRUs 102a, 102b, 102c and traditional land-line communications devices. For example, the CN 106 may include, or may communicate with, an IP gateway (e.g., an IP multimedia subsystem (IMS) server) that serves as an interface between the CN 106 and the PSTN 108. In addition, the CN 106 may provide the WTRUs 102a, 102b, 102c with access to the other networks 112, which may include other wired and/or wireless networks that are owned and/or operated by other service providers.

[0064] Although the WTRU is described in FIGS. 1A-1D as a wireless terminal, it is contemplated that in certain representative embodiments that such a terminal may use (e.g., temporarily or permanently) wired communication interfaces with the communication network.

[0065] In representative embodiments, the other network 112 may be a WLAN.

[0066] A WLAN in Infrastructure Basic Service Set (BSS) mode may have an Access Point (AP) for the BSS and one or more stations (STAs) associated with the AP. The AP may have an access or an interface to a Distribution System (DS) or another type of wired/wireless network that carries traffic in to and/or out of the BSS. Traffic to STAs that originates from outside the BSS may arrive through the AP and may be delivered to the STAs. Traffic originating from STAs to destinations outside the BSS may be sent to the AP to be delivered to respective destinations. Traffic between STAs within the BSS may be sent through the AP, for example, where the source STA may send traffic to the AP and the AP may deliver the traffic to the destination STA. The traffic between STAs within a BSS may be considered and/or referred to as peer-to-peer traffic. The peer-to-peer traffic may be sent between (e.g., directly between) the source and destination STAs with a direct link setup (DLS). In certain representative embodiments, the DLS may use an 802.11e DLS or an 802.11z tunneled DLS (TDLS). A WLAN using an Independent BSS (IBSS) mode may not have an AP, and the STAs (e.g., all of the STAs) within or using the IBSS may communicate directly with each other. The IBSS mode of communication may sometimes be referred to herein as an “ad-hoc” mode of communication.

[0067] When using the 802.11ac infrastructure mode of operation or a similar mode of operations, the AP may transmit a beacon on a fixed channel, such as a primary channel. The primary channel may be a fixed width (e.g., 20 MHz wide bandwidth) or a dynamically set width via signaling. The primary channel may be the operating channel of the BSS and may be used by the STAs to establish a connection with the AP. In certain representative embodiments, Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) may be implemented, for example in in 802.11 systems. For CSMA/CA, the STAs (e.g., every STA), including the AP, may sense the primary channel. If the primary channel is sensed/detected and/or determined to be busy by a particular STA, the particular STA may back off. One STA (e.g., only one station) may transmit at any given time in a given BSS.

[0068] High Throughput (HT) STAs may use a 40 MHz wide channel for communication, for example, via a combination of the primary 20 MHz channel with an adjacent or nonadjacent 20 MHz channel to form a 40 MHz wide channel.

[0069] Very High Throughput (VHT) STAs may support 20MHz, 40 MHz, 80 MHz, and/or 160 MHz wide channels. The 40 MHz, and/or 80 MHz, channels may be formed by combining contiguous 20 MHz channels. A 160 MHz channel may be formed by combining 8 contiguous 20 MHz channels, or by combining two non-contiguous 80 MHz channels, which may be referred to as an 80+80 configuration. For the 80+80 configuration, the data, after channel encoding, may be passed through a segment parser that may divide the data into two streams. Inverse Fast Fourier Transform (IFFT) processing, and time domain processing, may be done on each stream separately. The streams may be mapped on to the two 80 MHz channels, and the data may be transmitted by a transmitting STA. At the receiver of the receiving STA, the above described operation for the 80+80 configuration may be reversed, and the combined data may be sent to the Medium Access Control (MAC).

[0070] Sub 1 GHz modes of operation are supported by 802.11af and 802.11ah. The channel operating bandwidths, and carriers, are reduced in 802.11af and 802.11ah relative to those used in 802.11n, and 802.11ac. 802.11af supports 5 MHz, 10 MHz and 20 MHz bandwidths in the TV White Space (TVWS) spectrum, and 802.11ah supports 1 MHz, 2 MHz, 4 MHz, 8 MHz, and 16 MHz bandwidths using non-TVWS spectrum. According to a representative embodiment, 802.11ah may support Meter Type Control/Machine-Type Communications, such as MTC devices in a macro coverage area. MTC devices may have certain capabilities, for example, limited capabilities including support for (e.g., only support for) certain and/or limited bandwidths. The MTC devices may include a battery with a battery life above a threshold (e.g., to maintain a very long battery life).

[0071] WLAN systems, which may support multiple channels, and channel bandwidths, such as 802.11n, 802.11ac, 802.11af, and 802.11ah, include a channel which may be designated as the primary channel. The primary channel may have a bandwidth equal to the largest common operating bandwidth supported by all STAs in the BSS. The bandwidth of the primary channel may be set and/or limited by a STA, from among all STAs in operating in a BSS, which supports the smallest bandwidth operating mode. In the example of 802.11ah, the primary channel may be 1 MHz wide for STAs (e.g., MTC type devices) that support (e.g., only support) a 1 MHz mode, even if the AP, and other STAs in the BSS support 2 MHz, 4 MHz, 8 MHz, 16 MHz, and/or other channel bandwidth operating modes. Carrier sensing and/or Network Allocation Vector (NAV) settings may depend on the status of the primary channel. If the primary channel is busy, for example, due to a STA (which supports only a 1 MHz operating mode), transmitting to the AP, the entire available frequency bands may be considered busy even though a majority of the frequency bands remains idle and may be available.

[0072] In the United States, the available frequency bands, which may be used by 802.11ah, are from 902 MHz to 928 MHz. In Korea, the available frequency bands are from 917.5 MHz to 923.5 MHz. In Japan, the available frequency bands are from 916.5 MHz to 927.5 MHz. The total bandwidth available for 802.11ah is 6 MHz to 26 MHz depending on the country code.

[0073] FIG. 1D is a system diagram illustrating the RAN 113 and the CN 115 according to an embodiment. As noted above, the RAN 113 may employ an NR radio technology to communicate with the WTRUs 102a, 102b, 102c over the air interface 116. The RAN 113 may also be in communication with the CN 115.

[0074] The RAN 113 may include gNBs 180a, 180b, 180c, though it will be appreciated that the RAN 113 may include any number of gNBs while remaining consistent with an embodiment. The gNBs 180a, 180b, 180c may each include one or more transceivers for communicating with the WTRUs 102a, 102b, 102c over the air interface 116. In one embodiment, the gNBs 180a, 180b, 180c may implement MIMO technology. For example, gNBs 180a, 108b may utilize beamforming to transmit signals to and/or receive signals from the gNBs 180a, 180b, 180c. Thus, the gNB 180a, for example, may use multiple antennas to transmit wireless signals to, and/or receive wireless signals from, the WTRU 102a. In an embodiment, the gNBs 180a, 180b, 180c may implement carrier aggregation technology. For example, the gNB 180a may transmit multiple component carriers to the WTRU 102a (not shown). A subset of these component carriers may be on unlicensed spectrum while the remaining component carriers may be on licensed spectrum. In an embodiment, the gNBs 180a, 180b, 180c may implement Coordinated Multi-Point (CoMP) technology. For example, WTRU 102a may receive coordinated transmissions from gNB 180a and gNB 180b (and/or gNB 180c).

[0075] The WTRUs 102a, 102b, 102c may communicate with gNBs 180a, 180b, 180c using transmissions associated with a scalable numerology. For example, the OFDM symbol spacing and/or OFDM subcarrier spacing may vary for different transmissions, different cells, and/or different portions of the wireless transmission spectrum. The WTRUs 102a, 102b, 102c may communicate with gNBs 180a, 180b, 180c using subframe or transmission time intervals (TTIs) of various or scalable lengths (e.g., containing varying number of OFDM symbols and/or lasting varying lengths of absolute time).

[0076] The gNBs 180a, 180b, 180c may be configured to communicate with the WTRUs 102a, 102b, 102c in a standalone configuration and/or a non-standalone configuration. In the standalone configuration, WTRUs 102a, 102b, 102c may communicate with gNBs 180a, 180b, 180c without also accessing other RANs (e.g., such as eNode-Bs 160a, 160b, 160c). In the standalone configuration, WTRUs 102a, 102b, 102c may utilize one or more of gNBs 180a, 180b, 180c as a mobility anchor point. In the standalone configuration, WTRUs 102a, 102b, 102c may communicate with gNBs 180a, 180b, 180c using signals in an unlicensed band. In a non-standalone configuration WTRUs 102a, 102b, 102c may communicate with/connect to gNBs 180a, 180b, 180c while also communicating with/connecting to another RAN such as eNode-Bs 160a, 160b, 160c. For example, WTRUs 102a, 102b, 102c may implement DC principles to communicate with one or more gNBs 180a, 180b, 180c and one or more eNode-Bs 160a, 160b, 160c substantially simultaneously. In the non-standalone configuration, eNode-Bs 160a, 160b, 160c may serve as a mobility anchor for WTRUs 102a, 102b, 102c and gNBs 180a, 180b, 180c may provide additional coverage and/or throughput for servicing WTRUs 102a, 102b, 102c.

[0077] Each of the gNBs 180a, 180b, 180c may be associated with a particular cell (not shown) and may be configured to handle radio resource management decisions, handover decisions, scheduling of users in the UL and/or DL, support of network slicing, dual connectivity, interworking between NR and E-UTRA, routing of user plane data towards User Plane Function (UPF) 184a, 184b, routing of control plane information towards Access and Mobility Management Function (AMF) 182a, 182b and the like. As shown in FIG. 1D, the gNBs 180a, 180b, 180c may communicate with one another over an Xn interface.

[0078] The CN 115 shown in FIG. 1D may include at least one AMF 182a, 182b, at least one UPF 184a,184b, at least one Session Management Function (SMF) 183a, 183b, and possibly a Data Network (DN) 185a, 185b. While each of the foregoing elements are depicted as part of the CN 115, it will be appreciated that any of these elements may be owned and/or operated by an entity other than the CN operator.

[0079] The AMF 182a, 182b may be connected to one or more of the gNBs 180a, 180b, 180c in the RAN 113 via an N2 interface and may serve as a control node. For example, the AMF 182a, 182b may be responsible for authenticating users of the WTRUs 102a, 102b, 102c, support for network slicing (e.g., handling of different PDU sessions with different requirements), selecting a particular SMF 183a, 183b, management of the registration area, termination of NAS signaling, mobility management, and the like. Network slicing may be used by the AMF 182a, 182b to customize CN support for WTRUs 102a, 102b, 102c based on the types of services being utilized WTRUs 102a, 102b, 102c. For example, different network slices may be established for different use cases such as services relying on ultra-reliable low latency (URLLC) access, services relying on enhanced massive mobile broadband (eMBB) access, services for machine type communication (MTC) access, and/or the like. The AMF 162 may provide a control plane function for switching between the RAN 113 and other RANs (not shown) that employ other radio technologies, such as LTE, LTE-A, LTE-A Pro, and/or non-3GPP access technologies such as WiFi.

[0080] The SMF 183a, 183b may be connected to an AMF 182a, 182b in the CN 115 via an N11 interface. The SMF 183a, 183b may also be connected to a UPF 184a, 184b in the CN 115 via an N4 interface. The SMF 183a, 183b may select and control the UPF 184a, 184b and configure the routing of traffic through the UPF 184a, 184b. The SMF 183a, 183b may perform other functions, such as managing and allocating UE IP address, managing PDU sessions, controlling policy enforcement and QoS, providing downlink data notifications, and the like. A PDU session type may be IP-based, non-IP based, Ethernet-based, and the like.

[0081] The UPF 184a, 184b may be connected to one or more of the gNBs 180a, 180b, 180c in the RAN 113 via an N3 interface, which may provide the WTRUs 102a, 102b, 102c with access to packet-switched networks, such as the Internet 110, to facilitate communications between the WTRUs 102a, 102b, 102c and IP-enabled devices. The UPF 184, 184b may perform other functions, such as routing and forwarding packets, enforcing user plane policies, supporting multi-homed PDU sessions, handling user plane QoS, buffering downlink packets, providing mobility anchoring, and the like.

[0082] The CN 115 may facilitate communications with other networks. For example, the CN 115 may include, or may communicate with, an IP gateway (e.g., an IP multimedia subsystem (IMS) server) that serves as an interface between the CN 115 and the PSTN 108. In addition, the CN 115 may provide the WTRUs 102a, 102b, 102c with access to the other networks 112, which may include other wired and/or wireless networks that are owned and/or operated by other service providers. In one embodiment, the WTRUs 102a, 102b, 102c may be connected to a local Data Network (DN) 185a, 185b through the UPF 184a, 184b via the N3 interface to the UPF 184a, 184b and an N6 interface between the UPF 184a, 184b and the DN 185a, 185b.

[0083] In view of Figures 1A-1D, and the corresponding description of Figures 1A-1D, one or more, or all, of the functions described herein with regard to one or more of: WTRU 102a-d, Base Station 114a-b, eNode-B 160a-c, MME 162, SGW 164, PGW 166, gNB 180a-c, AMF 182a-b, UPF 184a-b, SMF 183a-b, DN 185a-b, and/or any other device(s) described herein, may be performed by one or more emulation devices (not shown). The emulation devices may be one or more devices configured to emulate one or more, or all, of the functions described herein. For example, the emulation devices may be used to test other devices and/or to simulate network and/or WTRU functions.

[0084] The emulation devices may be designed to implement one or more tests of other devices in a lab environment and/or in an operator network environment. For example, the one or more emulation devices may perform the one or more, or all, functions while being fully or partially implemented and/or deployed as part of a wired and/or wireless communication network in order to test other devices within the communication network. The one or more emulation devices may perform the one or more, or all, functions while being temporarily implemented/deployed as part of a wired and/or wireless communication network. The emulation device may be directly coupled to another device for purposes of testing and/or may perform testing using over-the-air wireless communications.

[0085] The one or more emulation devices may perform the one or more, including all, functions while not being implemented/deployed as part of a wired and/or wireless communication network. For example, the emulation devices may be utilized in a testing scenario in a testing laboratory and/or a non-deployed (e.g., testing) wired and/or wireless communication network in order to implement testing of one or more components. The one or more emulation devices may be test equipment. Direct RF coupling and/or wireless communications via RF circuitry (e.g., which may include one or more antennas) may be used by the emulation devices to transmit and/or receive data.

[0086] This application describes a variety of aspects, including tools, features, examples or embodiments, models, approaches, etc. Many of these aspects are described with specificity and, at least to show the individual characteristics, are often described in a manner that may sound limiting. However, this is for purposes of clarity in description, and does not limit the application or scope of those aspects. Indeed, all of the different aspects may be combined and interchanged to provide further aspects. Moreover, the aspects may be combined and interchanged with aspects described in earlier filings as well.

[0087] The aspects described and contemplated in this application may be implemented in many different forms. FIGS. 5-10 described herein may provide some embodiments, but other embodiments are contemplated. The discussion of FIGS. 5-10 does not limit the breadth of the implementations. At least one of the aspects generally relates to video encoding and decoding, and at least one other aspect generally relates to transmitting a bitstream generated or encoded. These and other aspects may be implemented as a method, an apparatus, a computer readable storage medium having stored thereon instructions for encoding or decoding video data according to any of the methods described, and/or a computer readable storage medium having stored thereon a bitstream generated according to any of the methods described.

[0088] In the present application, the terms “reconstructed” and “decoded” may be used interchangeably, the terms “pixel” and “sample” may be used interchangeably, the terms “image,” “picture” and “frame” may be used interchangeably.

[0089] The terms HDR (high dynamic range) and SDR (standard dynamic range) may be used in this disclosure. Those terms often convey specific values of dynamic range to those of ordinary skill in the art. However, additional embodiments are also intended in which a reference to HDR is understood to mean “higher dynamic range” and a reference to SDR is understood to mean “lower dynamic range.” Such additional embodiments are not constrained by any specific values of dynamic range that might often be associated with the terms “high dynamic range” and “standard dynamic range.”

[0090] Various methods are described herein, and each of the methods comprises one or more steps or actions for achieving the described method. Unless a specific order of steps or actions is required for proper operation of the method, the order and/or use of specific steps and/or actions may be modified or combined. Additionally, terms such as “first”, “second”, etc. may be used in various embodiments to modify an element, component, step, operation, etc., such as, for example, a “first decoding” and a “second decoding”. Use of such terms does not imply an ordering to the modified operations unless specifically required. So, in this example, the first decoding need not be performed before the second decoding, and may occur, for example, before, during, or in an overlapping time period with the second decoding.

[0091] Various methods and other aspects described in this application may be used to modify modules, for example, in-loop filters modules (265 and 365), of a video encoder 200 and decoder 300 as shown in FIG. 2 and FIG. 3, respectively. Moreover, the subject matter disclosed herein presents aspects that are not limited to VVC or HEVC, and may be applied, for example, to any type, format or version of video coding, whether described in a standard or a recommendation, whether pre-existing or future-developed, and extensions of any such standards and recommendations (e.g., including VVC and HEVC). Unless indicated otherwise, or technically precluded, the aspects described in this application may be used individually or in combination.

[0092] Various numeric values are used in examples described in the present application, such as coefficients, filter sizes, etc. These and other specific values are for purposes of describing examples and the aspects described are not limited to these specific values.

[0093] FIG. 2 is a diagram showing an example video encoder (e.g., an example block-based hybrid video encoder) 200. Variations of example encoder 200 are contemplated, but the encoder 200 is described below for purposes of clarity without describing all expected variations.

[0094] Before being encoded, the video sequence may go through pre-encoding processing (201), for example, applying a color transform to the input color picture (e.g., conversion from RGB 4:4:4 to YCbCr 4:2:0), or performing a remapping of the input picture components in order to get a signal distribution more resilient to compression (for instance using a histogram equalization of one of the color components). Metadata may be associated with the pre-processing, and attached to the bitstream.

[0095] In the encoder 200, a picture is encoded by the encoder elements as described below. The picture to be encoded is partitioned (202) and processed in units of, for example, coding units (CUs). Each unit is encoded using, for example, either an intra or inter mode. When a unit is encoded in an intra mode, it performs intra prediction (260). In an inter mode, motion estimation (275) and compensation (270) are performed. The encoder decides (205) which one of the intra mode or inter mode to use for encoding the unit, and indicates the intra/inter decision by, for example, a prediction mode flag. Prediction residuals are calculated, for example, by subtracting (210) the predicted block from the original image block.

[0096] The prediction residuals are then transformed (225) and quantized (230). The quantized transform coefficients, as well as motion vectors and other syntax elements, are entropy coded (245) to output a bitstream. The encoder can skip the transform and apply quantization directly to the non-transformed residual signal. The encoder can bypass both transform and quantization, i.e., the residual is coded directly without the application of the transform or quantization processes.

