



(51) International Patent Classification:

H04N 19/82 (2014.01) H04N 19/124 (2014.01)
H04N 19/86 (2014.01) H04N 19/119 (2014.01)
H04N 19/61 (2014.01) H04N 19/13 (2014.01)

(21) International Application Number:

PCT/KR2024/007733

(22) International Filing Date:

05 June 2024 (05.06.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

202341038884 06 June 2023 (06.06.2023) IN
202341038884 08 May 2024 (08.05.2024) IN

(71) Applicant: SAMSUNG ELECTRONICS CO., LTD.
[KR/KR]; 129, Samsung-ro, Yeongtong-gu, Suwon-si,
Gyeonggi-do 16677 (KR).

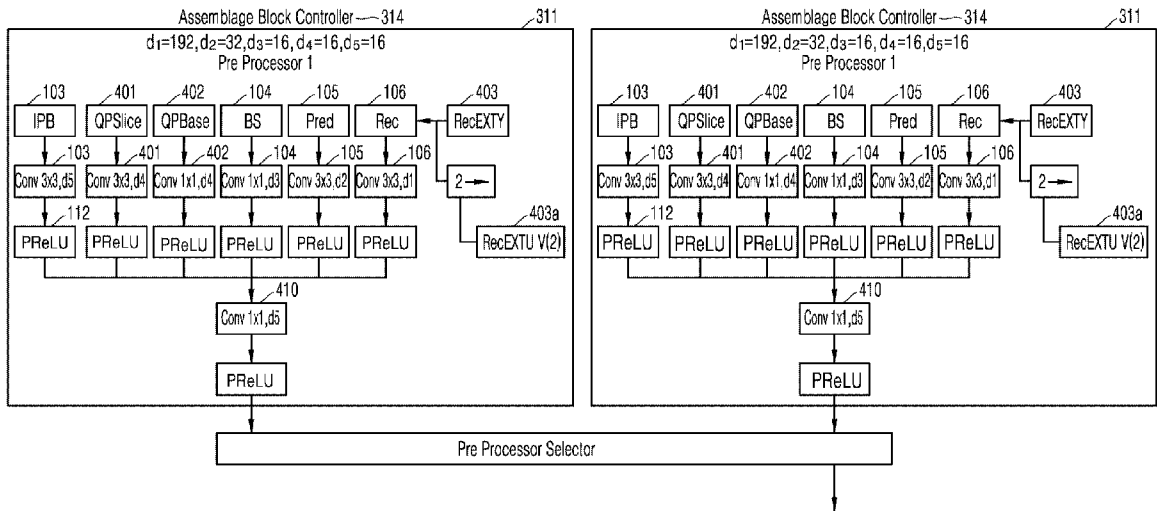
(72) Inventors: SINGH, Anubhav; 2041, Prestige kew
gardens, Yemalur, Bangalore, Karnataka 560037 (IN).
AGRAWAL, Aviral; O-003, Aditya Mega City, Vaib-
hav Khand, Indirapuram, Ghaziabad, Uttar Pradesh 201014

(IN). GADDE, Raj Narayana; #8, Saugandhikam, 7th
B2 Cross, Abhay reddy layout, Kaggadasapura, Banga-
lore, Karnataka 560093 (IN). PIAO, Yinji; 129, Sam-
sung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 16677
(KR). PARK, Minwoo; 129, Samsung-ro, Yeongtong-gu,
Suwon-si, Gyeonggi-do 16677 (KR). CHOI, Kwangpyo;
129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do
16677 (KR).

(74) Agent: Y.P.LEE, MOCK & PARTNERS; 12F Daelim
Acrotel, 13 Eonju-ro 30-gil, Gangnam-gu, Seoul 06292
(KR).

(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: METHOD AND ELECTRONIC DEVICE FOR COMPRESSING A VIDEO USING AI-BASED IN LOOP FILTER



(57) Abstract: Embodiments herein provide a method and an electronic device for compressing a video for AI-based in loop filter (AILF). The method includes obtaining a video comprising a plurality of frames, wherein each frame of the pluralities of frames comprises a plurality of channels. Further, the method includes extracting at least one feature from each frame of the plurality of frames of the video. Further, the method includes selecting, at least one pre-processor from the multi pre-processors (311) for the AILF (318) based on the at least one feature from each frame of the plurality of frames. Further, the method includes generating an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor.



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))*

Description

Title of Invention: METHOD AND ELECTRONIC DEVICE FOR COMPRESSING A VIDEO USING AI-BASED IN LOOP FILTER

Technical Field

- [1] The application is based on and derives the benefit of Indian Provisional Application 202341038884 filed on 6th June, 2023, the contents of which are incorporated herein by reference. The present disclosure relates to the field of video codec. More particularly proposed disclosure is related to a method and an electronic device for compressing a video AI-based in loop filter.

Background Art

- [2] In the realm of digital multimedia, video compression technology plays a vital role in efficient storage and transmission. Video data is commonly managed and conveyed as a series of bit streams. To achieve substantial compression efficiency, conventional video compression encoders and decoders, also referred to as "CODECs," generate a predictive reference picture for the picture being encoded. This encoding process entails representing the difference between the current picture and the predicted reference. The greater the correlation between the prediction and the current picture, the fewer bits are necessary for compressing an image, which ultimately improves the overall efficiency of the compression process. The creation of the most precise reference picture prediction is highly coveted.
- [3] Despite significant advancements in video compression technology, inherent limitations still persist. The AI based In Loop Filter, which employs a fixed ratio, struggles to keep pace with the ever-growing content, impeding the attainment of higher compression ratios and superior video quality. However, the Joint Video Experts Team (JVET) and Moving Picture Experts Group (MPEG), the standard bodies for video compression, are actively exploring the potential of artificial intelligence to enhance compression efficiency. This involves integrating AI into various stages of the video compression pipeline, such as prediction, transformation, quantization, and entropy processes.
- [4] The integration of Artificial Intelligence (AI) within the video compression process is currently under active consideration by a video compression standard body. This initiative demonstrates the vast potential of AI in transforming video compression processes. The standard body is striving to enhance the efficiency and performance of AI tools, with the ultimate goal of improving compression ratios, reducing data transmission bandwidth, and optimizing video quality. The usage of the AI in video compression is driven by the prospect of achieving superior compression technologies

in the future. The AI tools employed in this process include neural network models, which are trained based on video data to make decisions on video compression.