[0097] The encoder decodes an encoded block to provide a reference for further predictions. The quantized transform coefficients are de-quantized (240) and inverse transformed (250) to decode prediction residuals. Combining (255) the decoded prediction residuals and the predicted block, an image block is reconstructed. In-loop filters (265) are applied to the reconstructed picture to perform, for example, deblocking/SAO (Sample Adaptive Offset) filtering to reduce encoding artifacts. The filtered image is stored at a reference picture buffer (280).

[0098] FIG. 3 is a diagram showing an example of a video decoder 300. In example decoder 300, a bitstream is decoded by the decoder elements as described below. Video decoder 300 generally performs a decoding pass reciprocal to the encoding pass as described in FIG. 2. The encoder 200 also generally performs video decoding as part of encoding video data.

[0099] In particular, the input of the decoder includes a video bitstream, which may be generated by video encoder 200. The bitstream is first entropy decoded (330) to obtain transform coefficients, motion vectors, and other coded information. The picture partition information indicates how the picture is partitioned. The decoder may therefore divide (335) the picture according to the decoded picture partitioning information. The transform coefficients are de-quantized (340) and inverse transformed (350) to decode the prediction residuals. Combining (355) the decoded prediction residuals and the predicted block, an image block is reconstructed. The predicted block may be obtained (370) from intra prediction (360) or motion-compensated prediction (i.e., inter prediction) (375). In-loop filters (365) are applied to the reconstructed image. The filtered image is stored at a reference picture buffer (380). For a given picture, the contents of the reference picture buffer 380 on the decoder 300 side may be identical to the contents of the reference picture buffer 280 on the encoder 200 side for the same picture.

[0100] The decoded picture can further go through post-decoding processing (385), for example, an inverse color transform (e.g., conversion from YCbCr 4:2:0 to RGB 4:4:4) or an inverse remapping performing the inverse of the remapping process performed in the pre-encoding processing (201). The post-decoding processing can use metadata derived in the pre-encoding processing and signaled in the bitstream.

[0101] FIG. 4 is a diagram showing an example of a system in which various aspects and embodiments described herein may be implemented. System 400 may be embodied as a device including the various components described below and is configured to perform one or more of the aspects described in this document. Examples of such devices, include, but are not limited to, various electronic devices such as personal computers, laptop computers, smartphones, tablet computers, digital multimedia set top boxes, digital television receivers, personal video recording systems, connected home appliances, and servers. Elements of system 400, singly or in combination, may be embodied in a single integrated circuit (IC), multiple ICs, and/or discrete components. For example, in at least one example, the processing and encoder/decoder elements of system 400 are distributed across multiple ICs and/or discrete components. In various embodiments, the system 400 is communicatively coupled to one or more other systems, or other electronic devices, via, for example, a communications bus or through dedicated input and/or output ports. In various embodiments, the system 400 is configured to implement one or more of the aspects described in this document.

[0102] The system 400 includes at least one processor 410 configured to execute instructions loaded therein for implementing, for example, the various aspects described in this document. Processor 410 can include embedded memory, input output interface, and various other circuitries as known in the art. The system 400 includes at least one memory 420 (e.g., a volatile memory device, and/or a non-volatile memory device). System 400 includes a storage device 440, which can include non-volatile memory and/or volatile memory, including, but not limited to, Electrically Erasable Programmable Read-Only Memory (EEPROM), Read-Only Memory (ROM), Programmable Read-Only Memory (PROM), Random Access Memory (RAM), Dynamic Random Access Memory (DRAM), Static Random Access Memory (SRAM), flash, magnetic disk drive, and/or optical disk drive. The storage device 440 can include an internal storage device, an attached storage device (including detachable and non-detachable storage devices), and/or a network accessible storage device, as non-limiting examples.

[0103] System 400 includes an encoder/decoder module 430 configured, for example, to process data to provide an encoded video or decoded video, and the encoder/decoder module 430 can include its own processor and memory. The encoder/decoder module 430 represents module(s) that may be included in a device to perform the encoding and/or decoding functions. As is known, a device can include one or both of the encoding and decoding modules. Additionally, encoder/decoder module 430 may be implemented as a separate element of system 400 or may be incorporated within processor 410 as a combination of hardware and software as known to those skilled in the art.

[0104] Program code to be loaded onto processor 410 or encoder/decoder 430 to perform the various aspects described in this document may be stored in storage device 440 and subsequently loaded onto memory 420 for execution by processor 410. In accordance with various embodiments, one or more of processor 410, memory 420, storage device 440, and encoder/decoder module 430 can store one or more of various items during the performance of the processes described in this document. Such stored items can include, but are not limited to, the input video, the decoded video or portions of the decoded video, the bitstream, matrices, variables, and intermediate or final results from the processing of equations, formulas, operations, and operational logic.

[0105] In some embodiments, memory inside of the processor 410 and/or the encoder/decoder module 430 is used to store instructions and to provide working memory for processing that is needed during encoding or decoding. In other embodiments, however, a memory external to the processing device (for example, the processing device may be either the processor 410 or the encoder/decoder module 430) is used for one or more of these functions. The external memory may be the memory 420 and/or the storage device 440, for example, a dynamic volatile memory and/or a non-volatile flash memory. In several embodiments, an external non-volatile flash memory is used to store the operating system of, for example, a television. In at least one embodiment, a fast external dynamic volatile memory such as a RAM is used as working memory for video coding and decoding operations, such as, for example, MPEG-2 (MPEG refers to the Moving Picture Experts Group, MPEG-2 is also referred to as ISO/IEC 13818, and 13818-1 is also known as H.222, and 13818-2 is also known as H.262), HEVC (HEVC refers to High Efficiency Video Coding, also known as H.265 and MPEG-H Part 2), or VVC (Versatile Video Coding, a new standard being developed by JVET, the Joint Video Experts Team).

[0106] The input to the elements of system 400 may be provided through various input devices as indicated in block 445. Such input devices include, but are not limited to, (i) a radio frequency (RF) portion that receives an RF signal transmitted, for example, over the air by a broadcaster, (ii) a Component (COMP) input terminal (or a set of COMP input terminals), (iii) a Universal Serial Bus (USB) input terminal, and/or (iv) a High Definition Multimedia Interface (HDMI) input terminal. Other examples, not shown in FIG. 4, include composite video.

[0107] In various embodiments, the input devices of block 445 have associated respective input processing elements as known in the art. For example, the RF portion may be associated with elements suitable for (i) selecting a desired frequency (also referred to as selecting a signal, or band-limiting a signal to a band of frequencies), (ii) down-converting the selected signal, (iii) band-limiting again to a narrower band of frequencies to select (for example) a signal frequency band which may be referred to as a channel in certain embodiments, (iv) demodulating the down-converted and band-limited signal, (v) performing error correction, and (vi) demultiplexing to select the desired stream of data packets. The RF portion of various embodiments includes one or more elements to perform these functions, for example, frequency selectors, signal selectors, band-limiters, channel selectors, filters, downconverters, demodulators, error correctors, and demultiplexers. The RF portion can include a tuner that performs various of these functions, including, for example, down-converting the received signal to a lower frequency (for example, an intermediate frequency or a near-baseband frequency) or to baseband. In one set-top box embodiment, the RF portion and its associated input processing element receives an RF signal transmitted over a wired (for example, cable) medium, and performs frequency selection by filtering, down-converting, and filtering again to a desired frequency band. Various embodiments rearrange the order of the above-described (and other) elements, remove some of these elements, and/or add other elements performing similar or different functions. Adding elements can include inserting elements in between existing elements, such as, for example, inserting amplifiers and an analog-to-digital converter. In various embodiments, the RF portion includes an antenna.

[0108] Additionally, the USB and/or HDMI terminals can include respective interface processors for connecting system 400 to other electronic devices across USB and/or HDMI connections. It is to be understood that various aspects of input processing, for example, Reed-Solomon error correction, may be implemented, for example, within a separate input processing IC or within processor 410 as necessary. Similarly, aspects of USB or HDMI interface processing may be implemented within separate interface ICs or within processor 410 as necessary. The demodulated, error corrected, and demultiplexed stream is provided to various processing elements, including, for example, processor 410, and encoder/decoder 430 operating in combination with the memory and storage elements to process the data stream as necessary for presentation on an output device.

[0109] Various elements of system 400 may be provided within an integrated housing, Within the integrated housing, the various elements may be interconnected and transmit data therebetween using suitable connection arrangement 425, for example, an internal bus as known in the art, including the Inter-IC (I2C) bus, wiring, and printed circuit boards.

[0110] The system 400 includes communication interface 450 that enables communication with other devices via communication channel 460. The communication interface 450 can include, but is not limited to, a transceiver configured to transmit and to receive data over communication channel 460. The communication interface 450 can include, but is not limited to, a modem or network card and the communication channel 460 may be implemented, for example, within a wired and/or a wireless medium.

[0111] Data is streamed, or otherwise provided, to the system 400, in various embodiments, using a wireless network such as a Wi-Fi network, for example IEEE 802.11 (IEEE refers to the Institute of Electrical and Electronics Engineers). The Wi-Fi signal of these examples is received over the communications channel 460 and the communications interface 450 which are adapted for Wi-Fi communications. The communications channel 460 of these embodiments is typically connected to an access point or router that provides access to external networks including the Internet for allowing streaming applications and other over-the-top communications. Other embodiments provide streamed data to the system 400 using a set-top box that delivers the data over the HDMI connection of the input block 445. Still other embodiments provide streamed data to the system 400 using the RF connection of the input block 445. As indicated above, various embodiments provide data in a non-streaming manner. Additionally, various embodiments use wireless networks other than Wi-Fi, for example a cellular network or a Bluetooth network.

[0112] The system 400 can provide an output signal to various output devices, including a display 475, speakers 485, and other peripheral devices 495. The display 475 of various embodiments includes one or more of, for example, a touchscreen display, an organic light-emitting diode (OLED) display, a curved display, and/or a foldable display. The display 475 may be for a television, a tablet, a laptop, a cell phone (mobile phone), or other device. The display 475 can also be integrated with other components (for example, as in a smart phone), or separate (for example, an external monitor for a laptop). The other peripheral devices 495 include, in various examples of embodiments, one or more of a stand-alone digital video disc (or digital versatile disc) (DVR, for both terms), a disk player, a stereo system, and/or a lighting system. Various embodiments use one or more peripheral devices 495 that provide a function based on the output of the system 400. For example, a disk player performs the function of playing the output of the system 400.

[0113] In various embodiments, control signals are communicated between the system 400 and the display 475, speakers 485, or other peripheral devices 495 using signaling such as AV.Link, Consumer Electronics Control (CEC), or other communications protocols that enable device-to-device control with or without user intervention. The output devices may be communicatively coupled to system 400 via dedicated connections through respective interfaces 470, 480, and 490. Alternatively, the output devices may be connected to system 400 using the communications channel 460 via the communications interface 450. The display 475 and speakers 485 may be integrated in a single unit with the other components of system 400 in an electronic device such as, for example, a television. In various embodiments, the display interface 470 includes a display driver, such as, for example, a timing controller (T Con) chip.

[0114] The display 475 and speakers 485 can alternatively be separate from one or more of the other components, for example, if the RF portion of input 445 is part of a separate set-top box. In various embodiments in which the display 475 and speakers 485 are external components, the output signal may be provided via dedicated output connections, including, for example, HDMI ports, USB ports, or COMP outputs.

[0115] The embodiments may be carried out by computer software implemented by the processor 410 or by hardware, or by a combination of hardware and software. As a non-limiting example, the embodiments may be implemented by one or more integrated circuits. The memory 420 may be of any type appropriate to the technical environment and may be implemented using any appropriate data storage technology, such as optical memory devices, magnetic memory devices, semiconductor-based memory devices, fixed memory, and removable memory, as non-limiting examples. The processor 410 may be of any type appropriate to the technical environment, and can encompass one or more of microprocessors, general purpose computers, special purpose computers, and processors based on a multi-core architecture, as non-limiting examples.

[0116] Various implementations involve decoding. “Decoding”, as used in this application, can encompass all or part of the processes performed, for example, on a received encoded sequence in order to produce a final output suitable for display. In various embodiments, such processes include one or more of the processes typically performed by a decoder, for example, entropy decoding, inverse quantization, inverse transformation, and differential decoding. In various embodiments, such processes also, or alternatively, include processes performed by a decoder of various implementations described in this application, for example, estimating a relationship between residual of a first chrominance component of a block and residual of a second chrominance component of the block based on a prediction of the first chrominance component and a prediction of the second chrominance component; predicting the residual of the first chrominance component based on the estimated relationship and the residual of the second chrominance component; decoding the block based on the predicted residual of the first chrominance component, etc.

[0117] Other example processes that may be performed by a decoder include receiving an encoded set of filters, wherein the encoded set of filters comprises a plurality of encoded coefficients; receiving an indication of a size index associated with the encoded set of filters; and decoding the plurality of encoded coefficients based on the size index, etc.

[0118] Other example processes that may be performed by a decoder include receiving a set of encoded filter coefficients; decoding a first filter coefficient and a second filter coefficient of the set of encoded filter coefficients to determine a first value of the first filter coefficient and a second value of the second filter coefficient; determining a difference between the first value and the second value; and decoding a third filter coefficient based on the difference between the first value and the second value, etc.

[0119] As further embodiments, in one example "decoding" refers only to entropy decoding, in another embodiment "decoding" refers only to differential decoding, and in another embodiment "decoding" refers to a combination of entropy decoding and differential decoding. Whether the phrase "decoding process" is intended to refer specifically to a subset of operations or generally to the broader decoding process will be clear based on the context of the specific descriptions and is believed to be well understood by those skilled in the art.

[0120] Various implementations involve encoding. In an analogous way to the above discussion about "decoding", "encoding" as used in this application can encompass all or part of the processes performed, for example, on an input video sequence in order to produce an encoded bitstream. In various embodiments, such processes include one or more of the processes typically performed by an encoder, for example, partitioning, differential encoding, transformation, quantization, and entropy encoding. In various embodiments, such processes also, or alternatively, include processes performed by an encoder of various implementations described in this application, for example, estimating a relationship between residual of a first chrominance component and residual of a second chrominance component based on a prediction of the first chrominance component and a prediction of the second chrominance component; predicting the residual of the first chrominance component based on the estimated relationship and the residual of the second chrominance component; encoding a block based on the predicted residual of the first chrominance component, etc.

[0121] Other example processes that may be performed by an encoder include identifying, based on rate distortion, a set of filters, wherein the set of filters comprises a plurality of coefficients; determining a size index associated with the set of filters; encoding the plurality of coefficients to generate a plurality of encoded coefficients; and including, in video data, the plurality of encoded coefficients and an indication of the size index, etc.

[0122] Other example processes that may be performed by an encoder include identifying a set of filter coefficients; encoding a first value of a first filter coefficient of the set of encoded filter coefficients, and a second value of a second filter coefficient of the set of encoded filter coefficients; determining a difference between the first value and the second value; and encoding a third filter coefficient based on the difference between the first value and the second value, etc.

[0123] Other example processes that may be performed by an encoder include identifying a filter comprising a first coefficient and a second coefficient; encoding a value of the first coefficient using an exponential-Golomb code to generate a binary code having a first bitlength; identifying a second bitlength for representing a value of the second coefficient; and encoding the second coefficient with a modified exponential-Golomb code to generate an encoded second coefficient, wherein a prefix of the encoded second coefficient comprises n zero-bits, and wherein n is the difference between the second bitlength and the first bitlength, etc.

[0124] Other example processes that may be performed by an encoder include identifying a filter comprising a set of coefficients; encoding the set of coefficients using an entropy coder, wherein a probability associated with each coefficient of the set of coefficients depends on a corresponding spatial position of the coefficient; and including, in video data, the encoded set of coefficients, etc.

[0125] Other example processes that may be performed by an encoder include identifying a first sign of a first coefficient, and a second sign of a second coefficient; determining a third sign of a sum of the first coefficient and the second coefficient; encoding a fourth sign of a third coefficient based on the third sign of the sum; and including, in video data, the encoded fourth sign, etc.

[0126] As further examples, in one embodiment “encoding” refers only to entropy encoding, in another embodiment “encoding” refers only to differential encoding, and in another embodiment “encoding” refers to a combination of differential encoding and entropy encoding. Whether the phrase “encoding process” is intended to refer specifically to a subset of operations or generally to the broader encoding process will be clear based on the context of the specific descriptions and is believed to be well understood by those skilled in the art.

[0127] When a figure is presented as a flow diagram, it should be understood that it also provides a block diagram of a corresponding apparatus. Similarly, when a figure is presented as a block diagram, it should be understood that it also provides a flow diagram of a corresponding method/process.

[0128] Various embodiments refer to rate distortion optimization. In particular, during the encoding process, the balance or trade-off between the rate and distortion is usually considered, often given the constraints of computational complexity. The rate distortion optimization is usually formulated as minimizing a rate distortion function, which is a weighted sum of the rate and of the distortion. There are different approaches to solve the rate distortion optimization problem. For example, the approaches may be based on an extensive testing of all encoding options, including all considered modes or coding parameters values, with a complete evaluation of their coding cost and related distortion of the reconstructed signal after coding and decoding. Faster approaches may also be used, to save encoding complexity, in particular with computation of an approximated distortion based on the prediction or the prediction residual signal, not the reconstructed one. Mix of these two approaches can also be used, such as by using an approximated distortion for only some of the possible encoding options, and a complete distortion for other encoding options. Other approaches only evaluate a subset of the possible encoding options. More generally, many approaches employ any of a variety of techniques to perform the optimization, but the optimization is not necessarily a complete evaluation of both the coding cost and related distortion.

[0129] The implementations and aspects described herein may be implemented in, for example, a method or a process, an apparatus, a software program, a data stream, or a signal. Even if only discussed in the context of a single form of implementation (for example, discussed only as a method), the implementation of features discussed can also be implemented in other forms (for example, an apparatus or program). An apparatus may be implemented in, for example, appropriate hardware, software, and firmware. The methods may be implemented in, for example, a processor, which refers to processing devices in general, including, for example, a computer, a microprocessor, an integrated circuit, or a programmable logic device. Processors also include communication devices, such as, for example, computers, cell phones, portable/personal digital assistants ("PDAs"), and other devices that facilitate communication of information between end-users.

[0130] Reference to “one embodiment,” “an embodiment,” “an example,” “one implementation” or “an implementation,” as well as other variations thereof, means that a particular feature, structure, characteristic, and so forth described in connection with the embodiment is included in at least one embodiment. Thus, the appearances of the phrase “in one embodiment,” “in an embodiment,” “in an example,” “in one implementation,” or “in an implementation”, as well any other variations, appearing in various places throughout this application are not necessarily all referring to the same embodiment or example.

[0131] Additionally, this application may refer to “determining” various pieces of information. Determining the information can include one or more of, for example, estimating the information, calculating the information, predicting the information, or retrieving the information from memory. Obtaining may include receiving, retrieving, constructing, generating, and/or determining.

[0132] Further, this application may refer to “accessing” various pieces of information. Accessing the information can include one or more of, for example, receiving the information, retrieving the information (for example, from memory), storing the information, moving the information, copying the information, calculating the information, determining the information, predicting the information, or estimating the information.

[0133] Additionally, this application may refer to “receiving” various pieces of information. Receiving is, as with “accessing”, intended to be a broad term. Receiving the information can include one or more of, for example, accessing the information, or retrieving the information (for example, from memory). Further, “receiving” is typically involved, in one way or another, during operations such as, for example, storing the information, processing the information, transmitting the information, moving the information, copying the information, erasing the information, calculating the information, determining the information, predicting the information, or estimating the information.

[0134] It is to be appreciated that the use of any of the following “/”, “and/or”, and “at least one of”, for example, in the cases of “A/B”, “A and/or B” and “at least one of A and B”, is intended to encompass the selection of the first listed option (A) only, or the selection of the second listed option (B) only, or the selection of both options (A and B). As a further example, in the cases of “A, B, and/or C” and “at least one of A, B, and C”, such phrasing is intended to encompass the selection of the first listed option (A) only, or the selection of the second listed option (B) only, or the selection of the third listed option (C) only, or the selection of the first and the second listed options (A and B) only, or the selection of the first and third listed options (A and C) only, or the selection of the second and third listed options (B and C) only, or the selection of all three options (A and B and C). This may be extended, as is clear to one of ordinary skill in this and related arts, for as many items as are listed.

[0135] Also, as used herein, the word “signal” refers to, among other things, indicating something to a corresponding decoder. For example, in some embodiments the encoder signals (e.g., to a decoder) a size index of a filter or set of filters, etc. In this way, in an embodiment the same parameter is used at both the encoder side and the decoder side. Thus, for example, an encoder can transmit (explicit signaling) a particular parameter to the decoder so that the decoder can use the same particular parameter. Conversely, if the decoder already has the particular parameter as well as others, then signaling may be used without transmitting (implicit signaling) to simply allow the decoder to know and select the particular parameter. By avoiding transmission of any actual functions, a bit savings is realized in various embodiments. It is to be appreciated that signaling may be accomplished in a variety of ways. For example, one or more syntax elements, flags, and so forth are used to signal information to a corresponding decoder in various embodiments. While the preceding relates to the verb form of the word “signal”, the word “signal” can also be used herein as a noun.

[0136] Note that syntax elements as used herein, for example, *tu_joint_cbcr_residual_flag*, *tu_cbf_cb*, *tu_cbf_cr*, *joint_cb_cr*, *predicted_ residual*, , etc., are descriptive terms. As such, they do not preclude the use of other syntax element names.

[0137] As will be evident to one of ordinary skill in the art, implementations may produce a variety of signals formatted to carry information that may be, for example, stored or transmitted. The information can include, for example, instructions for performing a method, or data produced by one of the described implementations. For example, a signal may be formatted to carry the bitstream of a described embodiment. Such a signal may be formatted, for example, as an electromagnetic wave (for example, using a radio frequency portion of spectrum) or as a baseband signal. The formatting may include, for example, encoding a data stream and modulating a carrier with the encoded data stream. The information that the signal carries may be, for example, analog or digital information. The signal may be transmitted over a variety of different wired or wireless links, as is known. The signal may be stored on a processor-readable medium.

[0138] Many embodiments are described herein. Features of embodiments may be provided alone or in any combination, across various claim categories and types. Further, embodiments may include one or more of the features, devices, or aspects described herein, alone or in any combination, across various claim categories and types. For example, features described herein can be implemented in a bitstream or signal that includes information generated as described herein. The information can allow a decoder to decode a bitstream, the encoder, bitstream, and/or decoder according to any of the embodiments described. For example, features described herein can be implemented by creating and/or transmitting and/or receiving and/or decoding a bitstream or signal. For example, features described herein can be implemented a method, process, apparatus, medium storing instructions, medium storing data, or signal. For example, features described herein can be implemented by a TV, set-top box, cell phone, tablet, or other electronic device that performs decoding. The TV, set-top box, cell phone, tablet, or other electronic device can display (e.g., using a monitor, screen, or other type of display) a resulting image (e.g., an image from residual reconstruction of the video bitstream). The TV, set-top box, cell phone, tablet, or other electronic device can receive a signal including an encoded image and perform decoding.

[0139] A cross-component linear model (CCLM) may be used for intra prediction.

[0140] Chroma samples may be predicted based on reconstructed luma samples of a CU (e.g., the same CU). For example, the chroma samples may be predicted based on reconstructed luma samples of the same CU by using a linear model:

$$pred\_C\ (i,j) = \alpha \cdot rec\_L'(i,j) +\ \beta$$

where $pred\_C\ (i,j)$ represents the predicted chroma samples in a CU and $rec\_L\ (i,j)$ represents the down-sampled reconstructed luma samples of the same CU.

[0141] The CCLM parameters (α and β) may be derived from neighboring chroma samples in a reference area (e.g., top row and left column) and the corresponding down-sampled luma samples (e.g., LM mode). One or more (e.g., at most four) neighboring chroma samples may be used. The neighboring chroma samples may be selected from the top (e.g., top only LM-A), or the left (e.g., left only LM-L). The selected mode may be signaled.

[0142] The selected neighboring luma samples at the selected positions may be down-sampled. The selected neighboring luma samples at the selected positions may be compared to find smaller values (e.g., two smaller values: x^0_A and x^1_A) and larger values (e.g., two larger values: x^0_B and x^1_B). The corresponding chroma sample values may be denoted as y^0_A, y^1_A, y^0_B, and y^1_B. Then x_A, x_B, y_A, and y_B may be derived as:

$$X_a=(x^0_A + x^1_A +1)>>1$$

$$X_b=(x^0_B + x^1_B +1)>>1$$

$$Y_a=(y^0_A + y^1_A +1)>>1$$

$$Y_b=(y^0_B + y^1_B +1)>>1$$

[0143] The linear model parameters α and β may be obtained according to the following equations:

$$\mathrm{A} = \frac{Y_a - Y_b}{\mathrm{X}_a - X_b}$$

$$\beta = Y_b - \alpha \cdot X_b$$

[0144] FIG. 5 illustrates an example of the location of the left and above samples and the sample of the current block (e.g., involved in the CCLM mode). FIG. 5 illustrates example locations of the samples used for the derivation of α and β.

[0145] One or more (e.g., three) multi-model LM (MMLM) modes may be added. In an MMLM mode (e.g., each MMLM mode), the reconstructed neighboring samples may be classified into a plurality of classes (e.g., two classes). For example, the reconstructed neighboring samples may be classified using a threshold (e.g., the average of the luma reconstructed neighboring samples). The linear model of a class (e.g., each class) may be derived using the Least-Mean-Square (LMS) method. For the CCLM mode, the LMS method may be used to derive the linear model.

[0146] A slope adjustment may be applied to a cross-component linear model (CCLM) and to MMLM prediction. FIG. 6 illustrates the effect of the slope adjustment parameter “u”. The left side of FIG. 6 illustrates an example model created with CCLM (e.g., the regular CCLM). The right side of FIG. 6 illustrates an example model updated with slope adjustment. The adjustment may involve tilting the linear function. The linear function may map luma values to chroma values with respect to a center point. The center point may be determined by the average luma value of the reference samples, as illustrated in FIG. 6.

[0147] In some examples, a convolutional cross-component model (CCCM) may be used for intra prediction.

[0148] The CCCM may be used to predict chroma samples from reconstructed luma samples (e.g., in a similar spirit as done by CCLM). The reconstructed luma samples may be down-sampled to match the lower resolution chroma grid when chroma sub-sampling is used.

[0149] In examples, a single model or multi-model variant of CCCM may be used. Multi-model CCCM mode may be selected for prediction units (PUs) that have at least 128 reference samples available. The multi-model variant of CCCM may use a plurality of models (e.g., two models). For example, the multi-model variant of CCCM may use a model (e.g., one model) applied to samples with values higher than the average luma reference value. Another model may be used for the rest of the samples (e.g., samples with values less than the average luma reference value).

[0150] CCCM may use a convolutional filter made of parameters, which may be used to weight input samples. For example, a convolutional filter may be made of 7 parameters weighting 7 inputs samples $\{s_i\}, where\ i = 0, \ldots, 6$. FIG. 7 illustrates an example spatial part of a convolutional filter. One or more (e.g., five) coefficients may be applied to luminance pixel values (e.g., corresponding to a plus sign shape). For example, a coefficient (e.g., one coefficient) may be applied to a square term (P). A coefficient (e.g., the last coefficient) may be applied to a bias term (B). The input to the spatial 5-tap component of the filter may include a center (C) luma sample. The center luma sample may be collocated with the chroma sample to be predicted and above/north (N), below/south (S), left/west (W) and right/east (E) neighbors as illustrated in FIG. 7.

[0151] The square term P may be represented as power of two of the center luma sample C and scaled to the sample value range of the content:

$$P = (C * C + midVal) >> bitDepth$$

where $midVal$ is a rounding term. For example, for 10-bit content, the square term P may be calculated as:

$$P = (C * C + 512) >> 10$$

[0152] The bias term B may represent a scalar offset between the input and output (e.g., similar to the offset term in CCLM). The bias term B may be set to a middle chroma value (e.g., 512 for 10-bit content).

[0153] The output of the filter may be calculated as a convolution between the filter coefficients c_i and the input values, and clipped to the range of valid chroma samples:

$$predChromaVal = c_0C + c_1N + c_2S + c_3E + c_4W + c_5P + c_6B$$

[0154] The filter coefficients c_i may be calculated by minimizing a mean squared error (MSE) between predicted and reconstructed chroma samples in the reference area. FIG. 8 illustrates an example reference area 802 used to derive the filter coefficients. The reference area 802 may include 6 lines/columns of chroma samples above and left of a PU 804. The reference area 802 may extend one PU width to the right and one PU height below the PU boundaries. The reference area 802 may be adjusted to include available samples (e.g., only available samples). The extensions 806 to the reference area 802 may be used to support "side samples" of the plus-shaped spatial filter. The extensions 806 may be padded (e.g., if the extensions are in unavailable areas).

[0155] The MSE minimization may be performed by calculating an autocorrelation matrix for the luma input and a cross-correlation vector between the luma input and chroma output.

$$\text{denote } s = \begin{pmatrix} C \\ N \\ S \\ E \\ W \\ P \\ B \end{pmatrix} \text{ where } s_i = \begin{cases} i = 0: C \\ i = 1: N \\ i = 2: S \\ i = 3: E \\ i = 4: W \\ i = 5: P \\ i = 6: B \end{cases}$$

[0156] To derive the coefficients $\{ci\}$, the autocorrelation matrix may be inverted. For example, the autocorrelation matrix may be LDL^T decomposed. The final filter coefficients may be calculated using back-substitution. This calculation may roughly follow the calculation of the ALF filter coefficients. LDL^T decomposition (sometimes also referred to as alternative Cholesky decomposition) may be used (e.g., instead of Cholesky decomposition) to avoid using square root operations. A Gaussian elimination technique may be used to invert the matrix. The calculation may use integer 64-bits arithmetic.

[0157] Feature(s) associated with a gradient and location based convolutional cross-component model (GL-CCCM) are provided herein.

[0158] In GL-CCCM, the convolution may apply on gradient and location information. FIG. 9 illustrates an example spatial part of a GL-CCCM convolutional filter. The GL-CCCM filter for the prediction may be calculated as:

$$predChromaVal = c_0C + c_1Gy + c_2Gx + c_3Y + c_4X + c_5P + c_6B$$

where Gy and Gx are the luma vertical and horizontal gradients, respectively. With reference to FIG. 9, Gy and Gx may be calculated as:

$$Gy = (2N + NW + NE) - (2S + SW + SE)$$

$$Gx = (2W + NW + SW) - (2E + NE + SE)$$

[0159] The Y and X parameters may be the vertical and horizontal coordinates, respectively, of the center luma sample location. Other parameters may be the same as in CCCM. The reference area for the parameter calculation may be the same as in CCCM described herein.

[0160] In some examples, the reconstructed luma samples may not be down-sampled to match the lower resolution chroma grid. FIG. 10 illustrates an example spatial part of non-down sampled CCCM luma terms used for a convolutional filter. If the luma samples are not down-sampled, the co-located reconstructed luma samples (e.g., 6 luma samples) may be used (e.g., directly), as illustrated in FIG. 10. One or more (e.g., 4) terms may be built from L_0, L_1, L_2, L_4, and the bias term B. In this case, the number of parameters may be 10.

[0161] In some examples, one or more (e.g., 4) down-sampling filters may be used to derive the input samples to be weighted with the coefficients. The down-sampling filter model may be signaled (e.g., per CU). The prediction of the chroma samples may be derived as follows:

$$Model\ 1: predChroma = c_0 * H(C) + c_1 * G1(C) + c_2 * G2(C) + c_3 * G3(C) + c_4 * G4(C) + c_5 * P + c_6 * B$$

$$Model\ 2: predChroma = c_0 * H(C) + c_1 * H(W) + c_2 * H(E) + c_3 * G1(C) + c_4 * G1(W) + c_5 * G1(E) + c_6 * B$$

$$Model\ 3: predChroma = c_0 * H(C) + c_1 * H(N) + c_2 * H(S) + c_3 * G2(C) + c_4 * G2(N) + c_5 * G2(S) + c_6 * B$$

$$Model\ 4: predChroma = c_0 * H(C) + c_1 * H(NE) + c_2 * H(SW) + c_3 * G4(C) + c_4 * G4(NE) + c_5 * G4(SW) + c_6 * B$$

where $H(\cdot)$, $G1(\cdot)$, $G2(\cdot)$, $G3(\cdot)$, and $G4(\cdot)$ are downsampling filters applied on luma samples, as illustrated in FIG. 11. FIG. 11 illustrates example down-sampling filters that may be used.

[0162] Cross-component residual model (CCRM) may be used for inter prediction. FIG. 12 illustrates an example of inter prediction using a CCRM.

[0163] A filter (e.g., an 8-tap filter) may include one or more (e.g., 6) spatial luma samples, a nonlinear term, and a bias term. The 8-tap filter may be used for the prediction of the chroma using the luma channel (e.g., only the luma channel) for inter coded blocks. The spatial luma samples $(L0, \ldots, L5)$ may be obtained from the luma grid. For example, the 6 luma samples closest to the chroma position C without down sampling (e.g., as illustrated in Fig. 6) may be selected as the spatial luma samples. The predicted chroma value may be calculated as:

$$predChromaVal = c_0 L0 + c_1 L1 + c_2 L2 + c_3 L3 + c_4 L4 + c_5 L5 + c_6\, nonlinear((L0 + L3 + 1) >> 1) + c_7 B,$$

where $nonlinear$ is a nonlinear operator for CCCM and B is the bias term.

[0164] Fig. 13 illustrates the luma samples $(L0, \ldots, L5)$ in relation to the chroma sample C (e.g., illustrated in a half-pel luma grid).

[0165] The filter coefficients may be derived using a division-free Gaussian elimination technique. The offsets may be applied to samples (e.g., prior to filter derivation).

[0166] Intra reference samples may be used as input samples in filter derivation (e.g., if the block has less than 64 chroma samples). CCCM with one or more (e.g., up to 6) rows and columns of intra reference samples may be used.

[0167] Blocks having 256 chroma samples or more may be divided into subblocks. The subblocks may have at most 256 chroma samples. Subblocks containing zero luma residual may be skipped.

[0168] Feature(s) associated with joint coding of chroma residuals (JCCR) are provided herein.

[0169] JCCR may be used if the chroma residuals are coded jointly. The usage (e.g., activation) of the JCCR mode may be indicated by a flag (e.g., a TU-level flag *tu_joint_cbcr_residual_flag*). The selected mode may be indicated (e.g., implicitly indicated) by the chroma coded block flags (CBFs). The flag *tu_joint_cbcr_residual_flag* may be present if either or both chroma CBFs for a TU are equal to 1. In the PPS and slice header, chroma quantization scale factor (e.g., quantization parameter (QP)) offset values may be signaled for the JCCR mode (e.g., to differentiate from other chroma QP offset values signaled for each chroma component for CCCM). The chroma QP offset values may be used to derive the chroma QP values for some blocks (e.g., blocks coded using the JCCR mode). The JCCR mode may have one or more (e.g., 3) sub-modes. If a corresponding JCCR sub-mode (e.g., sub-mode 2 in Table 1) is active in a TU, the corresponding chroma QP offset may be added to the applied luma-derived chroma QP (e.g., during quantization and decoding of that TU). For some JCCR sub-modes (e.g., sub-modes 1 and 3 in Table 1), the chroma QPs may be derived in the same way as Cb or Cr blocks. The reconstruction process of the chroma residuals (e.g., resCb and resCr) from the transmitted transform blocks is illustrated in Table 1. If the JCCR mode is activated, a joint chroma residual block (e.g., resJointC[x][y] in Table 1) may be signaled. A residual block for Cb (e.g., resCb) and/or a residual block for Cr (e.g., resCr) may be derived by considering information indicated in one or more syntax elements (e.g., *tu_cbf_cb*, *tu_cbf_cr*, and/or *CSign*). *CSign* may be a sign value specified in the slice header.

[0170] At the encoder side, the joint chroma components may be derived as explained herein. Depending on the mode (e.g., the modes in Table 1), resJointC{1,2} may be generated by the encoder as follows: if the mode is equal to 2 (e.g., single residual with reconstruction $Cb = C, Cr = CSign * C$), the joint residual may be determined according to:

$$resJointC[x][y] = (resCb[x][y] + CSign * resCr[x][y])/2$$

[0171] Otherwise, if the mode is equal to 1 (e.g., single residual with reconstruction $Cb = C, Cr = {(CSign * C)}/{2}$), the joint residual may be determined according to:

$$resJointC[x][y] = (4 * resCb[x][y] + 2 * CSign * resCr[x][y])/5$$

[0172] Otherwise, if the mode is equal to 3 (e.g., single residual, reconstruction $Cr = C, Cb = {(CSign * C)}/{2}$), the joint residual may be determined according to:

$$resJointC[x][y] = (4 * resCr[x][y] + 2 * CSign * resCb[x][y])/5$$

[0173] The value of $CSign$ may be a sign value (e.g., +1 or −1). The sign value may be specified in the slice header. $resJointC[\,][\,]$ may be the transmitted residual.

tu_cbf_cb	tu_cbf_cr	reconstruction of Cb and Cr residuals	mode
1	0	resCb[x][y] = resJointC[x][y] resCr[x][y] = (CSign * resJointC[x][y]) >> 1	1
1	1	resCb[x][y] = resJointC[x][y] resCr[x][y] = Csign * resJointC[x][y]	2
0	1	resCb[x][y] = (Csign * resJointC[x][y]) >> 1 resCr[x][y] = resJointC[x][y]	3

Table 1 - Reconstruction of chroma residuals

[0174] The joint chroma coding sub-modes (e.g., the three joint chroma coding sub-modes described in Table 1) may be (e.g., may only be) supported in I slices. In P and B slices, mode 2 (e.g., only mode 2) may be supported. In P and B slices, the syntax element *tu_joint_cbcr_residual_flag* may be present (e.g., may only be present) if both chroma CBFs are 1.