Disclosure of Invention

Solution to Problem

- [5] In an embodiment of the disclosure, a method for managing multi pre-processors for an AI-based In Loop Filter(AILF) in a video codec, may include obtaining, by an electronic device, a video comprising a plurality of frames, wherein each frame of the pluralities of frames comprises a plurality of channels before applying the AILF, and wherein each of the channels comprises image information. The method may include extracting, by the electronic device, at least one feature from each frame of the plurality of frames. The method may include selecting, by the electronic device, at least one pre-processor from the multi pre-processors for applying the AILF based on the at least one extracted feature, wherein each pre-processor of the multi pre-processors comprises a set of ratios each of which corresponds to each channel of the plurality of channels of each frame of the plurality of frames. The method may include generating, by the electronic device, an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor for the AILF. The method may include transmitting, by the electronic device, the encoded video and an index corresponding to the at least one selected pre-processor to a decoder.
- [6] In an embodiment of the disclosure, a method for managing multi pre-processors for an AI-based In Loop Filter(AILF) in a video codec, may include obtaining, by an electronic device, encoded video and index corresponding to at least one selected pre-processor. The method may include determining, by the electronic device, the at least one selected pre-processor for decoding the encoded video based on the index corresponding to the at least one selected pre-processor. The method may include decoding, by the electronic device, the encoded video using a set of ratios of the at least one selected pre-processor.
- [7] In an embodiment of the disclosure, an electronic device for managing multi pre-processors for an AI-based In Loop Filter (AILF) in a video codec, comprising: a memory comprising a video to be encoded and storing one or more instructions; and at least one processor including the multi pre-processors, configured to execute the one or more instructions. The at least one processor may be configured to execute the one or more instructions to obtain a video comprising a plurality of frames, wherein each frame of the pluralities of frames comprises a plurality of channels before applying the AILF, and wherein each of the channels comprises image information. The at least one processor may be configured to extract at least one feature from each frame of

the plurality of frames. The at least one processor may be configured to select at least one pre-processor from the multi pre-processors for applying the AILF based on the at least one extracted feature, wherein each pre-processor of the multi pre-processors comprises a set of ratios each of which corresponds to each channel of the plurality of frames of each frame of the plurality of frames. The at least one processor may be configured to generate an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor for the AILF. The at least one processor may be configured to transmit the encoded video and an index corresponding to the at least one selected pre-processor to a decoder.

Brief Description of Drawings

- [8] These and other features, aspects, and advantages of the proposed disclosure are illustrated in the accompanying drawings, throughout which like reference letters indicate corresponding parts in the various figures. The embodiments herein will be better understood from the following description with reference to the drawings, in which:
- [9] Figure 1 is a block diagram that illustrates a video compression architecture, according to prior art;
- [10] Figure 2 is illustrates extracted feature maps from a video, according to prior art;
- [11] Figure 3a is a block diagram that illustrates electronic device with an assemblage block controller for the video compression, according to the embodiment;
- [12] Figure 3b is a block diagram that illustrates the assemblage block controller, according to the embodiment;
- [13] Figure 4a is a block diagram that illustrates the multiple pre-processors selector with backbone network for the video compression, according to the embodiment;
- [14] Figure 4b is a block diagram that illustrates the multiple pre-processors selector with backbone network for the video compression, according to the embodiment;
- [15] Figure 5 is a block diagram that illustrates a Versatile Video Codec (VVC) decoder with AILF, according to the embodiment;
- [16] Figure 6 is a block diagram that illustrates encoder for the VVC, according to the embodiment;
- [17] Figure 7 is a block diagram that illustrates video compressor architecture including unique pre-processors and the backbone network, according to the embodiment;
- [18] Figure 8 is a block diagram that illustrates an encoder and a decoder with the pre-processor for video codec, according to the embodiment;
- [19] Figure 9a is a block diagram that illustrates step by step details of the encoder for any codec, according to the embodiment;

- [20] Figure 9b is a block diagram that illustrates step by step details of the encoder for any codec, according to the embodiment;
- [21] Figure 9c is a block diagram that illustrates step by step details of the encoder for any codec, according to the embodiment;
- [22] Figure 10 is a block diagram that illustrates step by step details of the decoder for any codec, according to the embodiment;
- [23] Figure 11a illustrates a image frame with a CTU and slice numbers, according to the embodiment;
- [24] Figure 11b is a block diagram that illustrates the extracted feature maps with indices, according to the embodiment;
- [25] Figure 12 is a flow diagram illustrating a method for compressing the video using the AI model, according to the embodiment;
- [26] It may be noted that to the extent possible, like reference numerals have been used to represent like elements in the drawing. Further, those of ordinary skill in the art will appreciate that elements in the drawing are illustrated for simplicity and may not have been necessarily drawn to scale. For example, the dimension of some of the elements in the drawing may be exaggerated relative to other elements to help to improve the understanding of aspects of the invention. Furthermore, the elements may have been represented in the drawing by conventional symbols, and the drawings may show only those specific details that are pertinent to the understanding the embodiments of the invention so as not to obscure the drawing with details that will be readily apparent to those of ordinary skill in the art having benefit of the description herein.

Mode for the Invention

- [27] The principal object of the embodiments herein is to a method and an electronic device for compressing a video for AI based In Loop Filter (AILF).
- [28] Another object of the embodiments herein is to create a hybrid codec model through the utilization of the AILF. This innovative filter can be positioned at any point within the conventional in-loop filter, and its placement is crucial in ensuring the production of high-quality compressed video.
- [29] Another object of the embodiments herein is to utilize multiple pre-processors with varying channel ratios, which are allocated to each channel of the pre-processor. The optimal pre-processor is then chosen from this selection based on the channel ratios for both compressing and decompressing the video.
- [30] Another object of the embodiments herein is to provide optimal set of ratios of the at least one selected pre-processor is embedded in at least one of a sequence header, a picture header, a slice header, and a Coding Tree Unit (CTU) of each frame of the plurality of frames.

- [31] The embodiments herein and the various features and advantageous details thereof are explained more fully with reference to the non-limiting embodiments that are illustrated in the accompanying drawings and detailed in the following description. Descriptions of well-known components and processing techniques are omitted so as to not unnecessarily obscure the embodiments herein. Also, the various embodiments described herein are not necessarily mutually exclusive, as some embodiments can be combined with one or more other embodiments to form new embodiments. The term "or" as used herein, refers to a non-exclusive or, unless otherwise indicated. The examples used herein are intended merely to facilitate an understanding of ways in which the embodiments herein can be practiced and to further enable those skilled in the art to practice the embodiments herein. Accordingly, the examples are not be construed as limiting the scope of the embodiments herein.
- [32] As is traditional in the field, embodiments are described and illustrated in terms of blocks that carry out a described function or functions. These blocks, which referred to herein as managers, units, modules, hardware components or the like, are physically implemented by analog and/or digital circuits are logic gates, integrated circuits, microprocessors, microcontrollers, memory circuits, passive electronic components, active electronic components, optical components, hardwired circuits and the like, and optionally be driven by firmware and software. The circuits, for example, be embodied in one or more semiconductor chips, or on substrate supports are printed circuit boards and the like. The circuits constituting a block be implemented by dedicated hardware, or by a processor (e.g., one or more programmed microprocessors and associated circuitry), or by a combination of dedicated hardware to perform some functions of the block and a processor to perform other functions of the block. Each block of the embodiments be physically separated into two or more interacting and discrete blocks without departing from the scope of the proposed method. Likewise, the blocks of the embodiments be physically combined into more complex blocks without departing from the scope of the proposed method.
- [33] The accompanying drawings are used to help easily understand various technical features and it is understood that the embodiments presented herein are not limited by the accompanying drawings. As such, the proposed method is construed to extend to any alterations, equivalents and substitutes in addition to those which are particularly set out in the accompanying drawings. Although the terms first, second, etc. used herein to describe various elements, these elements are not be limited by these terms. These terms are generally used to distinguish one element from another.
- [34] Accordingly, the embodiment herein a method for compressing a video using multi pre-processors for an AILF in a video codec. The method includes receiving a video comprising a plurality of frames, wherein each frame of the pluralities of frames

comprises a plurality of channels, before applying to an Artificial Intelligence (AI) model. Each of the channels having image information to be encoded by an encoder of the electronic device. Further, the method includes extracting at least one feature from each frame of the plurality of frames of the video. Further, the method includes selecting at least one pre-processor from the multi pre-processors for applying the AILF to enhance the encoding the image information based on the at least one feature from each frame of the plurality of frames, wherein each pre-processor of the multi pre-processors comprises a set of ratios each of which corresponds to each channel of the plurality of frames of each frame of the plurality of frames, and wherein the at least one selected pre-processor comprises an optimal set of ratios for encoding the image information from each channel. The method includes generating an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor for the AILF. The method includes transmitting the encoded video and an index corresponding to the at least one selected pre-processor to a decoder of the electronic device.

- [35] Accordingly, the embodiment herein an electronic device for managing multi pre-processors for the AILF in a video codec. The electronic device includes the multi pre-processors connected to an encoder and a decoder, a memory comprising a video to be encoded and an assemblage block controller (314)assemblage block controller, coupled to the memory and the multi pre-processors. Further, the assemblage block controller receive the video comprising a plurality of frames. Each of the frames includes a plurality of channels before applying the AILF. Each of the channels having image information to be encoded by an encoder of the electronic device. Further, the assemblage block controller extract at least one feature from each frame of the plurality of frames of the video. Further, the assemblage block controller selects at least one pre-processor from the multi pre-processors for encoding the image information based on the at least one feature from each frame of the plurality of frames. Each of the multi pre-processors includes a set of ratios each of which corresponds to each channel of the plurality of frames of each frame of the plurality of frames. Further the selected pre-processor comprises an optimal set of ratios for encoding the image information from each channel. The assemblage block controller generates an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor. The assemblage block controller transmits the encoded video and an index corresponding to the at least one selected pre-processor to a decoder of the electronic device.

- [36] In the existing system, the optimization of video compression while enhancing its quality is achieved through a sophisticated video compression architecture. This

architecture comprises a pre-processor equipped with a backbone network, tailored for an AILF or AI model. The AILF employs a Convolutional Neural Network (CNN) to execute the video compression process. By learning from data, the CNN proves to be highly effective in identifying patterns in images, recognizing object classes, and categorizing images. Moreover, CNNs are also leveraged for classifying audio, time-series, and signal data.

- [37] The CNN extracts a feature from a video frame using a fixed feature extraction layer. Input channels, including luma reconstruction, prediction buffer, boundary strength, quantization parameter (QP) base, QP slice, partition average, reconstructed prediction, and block coding types, are utilized in the process. Each selected input channel is assigned a fixed channel ratio, with luma reconstruction, reconstruction, prediction buffer, QP, and boundary strength assigned ratios of 192, 24, 12, 48, and 48, respectively. The extracted feature is then passed to the backbone network, which is pre-processed with fixed channel ratios. However, this pre-processing method ignores the importance of specific feature characteristics and results in suboptimal use of input channels, leading to a higher network complexity due to redundant features.
- [38] The proposed solution offers a novel approach to compressing video, distinct from conventional systems and methods. By utilizing multiple detachable pre-processors with varying input channel ratios, the system can effectively learn and extract features from the video's image. The backbone network is trained with these pre-processors, ultimately selecting one for encoding based on either a heuristic method or the pre-processor that yields the best results in terms of Bitrate and Quality metrics. Quality is measured using Peak Signal-to-Noise Ratio (PSNR) and other ratios, with higher PSNR indicating better compressed or reconstructed image quality. The selected pre-processor is then signalled to a decoder with an index, allowing for decoding of the encoded images into a bit stream of high quality.
- [39] Presented in FIG 1 is a block diagram showcasing a video compression architecture as previously disclosed. Within this architecture, the pre-processor (100a) and the backbone network (100b) play integral roles. The pre-processor (100a) comprises of a convolution layers set and a series of Parametric Rectified Linear Unit (PRELU) layers (112). Meanwhile, the backbone network (100b) incorporates both a fuse technique (114) and a transition technique (117). The combination of these two components, the pre-processor (100a) and the backbone network (100b), effectively facilitates the AILF.
- [40] The input channels, carrying information from the frames of the video, are received by the pre-processor (100a) and undergo processing before being sent to the backbone network (100b). The pre-processor determines and assigns ratios to each input channel, based on their suitability for enhancing video quality. These ratios, including but not

limited to d_1 , d_2 , d_3 , d_4 , d_5 , and d_6 , are fixed and non-variable. The selected input frame of the video and these ratios are then processed through convolution layers with various filtering options. Finally, the processed output is transmitted through the backbone network (100b) to compress the video without any compromise to its quality.

[41] In the Figure 1, an unsqueeze expand (102) is depicted receiving a distinct buffer (DB) (101) as input, which is assigned a ratio of d_1 by the pre-processor. The unsqueeze expand (102) then expands the channel dimensions of DB (101). Subsequently, a first Conv 3x3 (1 d_3) layer (107) receives input from the unsqueeze expand (102) and applies a 3x3 filter to extract features. Similarly, other inputs, such as an Inter Prediction Block (IPB) (103), a base signal (BS) (104), a prediction (pred) (105), and a reconstruction (rec) (106), are received with ratios of d_3 , d_3 , d_2 , and d_1 , respectively, as defined by the pre-processor. The inputs with assigned ratios are processed through convolution layers, such as the second Conv 3x3 (1 d_3) layer (107), third Conv 3x3 (1 d_3) layer (107), Conv 3x3 (1 d_2) layer (110), and Conv 3x3 (1 d_1) layer (111). These convolution layers utilize 3x3 filtering to provide blur-free sharpening or feature extraction from the inputs. The output of these convolutional layers (107, 110, and 111) is then passed through their respective PReLU layers (112), which apply a parametric RELU activation function to capture nuanced features and patterns for enhancing video compression ability. Finally, the output of the PReLU layer (112) is passed through the backbone network (100b).

[42] The concatenated output from the concat layer (113) is received by the backbone network (100b) and undergoes a series of techniques to enhance its efficiency. The fuse technique (114) comprises a Conv 1x1 ($d_1+d_2+3*d_3$ d_4) layer (115) and a PReLU layer (112), which work together to reduce the dimensionality of the data and apply an activation function for non-linearity. The Conv 1x1 ($d_1+d_2+3*d_3$ d_4) layer (115) applies a filter to the concatenated input channels ratios, compressing the data for optimal processing. The output from this layer then passes through the PReLU layer (112) to determine non-linearity. Finally, the output from the fuse technique (114) is passed to the Conv 3x3 2 (d_4 96) layer (118), which applies a 3x3 convolutional filter to the previous layer's output, down-samples the data by a factor of 2 in both width and height dimensions, and further reduces the data's dimensionality.