[0175] The JCCR mode may be combined with the chroma transform skip (TS) mode. To speed up the encoder decision, the JCCR transform selection may depend on whether the independent coding of

Cb and Cr components selects the DCT-2 or the TS as the transform (e.g., as the best transform). The JCCR transform selection may depend on whether there are non-zero coefficients in independent chroma coding. For example, if a chroma component (e.g., one chroma component) selects DCT-2 (or TS) and the other component is zero (e.g., all zero), or if both chroma components select DCT-2 (or TS), then DCT-2 (or TS) (e.g., only DCT-2 or TS) may be considered in JCCR encoding. Otherwise, if one component selects DCT-2 and the other selects TS, then both DCT-2 and TS may be considered in JCCR encoding.

[0176] In some examples, joint Cb-Cr may be able to predict the Cr residual with a simple relationship (e.g., $Cb = -Cr$). CCRM may be used to express more complex relationship between the luma and the chroma (e.g., on the prediction).

[0177] In some examples, joint Cb-Cr may be improved by improving the model used to derive one residual based on another residual. For example, the relationship between residuals may be deduced at the encoder and/or decoder from the prediction of Cb and Cr. In some examples, the relationship between residuals may be deduced at the encoder and/or decoder from the luma. For example, luma residual may be used to predict Cb residuals. For example, luma residuals (and, in some cases, Cb residuals) may be used to predict Cr residuals.

[0178] Feature(s) associated with residual on top of residual prediction are provided herein. Feature(s) associated with signaling (e.g., signaling a QP offset) are provided herein. Feature(s) associated with encoding are provided herein.

[0179] A cross-component residual/prediction model may be used for encoding and/or decoding a video block. A prediction model for cross-component residual prediction may be derived based on the prediction of Cr and the prediction of Cb. The prediction model may be used to predict the residual Cr based on the residual of Cb. The prediction model may be used to predict the residual Cb based on the residual of Cr.

[0180] For example, the relationship f between the residual of Cb and the residual of Cr may be determined (e.g., computed) using the prediction of Cr and Cb, respectively. For example, the residual Cr may be predicted based on the residual of Cb and the relationship f.

[0181] FIG. 14 illustrates an example of Cr residual prediction. FIG. 15 illustrates an example of function support to predict Cr.

[0182] The function f may be a linear combination (e.g., a simple linear combination) of the samples x_i in the prediction of Cb (e.g., using a support around the current sample). FIG. 15 illustrates an example using 9 samples (e.g., labeled a0-a8) of the prediction of Cb to predict the sample y of the prediction of Cr at the center (e.g., at position a4). For example, the prediction of Cr at a position may be calculated as:

$$f(x_i) = \sum_{i=0}^{8} a_i x_i$$

where a_i represents the coefficients of the filter to compute, and x_i represents the value of the prediction Cb at the location of the coefficients a_i.

[0183] An encoder and/or decoder may estimate the values of the coefficients a_i by minimizing:

$$\arg\min \sum_{j} \left| f(x_j) - y_j \right|$$

where the sum over the samples (e.g., all samples) x_j and y_j inside the block are used to determine the coefficients a_i. The norm may be the L2 norm (or, in some cases, the L1 norm).

[0184] In some examples, the function may include an offset. For example, the function may be expressed by:

$$f(x_i) = \sum_{i=0}^{8} a_i x_i + a_9$$

where a9 is a bias term.

[0185] In some examples, the function f may be non-linear. For example, f may be expressed as:

$$f(x_i) = a_4 * x_4 + \sum_{i=0}^{8} a_i clip_r(x_i - x_4) + a_9$$

where the clipped difference between the collocated samples in Cb and its neighbors is used. In some examples, the function f may be expressed using the collocated sample value in Cb. The $clip_r$ function may be express as (e.g., determined by):

$$clip_r(x) = \min\left(\max(x, -r), r\right)$$

[0186] The bound r may be determined based on the best value to predict the samples of the prediction of Cr (e.g., from among a set of possible clipping values). For example, the possible clipping values may be {8, 16, 32, 128, 256, 512, 1024}. In examples, the bounds may be different. For example, the bounds may depend on the coefficient location inside the filter.

[0187] In some examples, the size of the filter support may be different (e.g., larger or smaller).

[0188] In some examples, an advanced relationship may be computed. In some examples, an order 1 (e.g., linear) relationship may be computed. In some examples, a higher order relationship may be computed. For example, an order 2 (e.g., up to square of the values) may be used.

[0189] In some examples, the role of Cb and Cr may be reversed (e.g., inverted). For example, the function f may be computed with Cr samples to predict Cb samples. The choice between Cb or Cr (e.g., whether f is computed based on Cb or Cr) may be signaled in the bitstream.

[0190] Cross-component residual may be predicted using luma.

[0191] FIG. 16 illustrates an example computation of Cb residual from Y residual.

[0192] The Cb residual (and/or Cr residual) may be predicted from the Y residual, as illustrated in FIG. 16. In this case, the prediction may be similar to the prediction of the Cr residual from the Cb residual (e.g., a relationship between the prediction of the Cb (or Cr) residual and the prediction of the Y residual is computed and applied on the Y residual to produce the Cb residual (or Cr residual)).

[0193] In the case of down-sampled chroma (e.g., for 420 YUV format), the set of collocated luma samples to use may be adapted (e.g., down-sampling of luma) to the down-sampled chroma format (e.g., as it is done in CCCM). The phase of the chroma signal with respect to the luma signal may be used to compute the collocated luma samples.

[0194] FIG. 17 illustrates an example of Cr residual derivation from Y and Cb residual.

[0195] The Cr residual may be predicted after the Cb residual has been predicted from the luma residuals. The Cr residual may be predicted using the luma and Cb residual (e.g., both luma and Cb residual). For example, the Cr residual may be calculated by:

$$f\left(x_i^L, x_j^{Cb}\right) = \sum_{i=0}^{N-1} a_i x_i^L + \sum_{i=0}^{M-1} a'_i x_i^{Cb}$$

where N is the number of samples used in the luma residual and M the number of samples used in the Cb residual. In some examples (e.g., for the same filter size and a 420 YUV format), a larger number of luma samples may be used (e.g., compared to the Cb samples). The luma residuals may be down-sampled (e.g., down-sampled first).

[0196] FIG. 18 illustrates example luma/Cb coefficient locations. FIG. 18 illustrates example locations of luma (ai) coefficients and Cb (a’i) coefficients (e.g., in the case of collocated chroma samples). As illustrated in FIG. 18, 9 Cb coefficients may be used (e.g., M = 9), and 49 luma coefficients may be used (e.g., N = 49).

[0197] A sub-sampled set (e.g., only a sub-sampled set) of the luma samples may be used. For example, the sub-sampled set of luma samples may include coefficients at locations {a8, a10, a12, a22, a24, a26, a36, a38, a40}.

[0198] In some examples, the luma samples may be down-sampled (e.g., first down-sampled) and used in the function f. For example, the average of luma samples at locations {a0, a1, a7, a8} may be used in the equation. The number of luma coefficients and the number of chroma coefficients may be the same.

[0199] Residual on top of residual prediction may be used. FIG. 19 illustrates an example of adding residual on the residual prediction.

[0200] In some examples, a residual on top of residual prediction may be added to the residual prediction mode (e.g., as illustrated in FIG. 19). In this case, the residual of residual (e.g., resresCr) may be computed as the difference between the original samples and the sum of the prediction Cr and the prediction of the residual Cr.

[0201] The residual on top of residual prediction may be applied to modes described herein (e.g., prediction of Cb residual from luma, prediction of Cr from luma and Cr residuals, etc.).

[0202] Feature(s) associated with signaling are provided herein. At block level, the chroma residual prediction may be signaled (e.g., after decoding the CBF of Cb and Cr). In some cases, the flag *predicted_residual* may be coded. For example, the flag *predicted_residual* may be coded if one or more of the following are true: both flags are true (e.g., a residual exists for both components), the mode is not *joint_cb_cr*, and/or the block is in inter mode or IBC/IntraTMP mode (e.g., intra modes where the prediction is computed from a motion compensation of the block). In some examples, the *predicted_residual* mode may replace the mode *joint_cb_cr* (e.g., if the block is inter coded or uses the IBC/IntraTMP mode).

[0203] If a residual is added on top of the residual prediction, the signaling may be inverted:

[0204] The flag *predicted_residual* may be coded (e.g., may be coded first). The CBF of the component for which a residual prediction is used may be signaled (e.g., may then be signaled). The CBF may indicate if a residual on top of the residual prediction is coded in the bitstream.

[0205] In some examples, the QP used to quantize the residual of the Cb component (e.g., if the residual prediction mode is used) may be adapted. For example, the QP for Cb may be computed as:

$$QP_{Cb} = QP_{luma} + QP_{CbOffset} + QP_{Cboffset2}$$

where $QPCbOffset2$ may be used if (e.g., only if) the residual prediction mode is used and signaled (e.g., at slice level).

[0206] Feature(s) associated with encoding are provided herein.

[0207] The encoder may compute the Cb residual R' as follows (e.g., in order to minimize the following distortion and optimally reconstruct the chroma components):

$$\min_{R'} \sum |Pred_{Cb} + R' - O_{Cb}| + \sum |Pred_{Cr} + f(R') - O_{Cr}|$$

where R' is the Cb residual to encode, $Predx$ are the predictions, O_x are the original samples, and $f(R')$ is the Cb residual prediction. The sum may be performed on samples (e.g., all samples) of the block. The norm may be, for example, an L2 norm.

[0208] The distortion may be computed (e.g., directly) using the original residuals:

$$\min_{R'} \sum |R_{Cb} - R'| + \sum |R_{Cr} - f(R')|$$

[0209] Although features and elements are described above in particular combinations, one of ordinary skill in the art will appreciate that each feature or element can be used alone or in any combination with the other features and elements. In addition, the methods described herein may be implemented in a computer program, software, or firmware incorporated in a computer-readable medium for execution by a computer or processor. Examples of computer-readable media include electronic signals (transmitted over wired or wireless connections) and computer-readable storage media. Examples of computer-readable storage media include, but are not limited to, a read only memory (ROM), a random access memory (RAM), a register, cache memory, semiconductor memory devices, magnetic media such as internal hard disks and removable disks, magneto-optical media, and optical media such as CD-ROM disks, and digital versatile disks (DVDs). A processor in association with software may be used to implement a radio frequency transceiver for use in a WTRU, UE, terminal, base station, RNC, or any host computer.

CLAIMS

What is Claimed:

1. A device for video decoding, the device comprising:
a processor configured to:
estimate a relationship between residual of a first chrominance component of a block and residual of a second chrominance component of the block based on a prediction of the first chrominance component and a prediction of the second chrominance component;
predict the residual of the first chrominance component based on the estimated relationship and the residual of the second chrominance component; and
decode the block based on the predicted residual of the first chrominance component.

2. The device of claim 1, wherein the block is a first block, the estimated relationship is a first estimated relationship, and the processor is further configured to:
estimate a second relationship between residual of the first chrominance component of a second block and residual of a luminance component of the second block based on a prediction of the first chrominance component of the second block and a prediction of the luminance component of the second block; and
predict the residual of the first chrominance component of the second block based on the second estimated relationship and the residual of the luminance component of the second block; and
decode the second block based, at least in part, on the predicted residual of the first chrominance component of the second block.

3. The device of claim 2, wherein the processor is further configured to predict residual of the second chrominance component of the second block based the residual of the first chrominance component of the second block and the residual of the luminance component of the second block, wherein the processor being configured to decode the second block based, at least in part, on the predicted residual of the first chrominance component of the second block comprises the processor being configured to decode the second block based on the predicted residual of the first chrominance component of the second block and the predicted residual of the second chrominance component of the second block.

4. The device of claim 2, wherein the processor is further configured to compute residual of the predicted residual of the first chrominance component of the second block based on a difference between an original sample of the first chrominance component of the second block and a sum of the prediction of the first chrominance component of the second block and the predicted residual of the first chrominance component of the second block, wherein the processor being configured to decode the second block based, at least in part, on the predicted residual of the first chrominance component of the second block comprises the processor being configured to decode the second block based on the predicted residual of the first chrominance component of the second block and the residual of the predicted residual of the first chrominance component of the second block.

5. The device of any one of claims 1-4, wherein the processor is further configured to compute residual of the predicted residual of the first chrominance component based on a difference between an original sample of the first chrominance component and a sum of the prediction of the first chrominance component and the predicted residual of the first chrominance component, wherein the processor being configured to decode the second block based, at least in part, on the predicted residual of the first chrominance component of the second block comprises the processor being configured to decode the second block based on the predicted residual of the first chrominance component of the second block and the residual of the predicted residual of the first chrominance component.

6. The device of any one of claims 1-5, wherein the first chrominance component comprises a red-difference chrominance component, and wherein the second chrominance component comprises a blue-difference chrominance component.

7. The device of any one of claims 1-6, wherein the processor is further configured to send a flag to an encoder, wherein the flag indicates the predicted residual of the first chrominance component.

8. A method for video decoding, the method comprising:

estimating a relationship between residual of a first chrominance component of a block and residual of a second chrominance component of the block based on a prediction of the first chrominance component and a prediction of the second chrominance component;

predicting the residual of the first chrominance component based on the estimated relationship and the residual of the second chrominance component; and

decoding the block based on the predicted residual of the first chrominance component.

9. The method of claim 8, wherein the block is a first block, the estimated relationship is a first estimated relationship, and the method further comprises:

estimating a second relationship between residual of the first chrominance component of a second block and residual of a luminance component of the second block based on a prediction of the first chrominance component of the second block and a prediction of the luminance component of the second block; and

predicting the residual of the first chrominance component of the second block based on the second estimated relationship and the residual of the luminance component of the second block; and

decoding the second block based, at least in part, on the predicted residual of the first chrominance component of the second block.

10. The method of claim 9, wherein the method further comprises predicting residual of the second chrominance component of the second block based the residual of the first chrominance component of the second block and the residual of the luminance component of the second block, wherein decoding the second block is further based on the predicted residual of the second chrominance component of the second block.

11. The method of claim 9, wherein the method further comprises computing residual of the predicted residual of the first chrominance component of the second block based on a difference between an original sample of the first chrominance component of the second block and a sum of the prediction of the first chrominance component of the second block and the predicted residual of the first chrominance component of the second block, wherein decoding the second block is further based on the residual of the predicted residual of the first chrominance component of the second block.

12. The method of any one of claims 8-11, wherein the method further comprises computing residual of the predicted residual of the first chrominance component based on a difference between an original sample of the first chrominance component and a sum of the prediction of the first chrominance component and the predicted residual of the first chrominance component, wherein decoding the block is further based on the residual of the predicted residual of the first chrominance component.

13. The method of any one of claims 8-12, wherein the first chrominance component comprises a red-difference chrominance component, and wherein the second chrominance component comprises a blue-difference chrominance component.

14. The method of any one of claims 8-13, wherein the method further comprises sending a flag to an encoder, wherein the flag indicates the predicted residual of the first chrominance component.

15. A device for video encoding, the device comprising:
a processor configured to:
estimate a relationship between residual of a first chrominance component and residual of a second chrominance component based on a prediction of the first chrominance component and a prediction of the second chrominance component;
predict the residual of the first chrominance component based on the estimated relationship and the residual of the second chrominance component; and
encode a block based on the predicted residual of the first chrominance component.

16. The device of claim 15, wherein the block is a first block, the estimated relationship is a first estimated relationship, and the processor is further configured to:
estimate a second relationship between residual of the first chrominance component of a second block and residual of a luminance component of the second block based on a prediction of the first chrominance component of the second block and a prediction of the luminance component of the second block; and
predict the residual of the first chrominance component of the second block based on the second estimated relationship and the residual of the luminance component of the second block; and
encode the second block based, at least in part, on the predicted residual of the first chrominance component of the second block.

17. The device of claim 16, wherein the processor is further configured to predict residual of the second chrominance component of the second block based the residual of the first chrominance component of the second block and the residual of the luminance component of the second block, wherein the processor being configured to encode the second block based, at least in part, on the predicted residual of the first chrominance component of the second block comprises the processor being configured to encode the second block based on the predicted residual of the first

chrominance component of the second block and the predicted residual of the second chrominance component of the second block.

18. The device of claim 16, wherein the processor is further configured to compute residual of the predicted residual of the first chrominance component of the second block based on a difference between an original sample of the first chrominance component of the second block and a sum of the prediction of the first chrominance component of the second block and the predicted residual of the first chrominance component of the second block, wherein the processor being configured to encode the second block based, at least in part, on the predicted residual of the first chrominance component of the second block comprises the processor being configured to encode the second block based on the predicted residual of the first chrominance component of the second block and the residual of the predicted residual of the first chrominance component of the second block.

19. The device of any one of claims 15-18, wherein the processor is further configured to compute residual of the predicted residual of the first chrominance component based on a difference between an original sample of the first chrominance component and a sum of the prediction of the first chrominance component and the predicted residual of the first chrominance component, wherein the processor being configured to encode the second block based, at least in part, on the predicted residual of the first chrominance component of the second block comprises the processor being configured to encode the second block based on the predicted residual of the first chrominance component of the second block and the residual of the predicted residual of the first chrominance component.

20. The device of any one of claims 15-19, wherein the first chrominance component comprises a red-difference chrominance component, and wherein the second chrominance component comprises a blue-difference chrominance component.

21. The device of any one of claims 15-20, wherein the processor is further configured to send a flag to a decoder, wherein the flag indicates the predicted residual of the first chrominance component.

22. A method for video encoding comprising:

estimating a relationship between residual of a first chrominance component and residual of a second chrominance component based on a prediction of the first chrominance component and a prediction of the second chrominance component;

predicting the residual of the first chrominance component based on the estimated relationship and the residual of the second chrominance component; and

encoding a block based on the predicted residual of the first chrominance component.

23. The method of claim 22, wherein the block is a first block, the estimated relationship is a first estimated relationship, and the method further comprises:

estimating a second relationship between residual of the first chrominance component of a second block and residual of a luminance component of the second block based on a prediction of the first chrominance component of the second block and a prediction of the luminance component of the second block; and

predicting the residual of the first chrominance component of the second block based on the second estimated relationship and the residual of the luminance component of the second block; and

encoding the second block based, at least in part, on the predicted residual of the first chrominance component of the second block.

24. The method of claim 23, wherein the method further comprises predicting residual of the second chrominance component of the second block based the residual of the first chrominance component of the second block and the residual of the luminance component of the second block, wherein encoding the second block is further based on the predicted residual of the second chrominance component of the second block.