[43] Figure 2 is illustrates extracted feature maps from a video, according to prior art disclosed herein. The video yields feature maps for an original patch (201), a partition patch (202), a reconstructed patch (203), a prediction patch (204), and the boundary strength of a patch (205). These extracted features serve as inputs to the pre-processor, which processes them to derive the optimal output based on the most fitting feature patch.

- [44] The initial patch (201) comprises unrefined and unprocessed video data extracted from the source. This patch (201) serves as the fundamental reference point for the subsequent stages of processing.
- [45] The partition patch (202) extracted from the video is . The partition map the encoder has decided based on Rate Distortion Optimization (RDO).
- [46] The reconstructed patch (203) is the outcome of a process that involves utilizing information from various patches, including but not limited to the partition patch (202) and prediction (204), to generate a refined version of the original patch (201) from the video bit stream.
- [47] The prediction patch (204) comprises a forecast of the initial patch (201) derived from pertinent information found in the original patch (201). This process may entail utilizing data from adjoining patches in frames of the video. The precision of the prediction amplifies the excellence of the reconstituted patch (203).
- [48] The patch (205) boundary strength denotes a measure of confidence in the precision of patch boundaries. This parameter is utilized in object detection, segmentation, and motion estimation. A high value of boundary strength signifies a robust separation between distinct regions in the reconstructed patch (203).
- [49] Figure 3a is a block diagram that illustrates electronic device with an assemblage block controller for video compression, according to the embodiment disclosed herein. The electronic apparatus (310) comprises a multitude of pre-processors (311), an I/O interface (312), a memory (313), and an assemblage block controller (314).
- [50] The video codec device (310) incorporates an assemblage block controller (314) to compress video. In addition, the electronic device (310) features multi pre-processors (311) that communicate with the memory (313), the I/O interface (312), and the assemblage block controller (314). These multi pre-processors (311) execute instructions stored in the memory (313) to perform a range of processes. They may include one or several processors, including general-purpose processors such as a central processing unit (CPU) or an application processor (AP), graphics-only processing units like a graphics processing unit (GPU) or a visual processing unit (VPU), and/or artificial intelligence (AI) dedicated processors such as a neural processing unit (NPU).
- [51] Further, the memory (313) of the electronic device (310) includes storage locations to be addressable through the multi pre-processors (311). The memory (313) is not limited to a volatile memory and/or a non-volatile memory. Further, the memory (313) can include one or more computer-readable storage media. The memory (313) can include non-volatile storage elements. For example, non-volatile storage elements can include magnetic hard discs, optical discs, floppy discs, flash memories, or forms of electrically programmable memories (EPROM) or electrically erasable

and programmable (EEPROM) memories. The memory (313) is capable of storing quantized coefficients for each frame of a video, along with pre-processor numbers assigned to each pixel value. This includes various values for each frame such as sequence header, picture header, slice header, slice numbers, and index numbers of the pre-processor. Additionally, the memory (313) stores ratios assigned to each input channel of the multi pre-processors (311), which are related to luma reconstruction, prediction buffer, boundary strength, QP base, QP slice, and block coding type. The memory (313) may store one or more instructions for operations performed by at least one processor.

- [52] The I/O interface (312) transmits the information between the memory (313) and external peripheral devices. The peripheral devices are the input-output devices associated with the electronic device (310). The I/O interface (312) receives the information from the memory (313) and the multi pre-processors (311). The I/O interface (312) communicates with the assemblage block controller (314) to fetch the data and process for the selection of the multi pre-processors (311).
- [53] Although not explicitly shown in the Figure 3a, the electronic device may include at least one processor. The at least one processor may execute one or more instructions stored in memory, and the at least one processor may include at least one of controller, an assemblage block controller, and multi pre-processors.
- [54] The assemblage block controller (314) is a cutting-edge hardware that incorporates both analog and digital circuits, including logic gates, integrated circuits, microprocessors, microcontrollers, memory circuits, passive and active electronic components, as well as optical components. It interfaces with the I/O interface (312) and memory (313) to receive ratios from the multi pre-processors (311) and stored data from the memory (313) that is configured to accommodate a video comprising multiple frames, each of which consists of various channels containing image information to be encoded by an encoder of the electronic device (310). Moreover, the assemblage block controller (314) extracts at least one feature from each frame of the video and selects at least one pre-processor from a group of multi pre-processors (311) based on the extracted features. Each pre-processor within the multi pre-processors (311) has a set of ratios that correspond to individual channels within every frame of the video. The chosen pre-processor is distinguished by an optimal set of ratios designed for the encoding of image information from each channel. The assemblage block controller (314) then generates an encoded video by encoding the image information from each channel of the frames of the video using the selected pre-processor. Finally, it transmits the encoded video and an index corresponding to the selected pre-processor to a decoder (317) of the electronic device (310).

- [55] The controller (314) encode the image information from each channel of every frame in the plurality of frames across multiple channels. The incoming video is then segmented into distinct blocks for intra prediction technique (610) and inter prediction technique (512). Transformation and quantization are applied to the video blocks based on the intra and inter prediction techniques (610 512) to produce quantized coefficients for the segmented video blocks that have undergone both intra and inter prediction (610 512). The assemblage block controller (314) then reconstructs the video frames using the generated quantized coefficients. The reconstructed video is passed through the AILF (318) model for artifact removal, resulting in an artifact-free video. The assembly block controller (314) of the assemblage block employs a selected pre-processor to generate a feature map for back bone of the AILF. Finally, the artifact-free reconstructed video is passed through entropy coding for encoding.
- [56] Moreover at the decoder, the assembly block controller (314) receives the encoded video and its associated index, pertaining to the chosen pre-processor. Subsequently, the controller uses this index to identify the selected pre-processor for decoding the encoded video. Further, the assembly block controller (314) determines the most suitable set of ratios for the selected pre-processor, based on the index, and utilizes this optimal set of ratios for decoding the encoded video. Additionally, the assembly block controller (314) incorporates the optimal set of ratios of the chosen pre-processor into the encoded image information, to be utilized by the decoder (317) for decoding the encoded video. Finally, the pre-processed image information from each channel and frame is inputted into an in-loop filter to generate the encoded video.
- [57] Figure 3b is a block diagram that illustrates the assemblage block controller, according to the embodiment disclosed herein. The assemblage block controller (314) includes a pre-processor selector (315), an encoder (316), a decoder (317), and the AILF (318). The pre-processor selector (315) chooses the optimal pre-processor from the multi-preprocessor (311) by considering both the quality of the video output and the pre-processor's index number.
- [58] Meanwhile, for convenience of explanation, the encoder and decoder are expressed as being stored in an electronic device, but this is not limited to the disclosed example. For example, the operation of transmitting from an encoder to a decoder may include an operation of transmitting from an encoder of the electronic device to a decoder of another electronic device. It may also include operations transmitted from the encoder to the decoder within one electronic device.
- [59] Each pre-processor within the multi pre-processors (311) is equipped with several input channels that carry video frame information. Prior to being sent to the backbone network (419), this information is processed by the pre-processor's input channels,

which are assigned various ratios, including but not limited to d1, d2, d3, d4, d5, and d6. These ratios may be adjusted as needed.

[60] The input channels themselves include an IPB (103), a QPSlice (401), a QPBase (402), a BS (104), a pred (105), and a rec (106), each of which is processed by convolution layers with different filtering options, such as Conv 3x3, Conv 3x1, Conv 1x3, and Conv 1x1. These convolution layers serve to eliminate blurring, sharpen details, and extract features from each input channel.

[61] The processed output is then passed through the PRELU (112), which performs a threshold operation based on the provided ratios and applies scaling coefficients to the input when it falls below the threshold or is zero. The PRELU layer (112) captures nuanced features and patterns from the input channels, potentially enhancing video compression quality.

[62] An index is generated for each pre-processor within the multi pre-processors (311) to determine which pre-processor provides the best output based on the assigned ratios and is suitable for the video bitstream. The encoder (316) selects the processed information of the video frame based on the selected pre-processor and encodes it before sending it to the bitstream.

[63] The pre-processor selector (315) expertly chooses the most optimal pre-processor from the multi pre-processors (311) and transmits the index of this chosen pre-processor to the decoder (317). The output of this selected pre-processor is then passed through Conv 2 3x3 C (418), which expertly utilizes a 3x3 convolutional filter for down-sampling operations and produces an output with a specific number of channels C. This output then proceeds to the PRELU layer (112), which applies a parametric RELU activation function to determine non-linearity and captures features and patterns for potentially enhancing video compression ability. Finally, the output passes through the backbone network (419).

[64] The backbone network (419) comprises convolution layers with various filtering functions and channels, such as Conv 1x1 CxC1 (411), Conv 3x1 CxC21 (412), and Conv 1x3 C21xC22 (413). These convolution layers (411, 412, and 413) process the provided data and pass it through the PRELU layer (112). The data processed by the PRELU layer (112) is then passed through Conv 1x1 (C1+C22) x C (414), Conv 1x3 CxC31 (415), and Conv 3x1 C31xC (416) for further processing. The [c h w] (417) is then used to determine the number of channels or depth, height, and width of each frame of the video.

[65] Once the backbone network (419) processes and extracts the required information, it passes through another convolution network, Conv 3x3 (420), for feature extraction, followed by the PRELU layer (112) for introducing adaptability. Subsequently, a second convolutional layer with Conv 3x3 6 layer (421) is employed for continuous

feature extraction. The processed output from the Conv 3x3 6 (421) is then passed through the crop technique (423), which involves extracting a specific region from the input. Pixel shuffling (422) is then employed, potentially for tasks like super-resolution or color space transformation.

- [66] The Rec UV operation (424) signifies the reconstruction of color information, the U and V chroma channels, followed by another crop operation (423). The final step involves the reconstruction of the luminance channel (Rec Y) (425).
- [67] The decoder (317) decodes the provided encoded frames of the video by using the selected pre-processor. The decoder (317) process the encoded frames of the video through an inverse quantization technique (502), an inverse transform technique (503), and the inverse prediction. The decoder (317) uses the AILF (318) generates only one artefact removed reconstructed frame to bit stream the video.
- [68] The AILF (318) is used in the encoder (316) and decoder (317) to process the input channels for encoding, selecting the best pre-processor. Further, decoding the encoded video and decoding the encoded data to bit stream the video on the display. The encoded video is passed through the AILF (318) reduces artefacts and improve the visual quality of the video. The AILF (318) operates by smoothing out discontinuities or artefacts that may arise during the compression and decoding process. The AILF processed decoded quality video is stored in the memory by the decoded picture buffer technique (608).
- [69] Figure 4a is a block diagram that illustrates the multiple pre-processors selector with backbone network for the video compression, according to the embodiment disclosed herein. Figure 4b is also a block diagram that illustrates the multiple pre-processors selector with backbone network for the video compression, according to the embodiment disclosed herein.
- [70] The Figure 4a shows the assemblage block controller (314) includes the multi pre-processors (311) and the pre processor selector. The multi pre-processors (311) includes n number of pre-processor with the input channels having a different ratios.
- [71] Each of the pre-processors within the multi pre-processors (311) is equipped with distinct input channels that capture information from the video frames. Prior to transmission to the backbone network (419), the pre-processor undertakes information processing. Each input channel is associated with a range of ratios, including but not limited to d1, d2, d3, d4, d5, and d6, among others. These ratios are variable and can be adjusted to meet specific requirements, contributing to the system's adaptability. Input channels, such as the IPB (103), the QPSlice (401), the QPBase (402), the BS (104), the pred (105), and the rec (106), undergo processing via convolution layers, utilizing various ratios. In an embodiment, a value of d1 may be 192, a value of d2

may be 32, a value of d3 may be 16, a value of d4 may be 16, a value of d5 may be 16, and a value of d6 may be 48, but, it is not limited to the disclosed examples.

- [72] The input channel's IPB (103) represents a predicted block based on neighboring pixels within the video frame, while the QPSlice (401) performs the quantization Parameter (QP) for each frame of the video. The QP measures the quality of the quantization process, with higher QP resulting in better quality but also a higher bitrate. The QPbase (402) determines the base QP for the video, which serves as a starting point for determining the QP for each frame. The BS (104) represents the video signal in the frame, while pred (105) reduces the bitrate of each frame by using redundancy between frames. The rec (106) represents the reconstructed pixels after quantization and entropy coding.
- [73] To improve the quality of the reconstructed video, the RecEXTY (403) and RecEXTUV (403a) blocks perform reconstruction extension by down-sampling the data and reducing dimensionality. The input channels of the multi pre-processors (311) are passed through convolution layers (404-409) and respective PRELU layers (112) to process the input channels with defined ratios. The combined PRELU layer output is then passed through another convolutional layer (410) and the PRELU layer to further process features from the input channels, contributing to overall transformations and representations in the AILF (318).
- [74] Each pre-processor generates an output frame of the video, with an index generated for each pre-processor. The output of the multi pre-processors (311) is input to the pre-processor selector (315) to select the best pre-processor based on the best output and the index of the pre-processor. The encoder (316) then selects the best pre-processor to encode and process the input channels, informing the decoder (317) of the selected pre-processor.
- [75] The Figure 4a shows the assemblage block controller (314) includes some convolution layers and backbone network. The chosen pre-processor output undergoes a down-sampling operation through Conv 2 3x3 C (418), which utilizes a 3x3 convolutional filter to generate an output with a specific number of channels C. This output then proceeds through the PRELU layer (112), which employs a parametric RELU activation function to identify non-linearity and detect features and patterns that could potentially improve video compression capabilities. Finally, the output passes through the backbone network (419).
- [76] The backbone network (419) comprises convolution layers with diverse filtering functions and channels, including but not limited to Conv 1x1 CxC1 (411), Conv 3x1 CxC21 (412), and Conv 1x3 C21xC22 (413). The provided data is processed by these convolution layers (411, 412, and 413) and then passed through the PRELU layer (112). The processed data from the PRELU layer (112) is further passed through

Conv 1x1 (C1+C22) x C (414), Conv 1x3 CxC31 (415), and Conv 3x1 C31xC (416) to obtain the number of channels or depth, height, and width of each frame of the video [c h w] (417). In an embodiment, a value of C may be 64, a value of C1 may be 160, a value of C21 may be 32, a value of C22 may be 32, and a value of C31 may be 64, but, it is not limited to the disclosed examples.