25. The method of claim 23, wherein the method further comprises computing residual of the predicted residual of the first chrominance component of the second block based on a difference between an original sample of the first chrominance component of the second block and a sum of the prediction of the first chrominance component of the second block and the predicted residual of the first chrominance component of the second block, wherein encoding the second block is further based on the residual of the predicted residual of the first chrominance component of the second block.

26. The method of any one of claims 22-25, wherein the method further comprises computing residual of the predicted residual of the first chrominance component based on a difference between an original sample of the first chrominance component and a sum of the prediction of the first chrominance component and the predicted residual of the first chrominance component, wherein

encoding the block is further based on the residual of the predicted residual of the first chrominance component.

27. The method of any one of claims 22-26, wherein the first chrominance component comprises a red-difference chrominance component, and wherein the second chrominance component comprises a blue-difference chrominance component.

28. The method of any one of claims 22-27, wherein the method further comprises sending a flag to a decoder, wherein the flag indicates the predicted residual of the first chrominance component.

29. Video data comprising information representative of the encoded block generated according to the method of any one of claims 22-28.

100

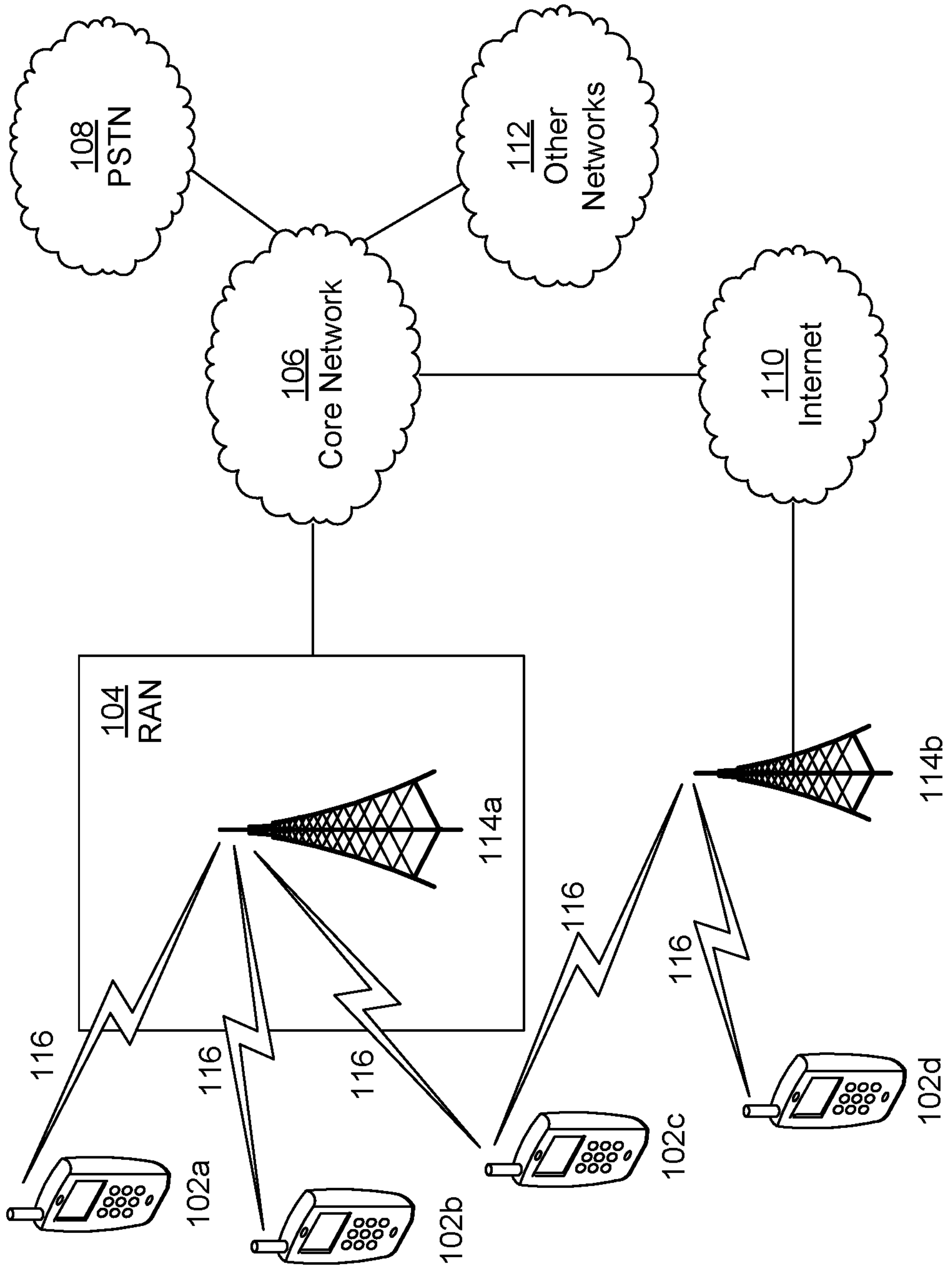

108
PSTN
112
Other
Networks
106
Core Network
110
Internet
104
RAN
114a
114b
116
116
116
116
116
102a
102b
102c
102d

FIG. 1A

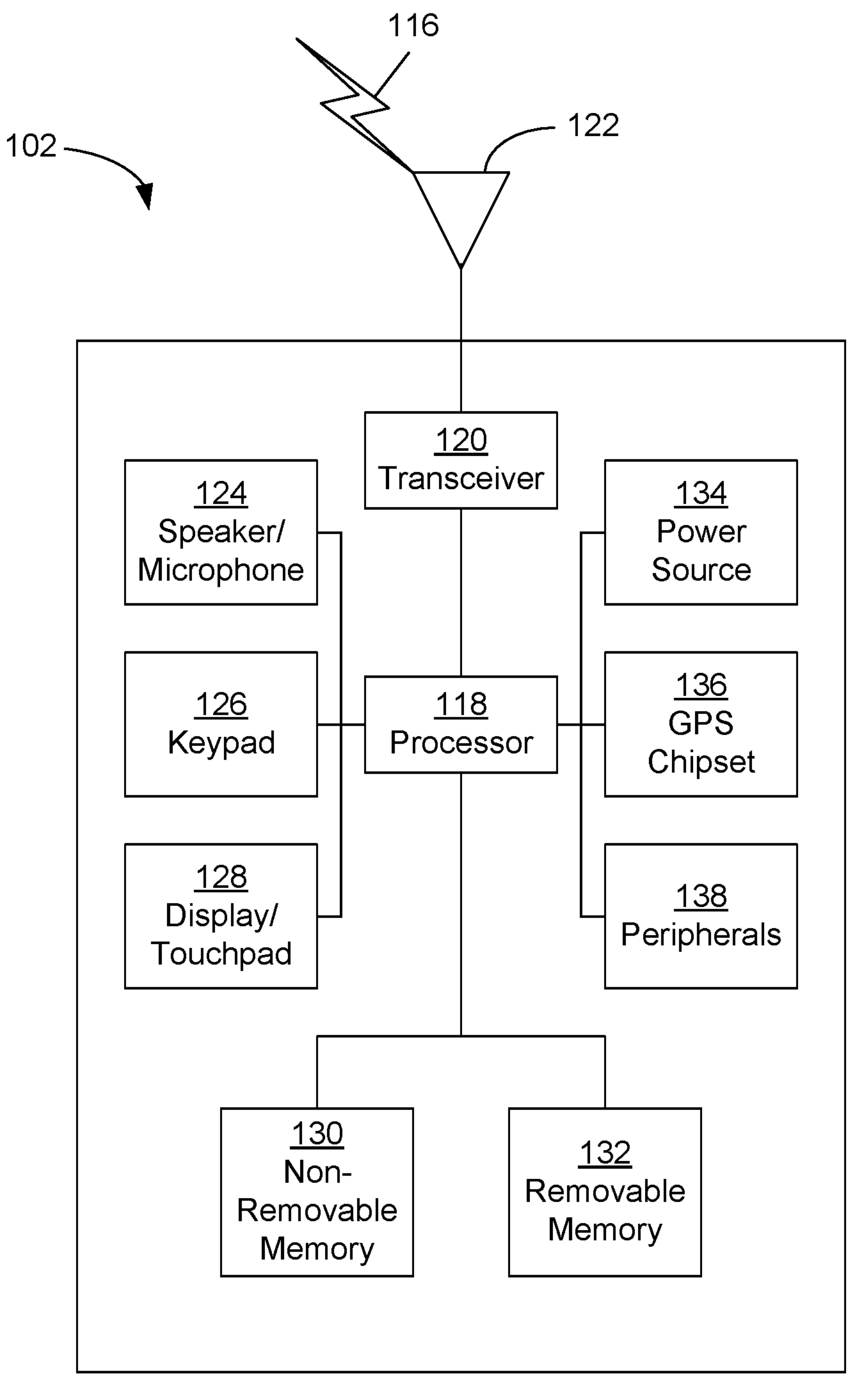

116
102
122
120
Transceiver
124
Speaker/
Microphone
134
Power
Source
126
Keypad
118
Processor
136
GPS
Chipset
128
Display/
Touchpad
138
Peripherals
130
Non-
Removable
Memory
132
Removable
Memory

FIG. 1B

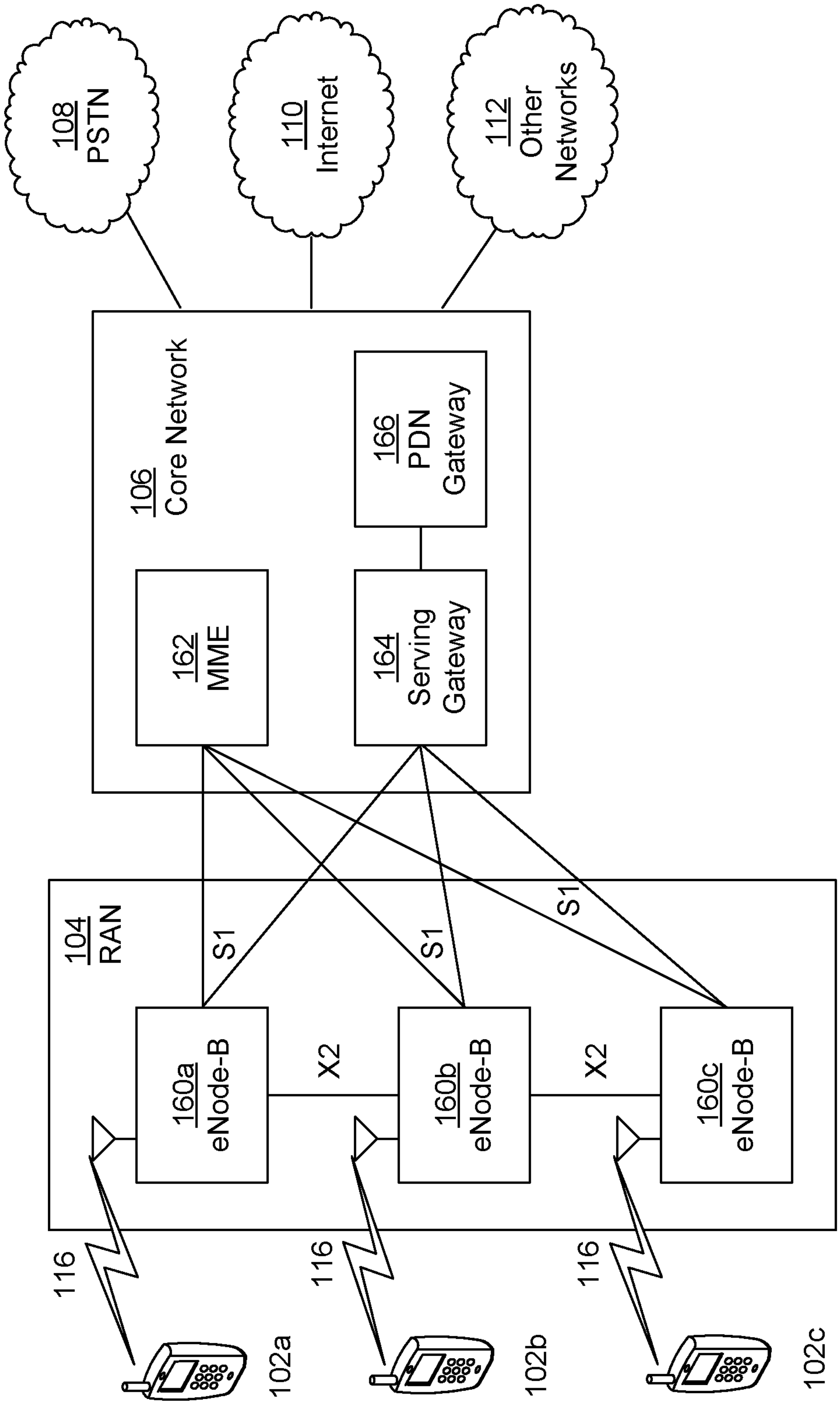

108
PSTN
110
Internet
112
Other
Networks
106
Core Network
162
MME
164
Serving
Gateway
166
PDN
Gateway
104
RAN
S1
S1
S1
160a
eNode-B
160b
eNode-B
160c
eNode-B
X2
X2
116
116
116
102a
102b
102c

FIG. 1C

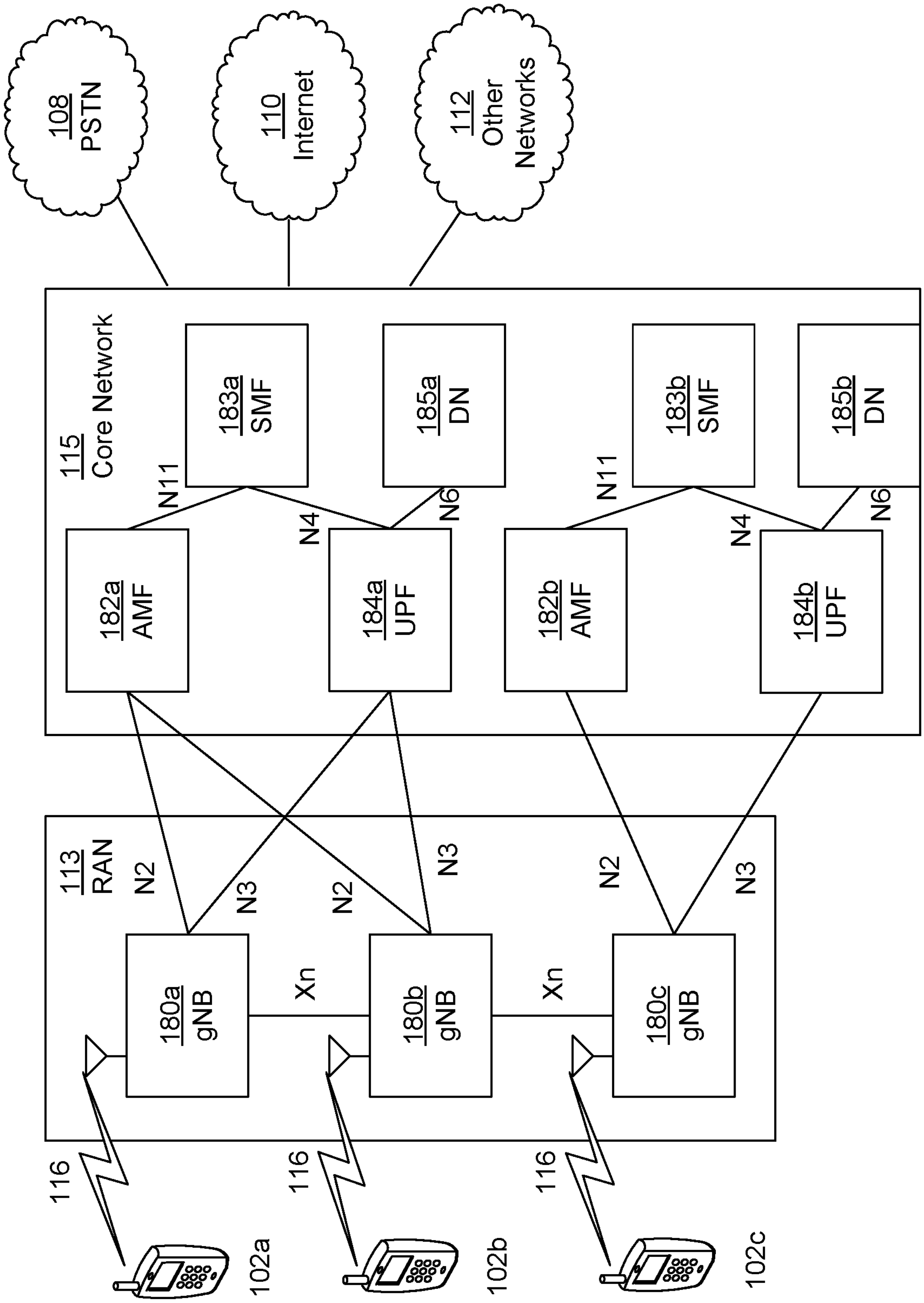
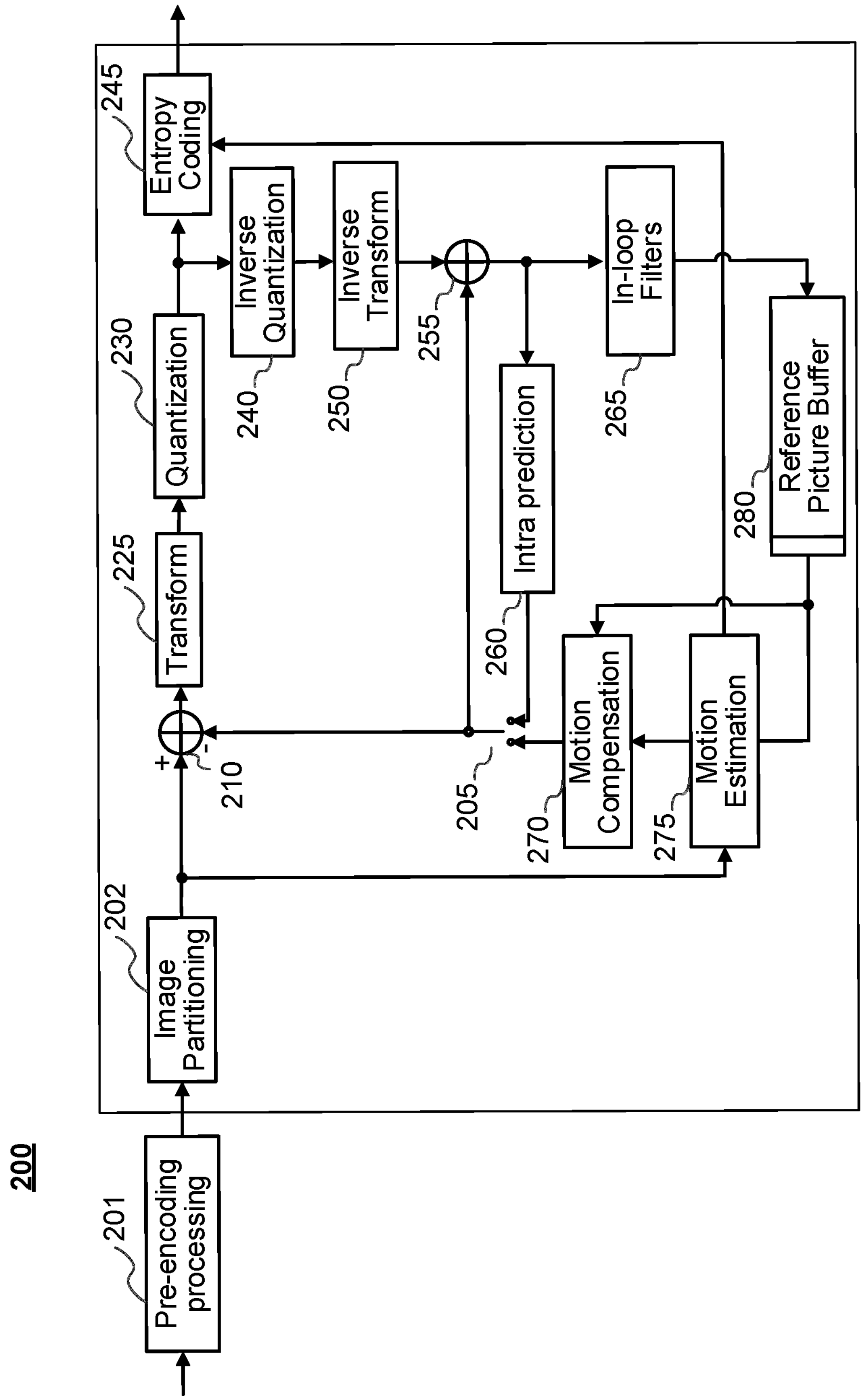