[77] Once the backbone network (419) extracts the required information, it passes through another convolution network, Conv 3x3 (420), for feature extraction, followed by the PRELU layer (112) activation function, which introduces adaptability. Subsequently, a second convolutional layer with Conv 3x3 6 layer (421) is employed for continuous feature extraction. The processed output from Conv 3x3 6 (421) is then passed through the crop technique (423), which involves extracting a specific region from the input. Pixel shuffling (422) is subsequently employed, potentially for tasks such as super-resolution or colour space transformation.

[78] The Rec UV technique (424) signifies the reconstruction of colour information, the U and V chroma channels, followed by another crop technique (423). The final step involves the reconstruction of the luminance channel (Rec Y) (425). Once the complete process is done, the best artifact-removed reconstructed frame of the video is generated. A frame using or merging The Rec UV and the Rec Y may be the frame used in next process or final frame.

[79] Figure 5 is a block diagram that illustrates a Versatile Video Codec (VVC) decoder with AILF, according to the embodiment disclosed herein. The AILF (318) pertains to a sophisticated video decoding system. The VVC, a widely accepted video compression standard, relies on the decoder (317) to reconstruct the video frames from the encoded data. With the integration of AILF, cutting-edge AI technology is utilized to enhance the in-loop filtering process, which aims to intelligently improve the video quality by minimizing artifacts and augmenting visual fidelity.

[80] The compressed bitstream is first subjected to processing by a Context-Adaptive Binary Arithmetic Coding (CABAC) (501), which effectively decodes the entropy-encoded data. Next, the decoded information is subjected to the inverse quantization technique (502), which restores transformed coefficients to their original precision. The inverse transform technique (503) is then employed to further reconstruct the spatial representation. Subsequently, a Luma Mapping and Chroma Scaling (chroma residue scaling) (LMCS) (504) operation is performed. Specifically, the LMCS (chroma residue scaling) (504) operation refines the colour components, effectively addressing chroma-related artifacts and contributing to the enhancement of visual quality in the reconstructed video frames.

[81] In video coding, a Closed Intra Prediction (CIIP) (510) technique is utilized for intra-frame compression, which involves making predictions within the same frame.

To further enhance the process, the CIIP (510) is combined with the LMCS (chroma residue scaling) (504) technique. This combination is activated when certain conditions are met or when the circuit is closed. The output of the LMCS (chroma residue scaling) (504) and the CIIP (510) combined is then fed into the intra prediction technique (610), which estimates pixel values based on neighbouring pixels within the frame of the video. The resulting output of the intra prediction technique (610) is then fed back into the CIIP (510), which has the potential to greatly improve the accuracy and quality of the predicted intra-frame content.

[82] The combined output of the LMCS (chroma residue scaling) (504) and CIIP (510) undergoes a series of processing stages to further refine video decoding. Initially, an LMCS operation is applied for inverse luma mapping (506), aimed at adjusting the luma component in reverse. The video signal then passes through a Deblocking Filter (507) to smooth block boundaries and reduce compression artifacts, followed by Sample Adaptive Offset (SAO) (508) and Adaptive Loop Filter (ALF) with Cross-Component Adaptive Loop Filter (CC-ALF) operations (509). These steps collectively contribute to enhancing the visual quality of the reconstructed frames of the video by addressing various distortions and artifacts in the decoded signal. The AILF (318) is placed strategically between the traditional In-Loop Filters, depending on where the maximum quality reconstructed frames of the video can be realized. The AILF (318) refines the reconstructed frames of the video by reducing noise and artifacts to provide a restored image. It adapts and intelligently adjusts to the decoded picture buffer (514) and input to an Inter prediction technique (512) to efficiently predict image blocks by leveraging information from previous frames. A Forward Luma Mapping through LMCS (511) is used for adjusting the luma component for colour and contrast enhancements. Finally, Closed Intra Prediction (CIIP) (510) is employed for intra-frame compression, predicting pixel values within the same frame. These operations collectively contribute to refining the visual quality of the decoded video frames, addressing temporal redundancies and enhancing colour representation during the decoding process.

[83] Figure 6 is a block diagram that illustrates encoder for the VVC, according to the embodiment disclosed herein. The detailed block diagram of the VVC encoder is used to encode the video through encoding process introduced in the VVC encoder block diagram. The video is input to a residual technique (602), an intra prediction technique (610) and a motion estimation technique (609).

[84] The intra prediction technique (610) in video encoding harnesses spatial redundancies within an individual video frame. The intra prediction technique (610) determines the pixel values of each frame of the video (610) by analysing the values of neighbouring pixels in the video. The motion estimation technique (609) identifies

and quantifies the motion of the objects between consecutive frames in a video sequence. By detecting the motion of the objects in each frame of the video represents displacement of pixels between frames. The encoder (316) predicts the location of objects in subsequent frames. The motion estimation technique (609) significantly reduces the amount of data needed to describe moving objects in each frame of the video and enhances compression efficiency of the video with video quality.

[85] The output of the intra prediction technique (610) and the motion estimation technique (609) is input to the residual technique (602), the inverse quantization technique (502) and the inverse transform technique (503). The residual technique (602) provides difference between the input video and the intra predicted and motion estimated for each frame of the video. The output of the residual technique (602) is provided as input to a transform technique (603a) and a quantization technique (603b).

[86] The transform technique (603a) transforms the video. Each pixel values of each frame of the video undergone transform operation to transform spatial information into a frequency domain, highlighting important frequency components and allowing for efficient compression. By concentrating signal energy in a reduced set of coefficients, the transform facilitates subsequent quantization for contributing to the overall compression of frame of the video. The transformed video is quantized by dividing into number of frames of the video and generating quantized coefficients for the divided each frames of the video by the quantized technique (603b). The quantized coefficients of each frames of the video reduces the amount of data that needs to be stored and transmitted. An entropy coding technique (604) receives the quantized technique (603b) output as input and removes any statistical redundancy.

[87] The inverse transform technique (503) converts the quantized coefficients into spatial domain pixel values. The quantized coefficients are multiplied by the quantization step size to recover the original transformed coefficients. The inverse quantization technique (502) reduces degree of distortion or loss of information produced by the quantized coefficients in the quantised video. The residual technique (602) differentiates the received inverse quantized frames of the video and combined intra predicted and motion estimation of each frame of the video. The encoded video is passed through the AILF (318) reduces artefacts and improve the visual quality of the video. The AILF (318) operates by smoothing out discontinuities or artefacts that may arise during the compression and decoding process. The AILF (318) processed decoded quality video is stored in the memory by the decoded picture buffer technique (608).