108
PSTN
110
Internet
112
Other Networks
115
Core Network
183a
SMF
185a
DN
183b
SMF
185b
DN
182a
AMF
184a
UPF
182b
AMF
184b
UPF
N11
N4
N6
N11
N4
N6
113
RAN
N2
N3
N2
N3
N2
N3
180a
gNB
180b
gNB
180c
gNB
Xn
Xn
116
116
116
102a
102b
102c

FIG. 1D



200
Pre-encoding processing
201
Image Partitioning
202
Transform
225
Quantization
230
Entropy Coding
245
Inverse Quantization
240
Inverse Transform
250
255
Intra prediction
260
In-loop Filters
265
Reference Picture Buffer
280
Motion Compensation
270
Motion Estimation
275
205
210
+
-

FIG. 2

300

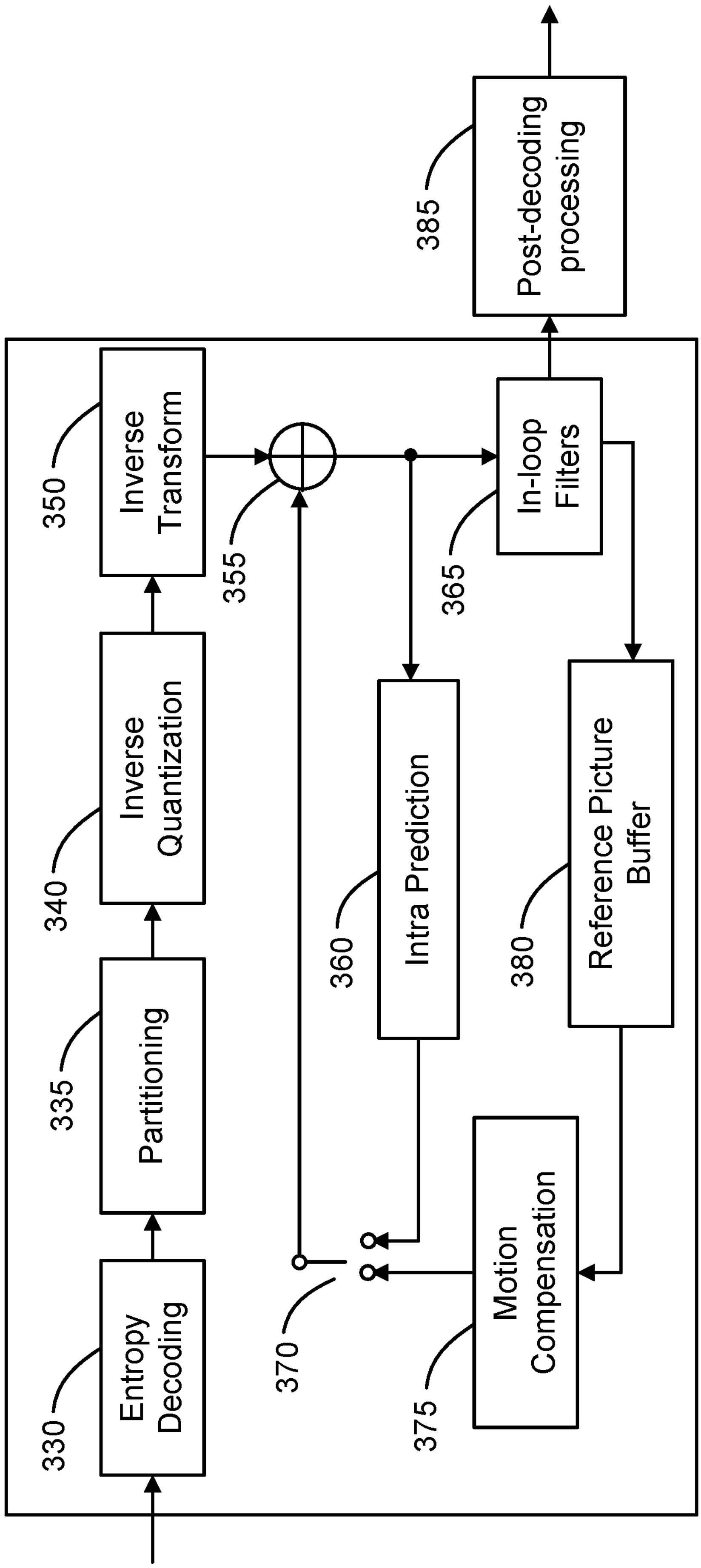

330
Entropy Decoding
335
Partitioning
340
Inverse Quantization
350
Inverse Transform
355
360
Intra Prediction
365
In-loop Filters
370
375
Motion Compensation
380
Reference Picture Buffer
385
Post-decoding processing

FIG. 3

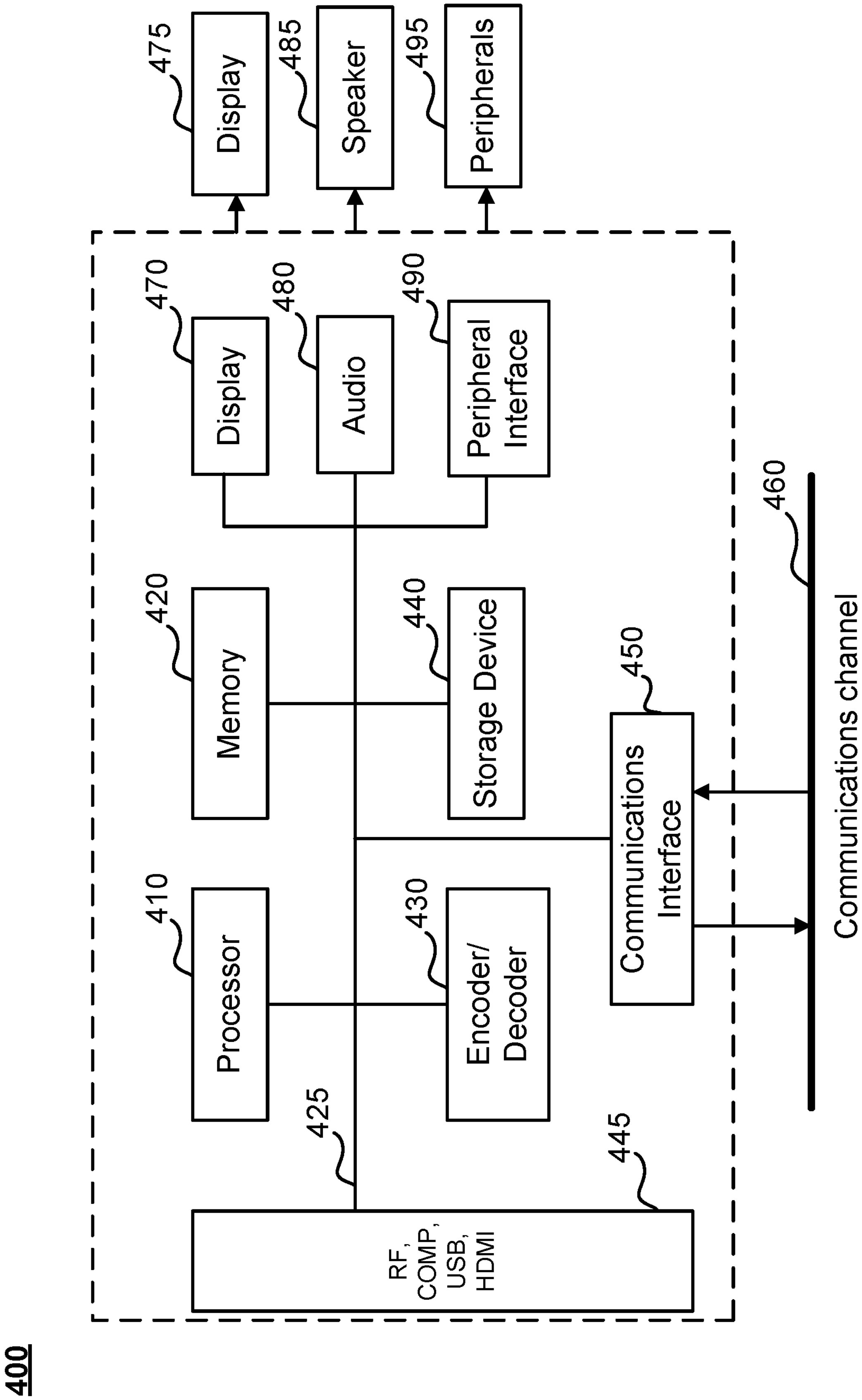

400
475
Display
485
Speaker
495
Peripherals
470
Display
480
Audio
490
Peripheral Interface
420
Memory
440
Storage Device
410
Processor
430
Encoder/ Decoder
450
Communications Interface
460
Communications channel
425
445
RF, COMP, USB, HDMI

FIG. 4

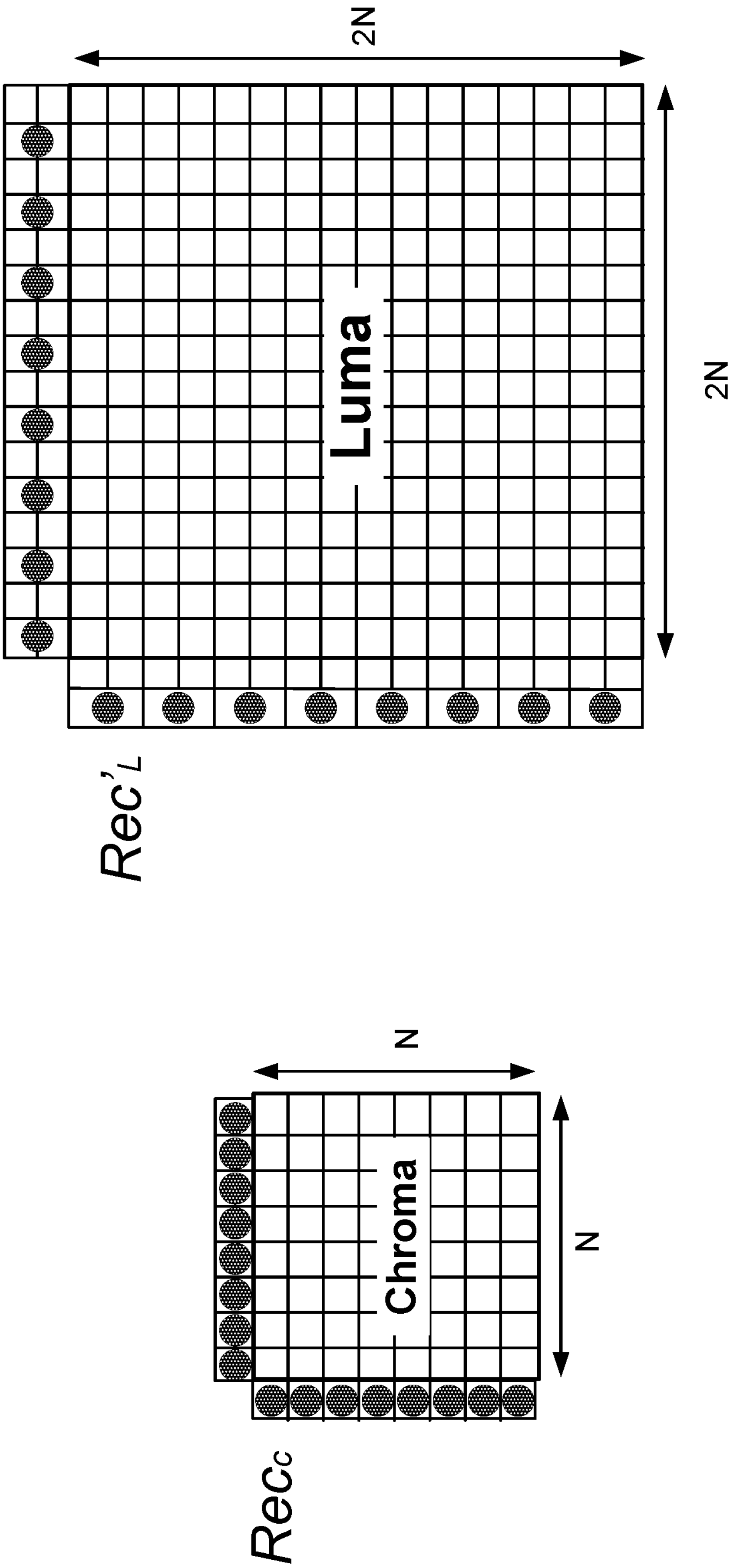


FIG. 5

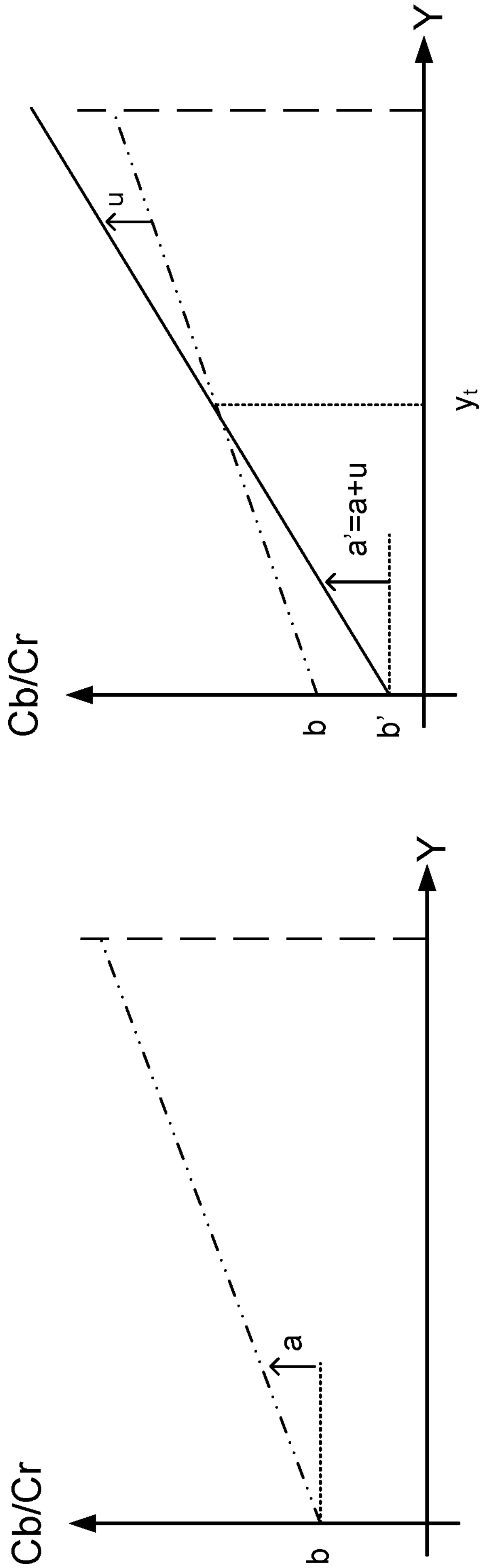


FIG. 6

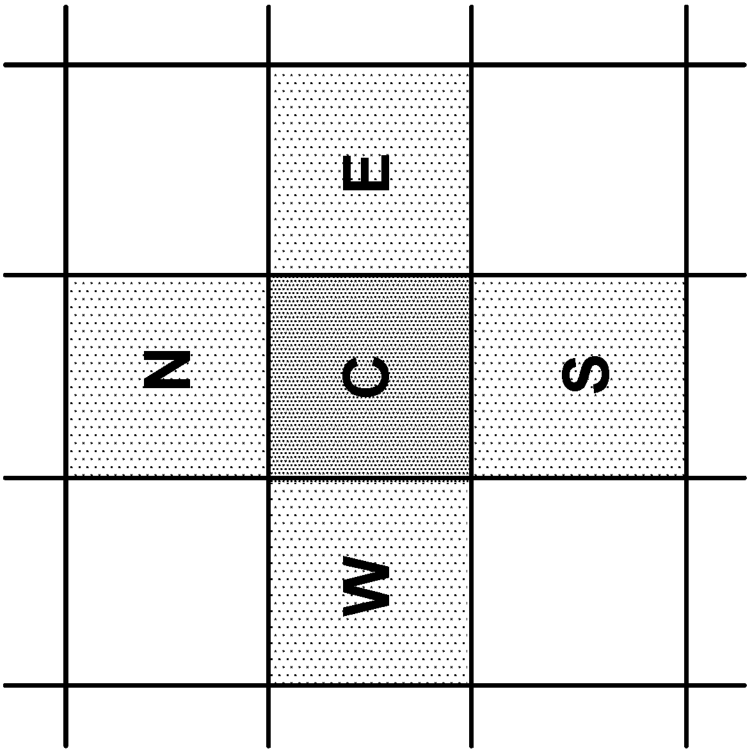


FIG. 7

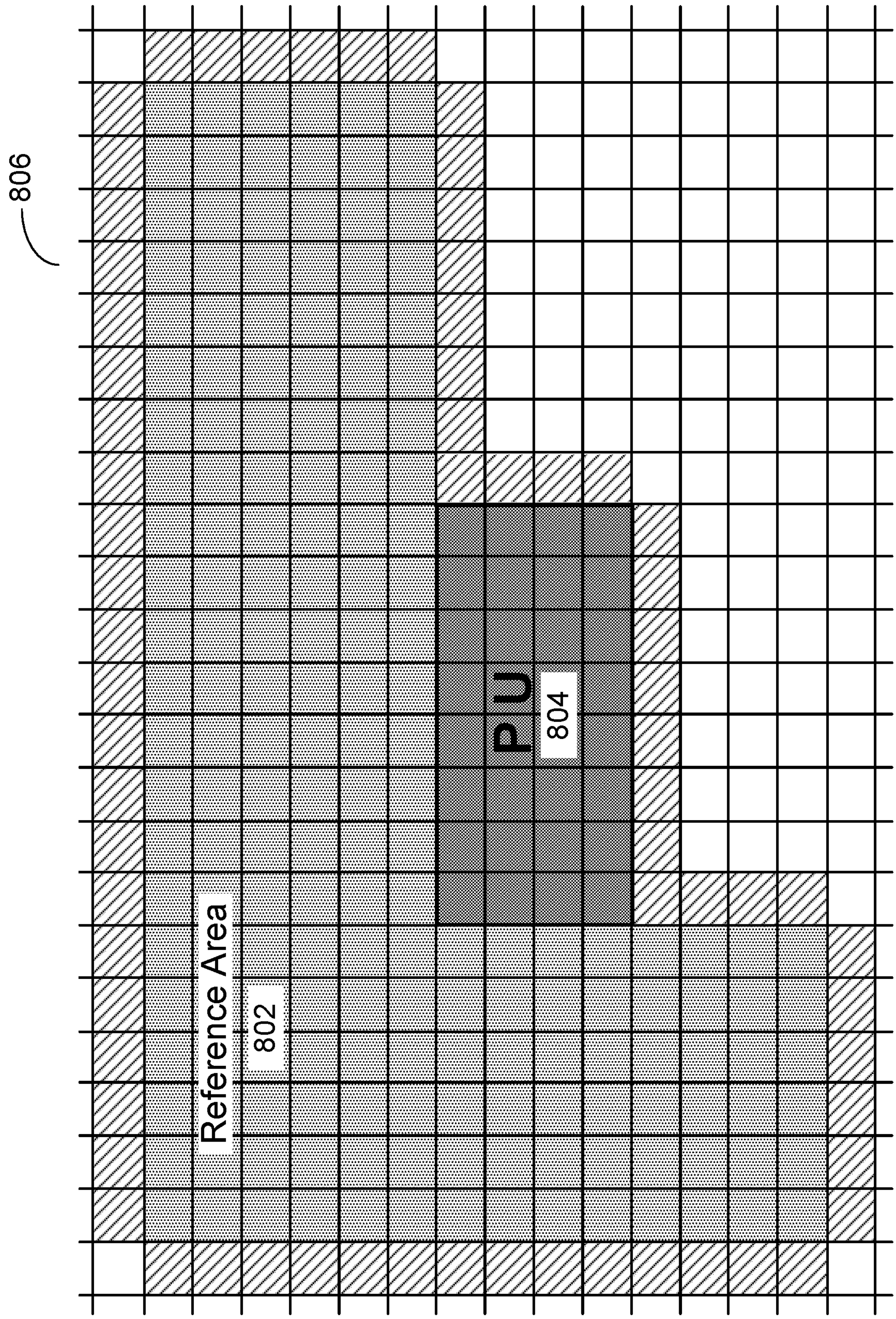


FIG. 8

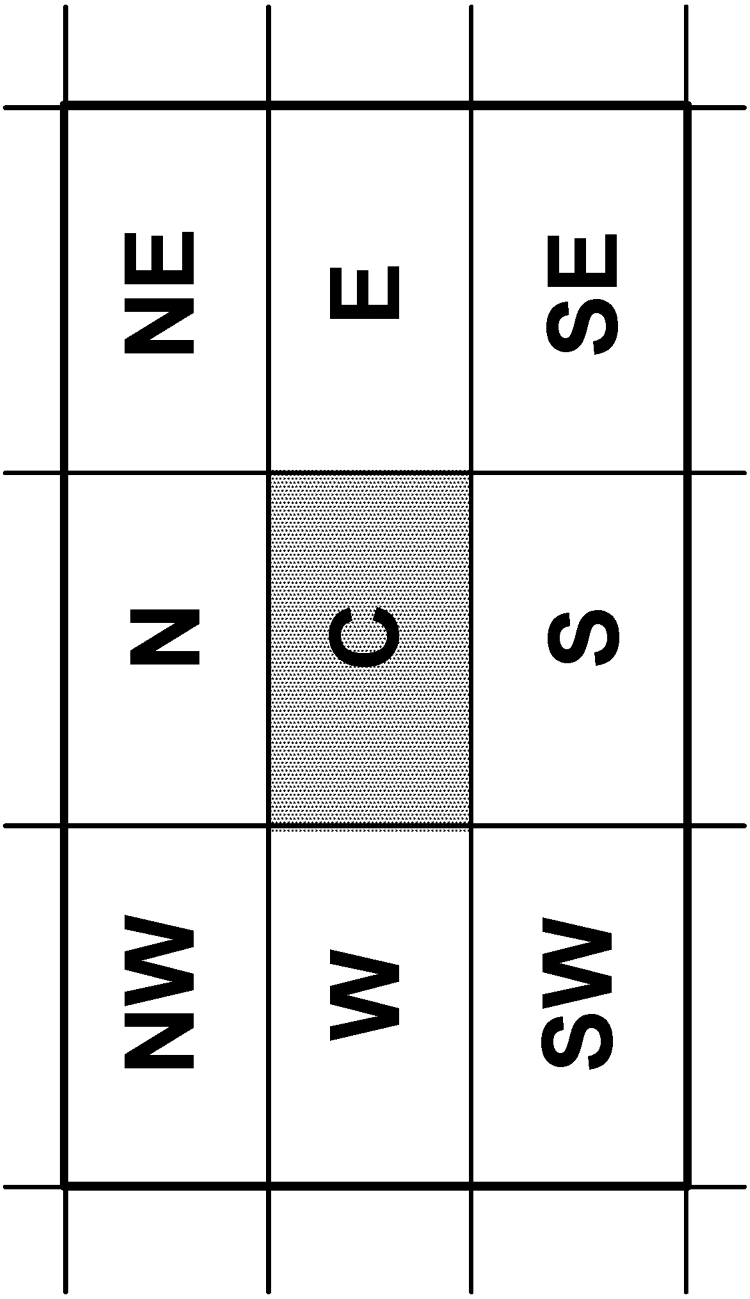


FIG. 9

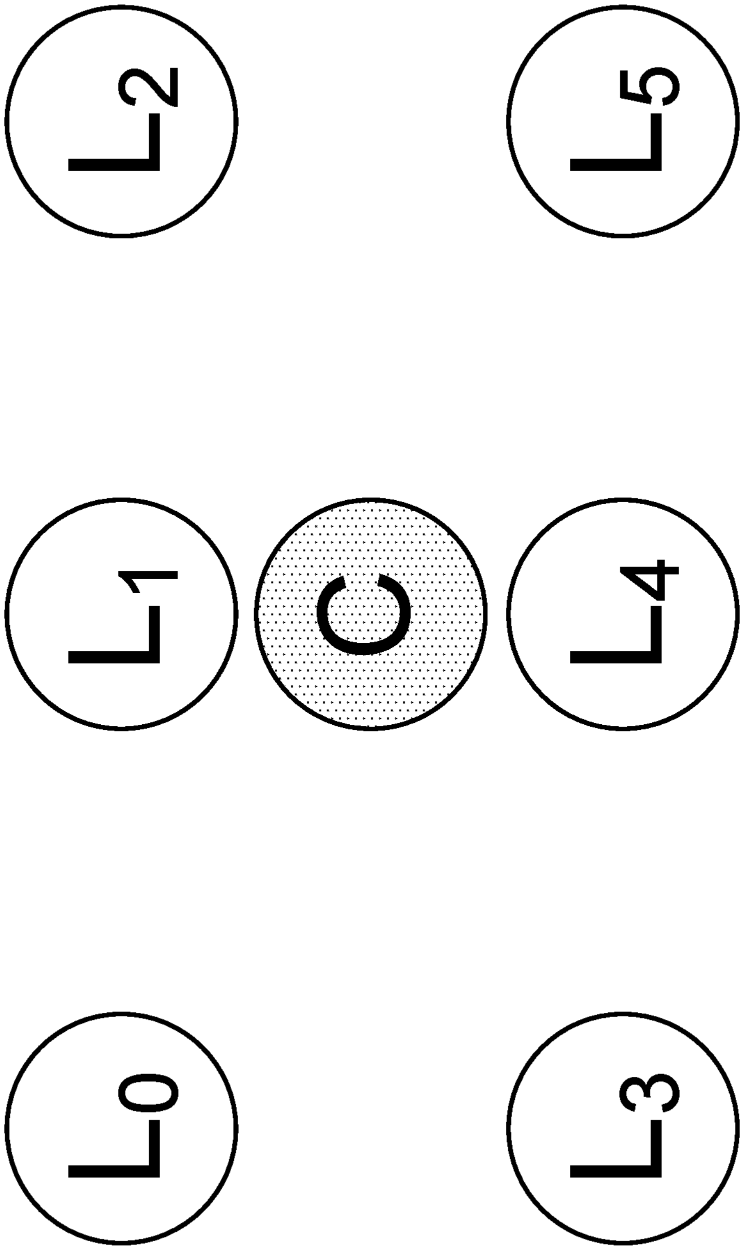


FIG. 10

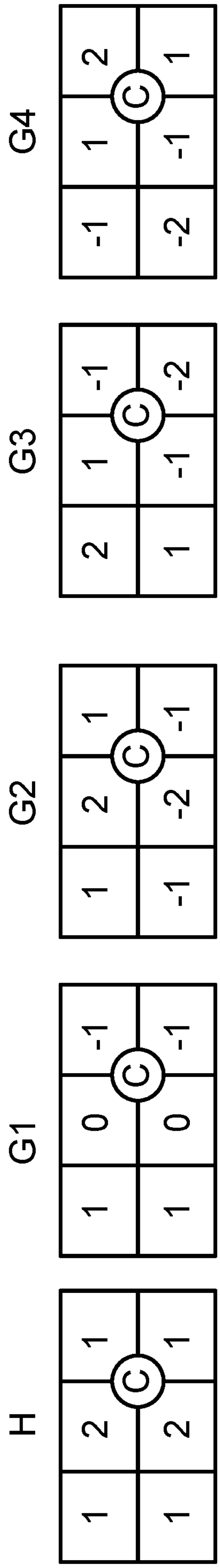


FIG. 11

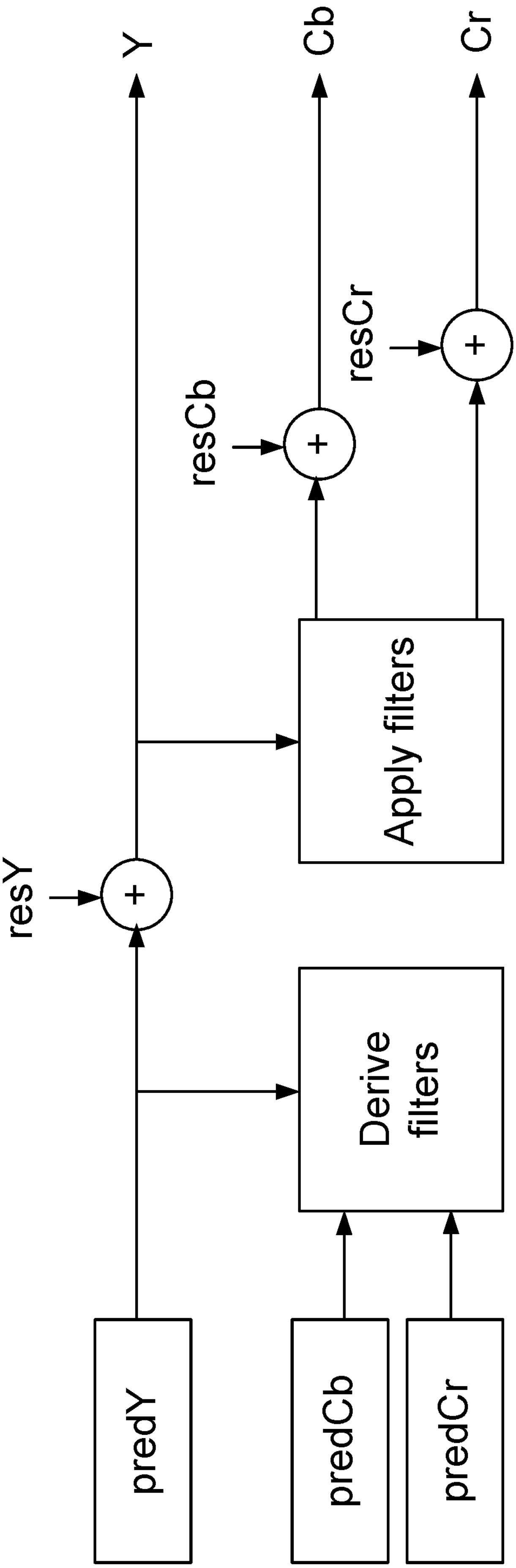

Y
Cb
Cr
resCb
resCr
resY
Apply filters
Derive filters
predY
predCb
predCr

FIG. 12

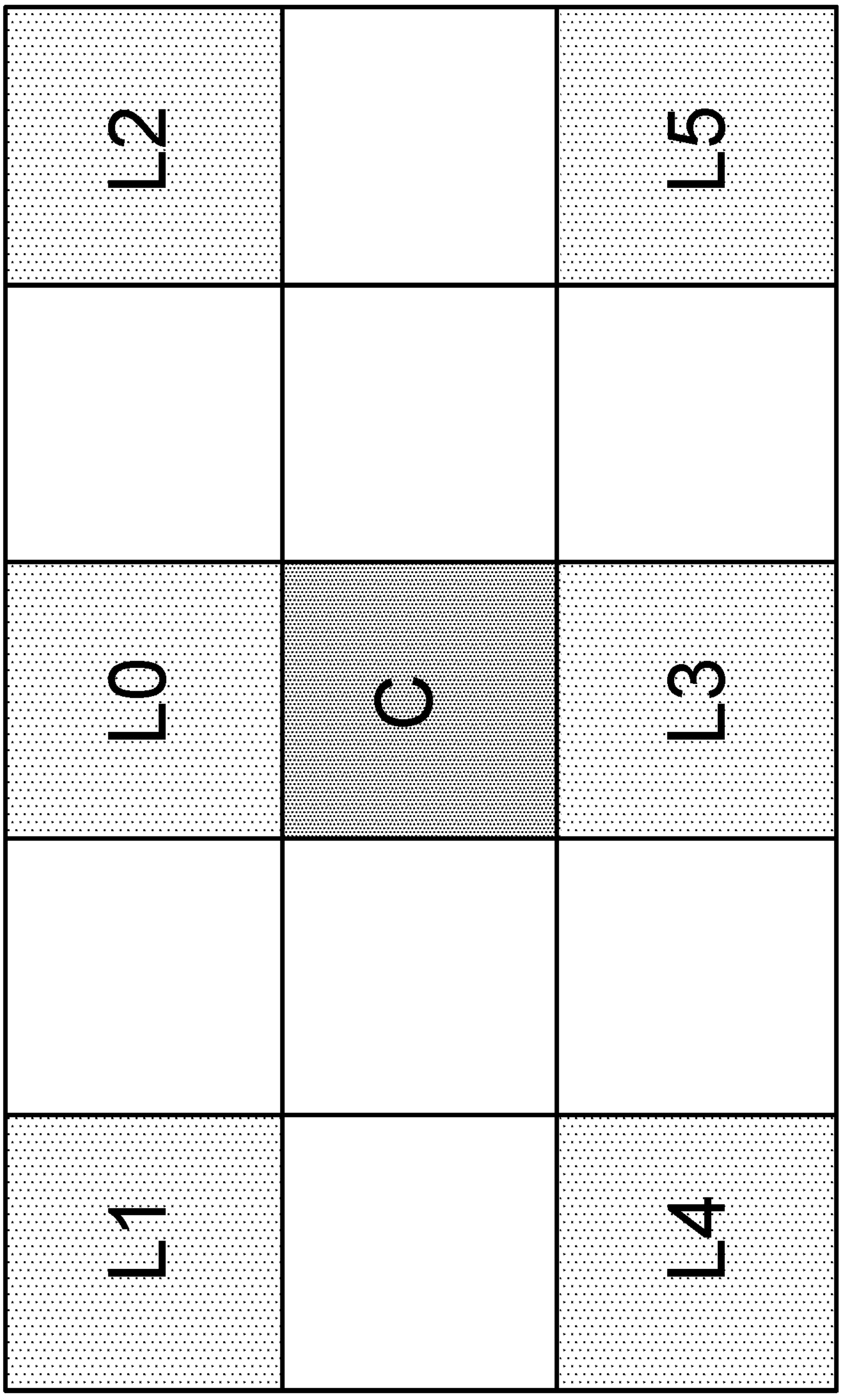


FIG. 13

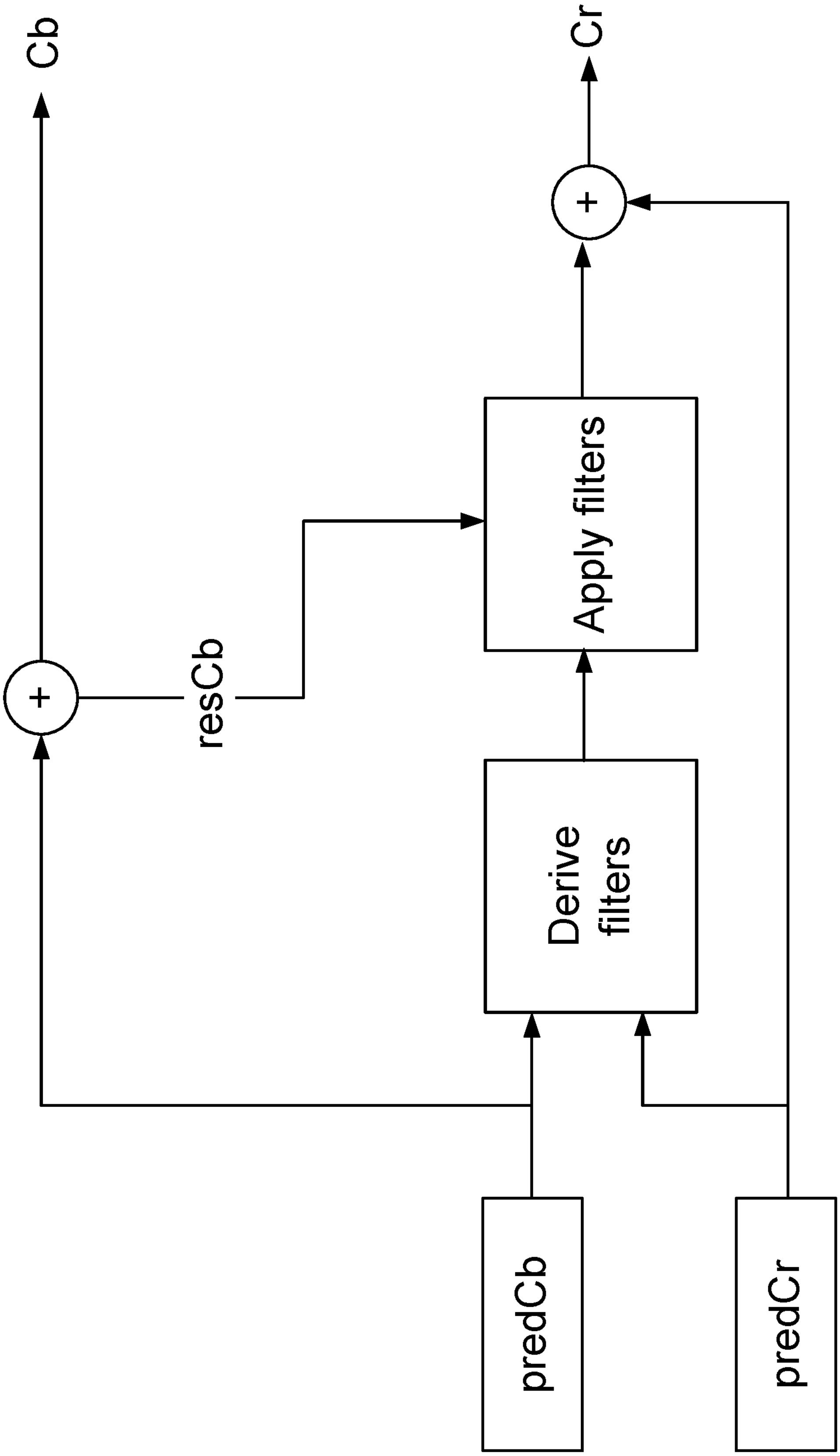


FIG. 14

a0	a1	a2
a3	a4	a5
a6	a7	a8

FIG. 15

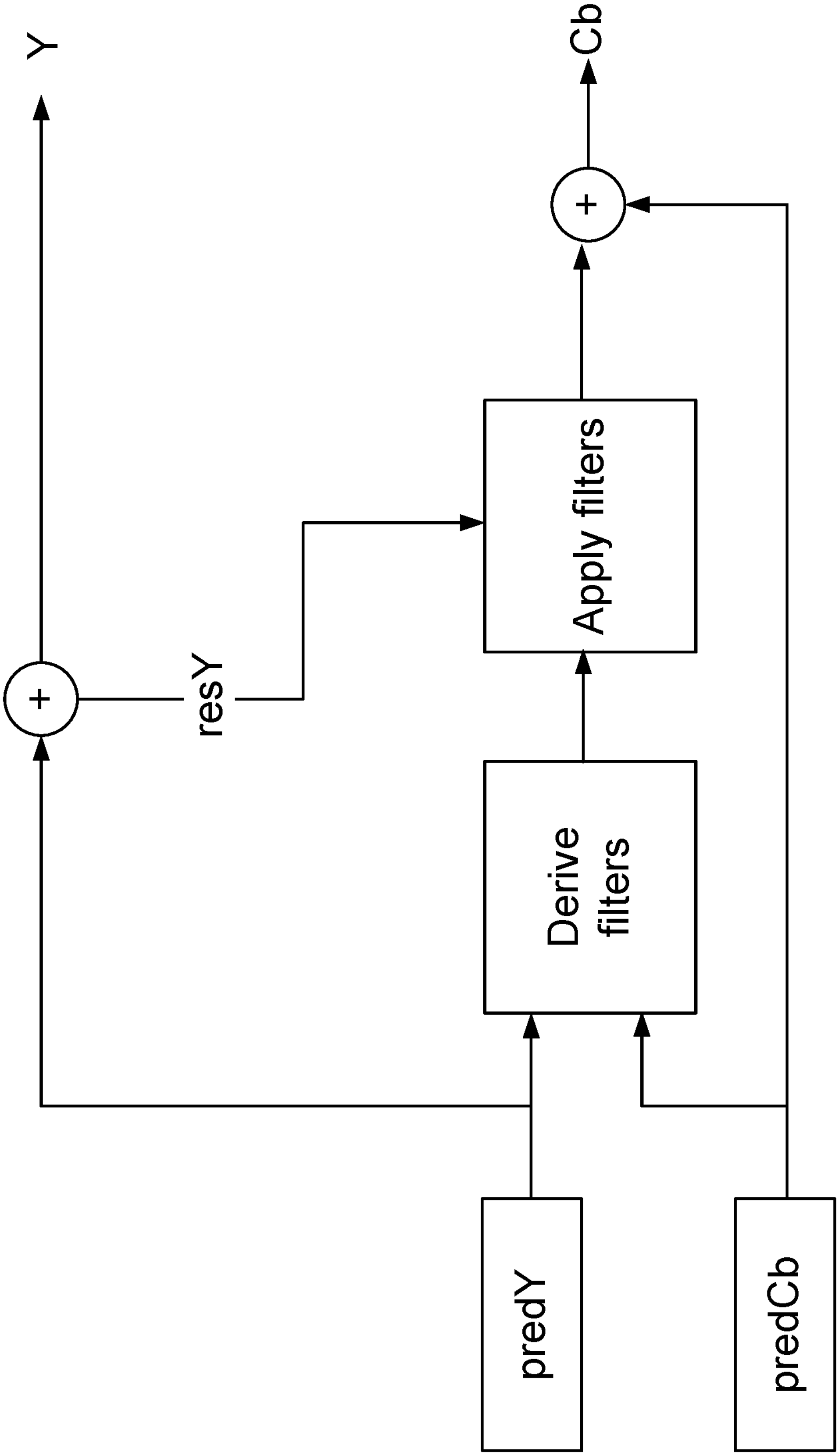


FIG. 16

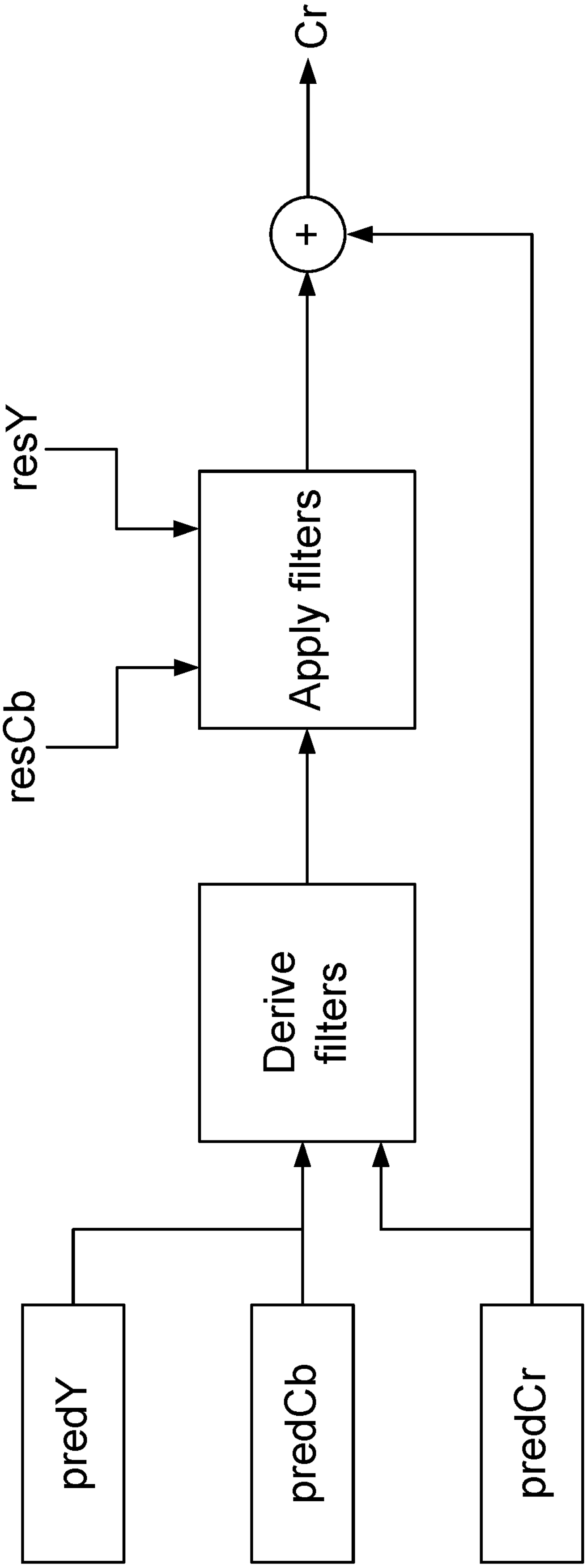

Cr
+
resY
resCb
Apply filters
Derive filters
predY
predCb
predCr

FIG. 17

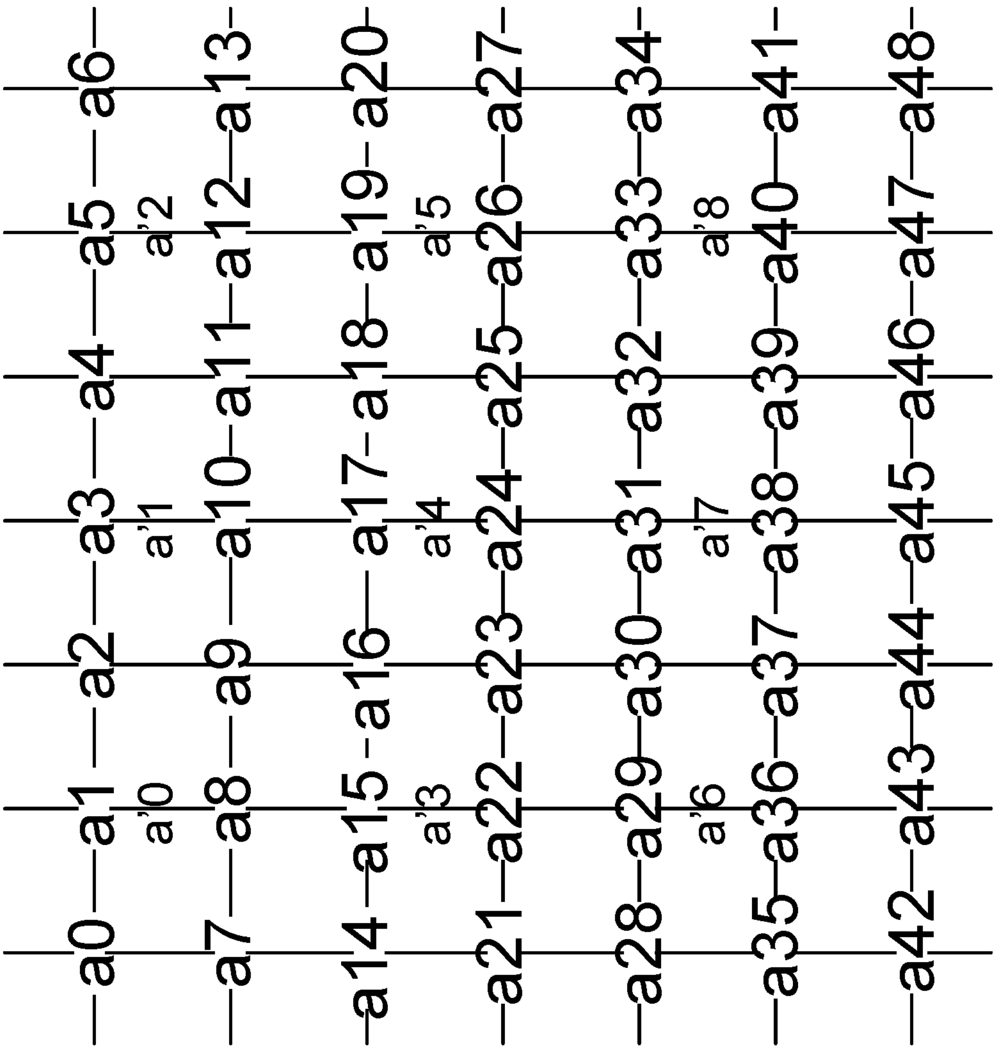

a0 a1 a2 a3 a4 a5 a6
a7 a8 a9 a10 a11 a12 a13
a14 a15 a16 a17 a18 a19 a20
a21 a22 a23 a24 a25 a26 a27
a28 a29 a30 a31 a32 a33 a34
a35 a36 a37 a38 a39 a40 a41
a42 a43 a44 a45 a46 a47 a48
a'0 a'1 a'2
a'3 a'4 a'5
a'6 a'7 a'8

FIG. 18

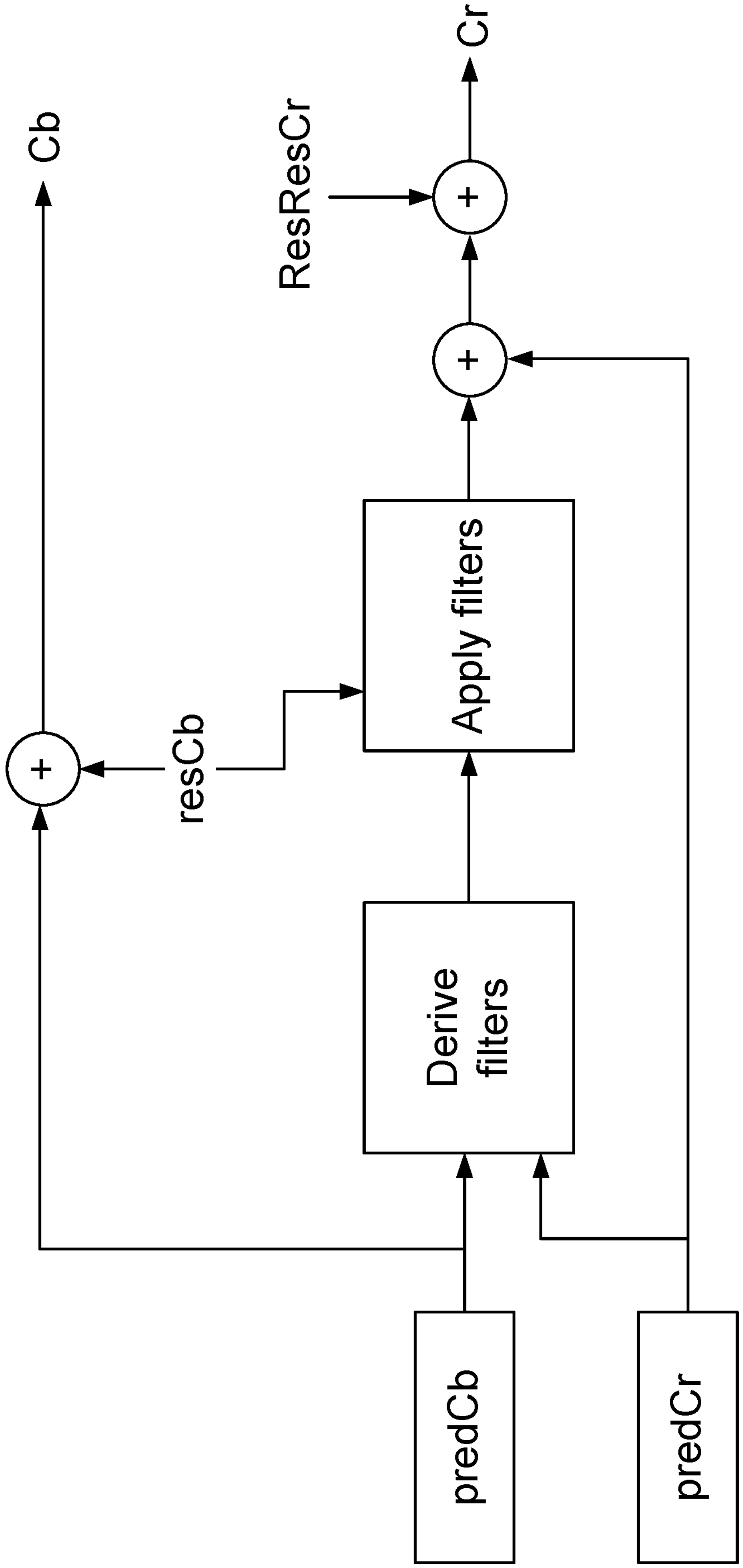

Cb
Cr
ResResCr
+
+
+
resCb
Apply filters
Derive filters
predCb
predCr

FIG. 19

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/068438

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N19/103 H04N19/186 H04N19/82
ADD.

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

B. FIELDS SEARCHED

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

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

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

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2021/138476 A1 (BEIJING DAJIA INTERNET INFORMATION TECH CO LTD [CN] ET AL.) 8 July 2021 (2021-07-08)	1,2,6-9, 13-16, 20-23, 27-29