[88] Figure 7 is a block diagram that illustrates video compressor architecture including pre-processor and the backbone network (419), according to the embodiment disclosed herein. The pre-processors within the multi pre-processors (311) are designed with

diverse input channels that carry information derived from frames of the video. Before sending to the backbone network (419), each pre-processor handles and refines input channels. Each input channel in the pre-processor is provided with various ratios, including but not limited to d1, d2, d3, d4, d5, and d6, among others. These ratios can be adjusted based on requirements. Input channels encompass elements like the IPB (103), the QPSlice (401), the QPBase (402), the BS (104), the pred (105), and the rec (106).

- [89] The input channels of the multi pre-processors (311) undergo convolutional processing through each pre-processor's convolution layers, including but not limited to Conv 3x3 d5 layer (404), Conv 3x3 d4 layer (405), Conv 1x1 d4 layer (406), Conv 1x1 d3 layer (407), Conv 3x3 d2 layer (408), and Conv 3x3 d1 layer (409), to process the input channels with the defined ratios. Following this, the respective PRELU layer (112) processes the input channels, and their combined output is passed through the Conv 1x1 d6 layer (410) and the PRELU layer (112). The output of each pre-processor and their index are utilized to select the best pre-processor from the multi pre-processors (311).
- [90] The selected pre-processor output is channelled through the backbone network (419), which comprises convolution layers with a variety of filtering functions and channels, including but not limited to Conv 1x1 CxC1 (411), Conv 3x1 CxC21 (412), Conv 1x3 and C21xC22 (413). These convolution layers (411, 412, and 413) process the provided data and pass it through the PRELU layer (112). The processed data from the PRELU layer (112) is further passed through Conv 1x1 (C1+C22) x C (414), Conv 1x3 CxC31 (415), and Conv 3x1 C31xC (416). The [c h w] (417) is then used to determine the number of channels, depth, height, and width of each frame of the video.
- [91] Once the backbone network (419) has processed and extracted the required information, it is passed through the Conv 3x3 layer (420) for feature extraction, followed by the PRELU layer (112) for activation function and adaptability. This is subsequently followed by a second convolutional layer with Conv 3x3 6 layer (421) for continuous feature extraction. The output from the Conv 3x3 6 (421) is then passed through the crop technique (423), which involves extracting a specific region from the input. Pixel shuffling (422) is then employed, potentially for tasks like super-resolution or colour space transformation.
- [92] The Rec UV operation (424) signifies the reconstruction of colour information, the U and V chroma channels, followed by another crop operation (423). The final step involves the reconstruction of the luminance channel (Rec Y) (425). Once the complete process is done, the best artifact-removed reconstructed frame of the video is generated.

- [93] Figure 8 is a block diagram that illustrates an encoder and a decoder with the pre-processor for video codec, according to the embodiment disclosed herein. Within the encoder (316), the input channels that transmit information from the frames of the video consist of the original patch (201), partition patch (202), reconstructed patch (203), prediction patch (204), and boundary strength patch (205). The video's extracted features are then fed into the pre-processor for processing and optimal output based on the most appropriate feature patch of the video's frames. These inputs are then directed to the multi pre-processors (311), each with varying ratios assigned to their input channels, such as d1, d2, d3, d4, d5, and d6, among others. The ratios are flexible and can be changed according to the need. The pre-processors employ various techniques, as illustrated in FIG6, to process the input channels and generate output frames of the video and an index for each pre-processor in the multi pre-processors (311).
- [94] The pre-processor selector (315) chooses the best pre-processor from the multi pre-processors (311) based on the video's output quality and pre-processor index number. The selected pre-processor index is then signalled to the decoder (317), which is input to the backbone network (419). This network includes convolution layers with various filtering functions and channels (411, 412, 413, 414, 415, 416, and 417). The processed input frames of the video are encoded by the encoder (316) and input into the bitstream.
- [95] The decoder (317) uses the selected pre-processor from the multi pre-processors (311) to decode the encoded frames of the video. The decoding process begins with the inverse quantization technique (502), inverse transform technique (503), and inverse prediction, as explained in FIG5. The decoded frames of the video are then bit streamed with high-quality frames of the video.
- [96] Figure 9a is a block diagram that illustrates details of the encoder for any codec, according to the embodiment disclosed herein. Figure 9b is a block diagram that illustrates details of the encoder for any codec, according to the embodiment disclosed herein. Figure 9c is a block diagram that illustrates details of the encoder for any codec, according to the embodiment disclosed herein.
- [97] Figure 9a is a block diagram that illustrates the typical encoder for any codec. The process of encoding includes the selection of frames of the video and process through the multi pre-processors (311). Further, send through the backbone network (419) as disclosed in the Figure 4. The reconstructed frame of the video, devoid of any output artefacts, is obtained by utilizing a multi pre-processor system (311). Each pre-processor is assigned a specific index and the resulting reconstructed frame is fed into the AILF (318). Meanwhile, for convenience of explanation, the in-loop filter that can include AILF is described as AILF. However, AILF may be configured in series

or parallel with at least one of LMCS, deblocking filter, SAO, ALF, and CC-ALF included in the in-loop filter, and is not limited to the disclosed example.

[98] Figure 9b is a block diagram that illustrates the operations performed in AILF according to an embodiment of the disclosure. The AILF (318) then selects the most optimal frame, along with the corresponding index of the pre-processor that produced it. This final output is then directed towards further usage, potentially sent to the decoded picture buffer, while the pre-processor index is directed towards entropy coding for bitstream encoding. This process contributes significantly towards improving the video quality during the decoding process. The block diagram of AILF in FIG. 9b shows operations that can be performed in AI in loop filtering included in in loop filtering after de-quantization and inverse transform are performed, and the output of FIG. 9b is stored in the decoded buffer, or Can be used for prediction of other frames.

[99] The video input utilized in prediction techniques incorporates the intra prediction technique (610) in video encoding, which effectively exploits spatial redundancies within an individual video frame. The intra prediction technique (610) determines the pixel values of each frame of the video (610) by analyzing the values of neighbouring pixels in the video. The output of the intra prediction technique (610) is then fed into the video transform technique (603a), which operates on each frame of the video, converting pixel values from spatial information to the frequency domain. This transformation enhances crucial frequency components, enabling efficient compression by concentrating signal energy in a condensed set of coefficients. The transformed video is subsequently quantized, where each frame undergoes division, generating quantized coefficients by the quantization technique (603b). This quantization minimizes the data volume required for storage and transmission. The output from the quantization process (603b) is then fed into the entropy coding technique (604), the inverse transform technique (503), and the inverse quantization technique (502). The entropy coding technique (604) eliminates any statistical redundancy, enhancing the overall compression efficiency of the frames of the video. This technique encodes frames of the video based on the received index of the selected pre-processor.