Y	paragraph [0068] - paragraph [0099] figure 7	3-5, 10-12, 17-19, 24-26

Y	US 2016/219283 A1 (CHEN JIANLE [US] ET AL) 28 July 2016 (2016-07-28) paragraph [0080] - paragraph [0081]	3,10,17, 24

-/--

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

[x] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 July 2024	09/08/2024

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

1

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/068438

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ASTOLA (NOKIA) P ET AL: "EE2-3.1: Cross-component residual model (CCRM) for inter prediction", 31. JVET MEETING; 20230711 - 20230719; GENEVA; (THE JOINT VIDEO EXPLORATION TEAM OF ISO/IEC JTC1/SC29/WG11 AND ITU-T SG.16), , no. JVET-AE0059 4 July 2023 (2023-07-04), XP030311213, Retrieved from the Internet: URL:https://jvet-experts.org/doc_end_user/documents/31_Geneva/wg11/JVET-AE0059-v1.zip JVET-AE0059-v1.docx [retrieved on 2023-07-04] figure 1	4,5,11, 12,18, 19,25,26

1

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

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2024/068438

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2021138476 A1	08-07-2021	CN 114846807 A	02-08-2022
		WO 2021138476 A1	08-07-2021
US 2016219283 A1	28-07-2016	BR 112017016111 A2	27-03-2018
		CA 2972501 A1	04-08-2016
		CN 107211124 A	26-09-2017
		EA 201791476 A1	30-11-2017
		EP 3092803 A1	16-11-2016
		ES 2715555 T3	04-06-2019
		HU E041885 T2	28-06-2019
		JP 6543716 B2	10-07-2019
		JP 2018509035 A	29-03-2018
		KR 20170107448 A	25-09-2017
		TN 2017000284 A1	16-01-2019
		TW 201640895 A	16-11-2016
		US 2016219283 A1	28-07-2016
		WO 2016123219 A1	04-08-2016