[100] The technique of inverse transform (503) includes the quantized coefficients with spatial domain pixel values, while the original transformed coefficients are regained by multiplying the quantized coefficients with the quantization step size. The inverse quantization technique (502) effectively mitigates the distortion or loss of information caused by the quantized coefficients in the quantized video. The AILF (318) receives input from the inverse quantization technique (502) and subsequently processes it to furnish an encoded, artifact-free reconstructed frame to the decoded picture buffer.

- [101] Figure 9c is a block diagram that illustrates the backbone network according to an embodiment of the disclosure. The backbone network may be used in AILF after processing by pre-processor. And, the backbone network may be used network of Figure 4b or 7, but, it is not limited to the disclosed examples.
- [102] Figure 10 is a block diagram that illustrates details of the decoder for any codec, according to the embodiment disclosed herein. The decoder (317) receives the selected pre-processor from the multi pre-processors (311) with index and used it for decoding the encoded frames of the video. The electronic device can determine a pre-processor according to the obtained index and obtain a set of ratios for each channel according to the determined pre-processor. Additionally, the obtained set of ratios for each channel can be used for AILF.
- [103] The encoded frames of the video passes through the inverse transform technique (503) converts the quantized coefficients into spatial domain pixel values. The quantized coefficients are multiplied by the quantization step size to recover the original transformed coefficients. The inverse quantization technique (502) reduces degree of distortion or loss of information produced by the quantized coefficients. The inverse prediction reconstructs pixel values based on previously encoded frames of the video, aiming to generate the original frames of the video by using predictions from the encoded data.
- [104] Figure 11a illustrates an image frame with a CTU and slice numbers, according to the embodiment disclosed herein. Coding Tree Units (CTUs) (815) can derive indices by referencing the context of CTUs (815). Additionally, CTUs (815) may exhibit signalling based on predefined conditions, such as sending an index only when the CTU (815) demonstrates specific characteristics like edge magnitude or variance. In cases where these conditions are not met, default indices, either hardcoded or sent at a higher level of abstraction such as a slice header (811), a picture header (810), and a sequence header (801), are employed. Indices can be transmitted in slice headers (811), the picture header (810), and the sequence header (801), as exemplified in a specific instance within the Picture Parameter Set (PPS). The CTUs (815), the slice header (811), the picture header (810) and the index pertains to the coding parameters or characteristics during encoding process. Meanwhile, when an index for pre-processing is obtained from the sequence header, a pre-processor may be determined for each sequence. If an index for pre-processing is obtained from the picture header, a pre-processor may be determined for each picture. If an index for preprocessing is obtained from the slice header, a preprocessor may be determined for each slice. Alternatively, the pre-processor may be determined for each CTU, or it may be determined for each PU, TU, and CU.

- [105] Figure 11b is a block diagram that illustrates the extracted feature maps with indices, according to the embodiment disclosed herein. The sequence header (801) that incorporates a Sequence Pre-Processor flag is 0 (802). When the flag is set to 0, it signifies the utilization of the default Pre-Processor, while a flag value of 1 indicates the selection of the pre-processor from the multi pre-processors (311) with X indexed. Additionally, when the flag is set to 2, it functions as input to the Picture header (810). Similarly, the picture header (810) uses the picture pre-processor flag with 0 or 1 or 2 (807, 808, 809), the slice header (811) uses picture pre-processor flag with 0 or 1 or 2 (812, 813, 814), and CTU (815) uses CTU pre-processor flag with 0 or slice pre-processor flag with 0 or 1.
- [106] Figure 12 is a flow diagram illustrating a method for compressing the video using the AI model, according to the embodiment disclosed herein. At step 121, the method includes obtaining a video comprising a plurality of frames, wherein each frame of the pluralities of frames comprises a plurality of channels. Each frames of the plurality of the frames are having image information to be encoded by an encoder (316). The encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor.
- [107] At step 122, the method includes extracting at least one feature from each frame of the plurality of frames. The extracted features including the original patch (201), the partitioned patch (202), the reconstructed patch (203), the prediction patch (204), and the boundary strength of a patch (205). These features serve as inputs to the multi pre-processors (311), which analyses and produces the optimal output based on the most relevant feature patch. The original patch (201) represents unprocessed video data and acts as the baseline for subsequent processes. The partition patch (202) undergoes segmentation, dividing the patch into smaller regions to isolate foreground from background or extract specific features. The reconstructed patch (203) results from reconstructing the video to a bitstream using information from patches, including the partition patch (202) and the prediction (204). The prediction patch (204) anticipates the original patch (201) based on available information, enhancing the quality of the reconstructed patch (203). The boundary strength of patch (205) indicates confidence in patch boundaries, crucial for segmentation, object detection, or motion estimation, with higher strength indicating a clearer distinction between regions in the reconstructed patch (203).
- [108] At step 123, the method includes selecting at least one pre-processor from the multi pre-processors (311) for applying the AILF (318) based on the at least one feature from each frame of the plurality of frames. Each pre-processor of the multi pre-processors (311) comprises a set of ratios each of which corresponds to each channel of the plurality of frames of each frame of the plurality of frames, and wherein the at least

one selected pre-processor (311) comprises an optimal set of ratios for encoding the image information from each channel. The electronic device may select the at least one pre-processor from the multi pre-processors for applying the AILF to enhance the each frame of the pluralities of frames.

[109] At step 124, the method may include generating an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor for the AILF (318). The method includes generating an encoded video by encoding the frame of the video at an AI-based In Loop stage. In this step, the intra prediction technique (610) and inter prediction technique (512) processes. Subsequently, it performs transformation and quantization on the video blocks based on the predictions. The device generates quantized coefficients for the frames of the video. Following this, it reconstructs the frames of the video using the produced quantized coefficients. The reconstructed video is then transmitted through the AILF (318) for artifact removal, resulting in the artifact-free reconstructed video. The electronic device perform in-loop filtering the reconstructed video with AILF to get an artefact removed reconstructed video, wherein the artefact removed reconstructed video is quality enhanced video. And, the electronic device may perform entropy coding to encode the artefact removed reconstructed video.

[110] At step 125, the method includes transmitting the encoded video and an index corresponding to the at least one selected pre-processor to a decoder. The decoder (317) receives the selected pre-processor from the multi pre-processors (311) with index and used it for decoding the encoded frames of the video. The encoded frames of the video passes through the inverse transform technique (503), transforming quantized coefficients back into spatial domain pixel values. By multiplying the quantized coefficients with the quantization step size, the original transformed coefficients are recovered. The inverse quantization technique (502) minimizes distortion or loss of information caused by the quantized coefficients. Additionally, the inverse prediction reconstructs pixel values by utilizing information from previously encoded frames, with the goal of generating the original frames of the video through predictions derived from the encoded data.

[111] In an embodiment of the disclosure, the objectives are achieved by providing a method managing multi pre-processors for AILF. The method includes receiving, by an electronic device, a video comprising a plurality of frames. Each of the frames includes a plurality of channels before applying the AILF and each of the channels includes image information to be encoded by an encoder of the electronic device. Further, the method includes extracting, by the electronic device, at least one feature from each frame of the plurality of frames of the video. Further, the method includes

selecting, by the electronic device, at least one pre-processor from the multi pre-processors for applying the AILF to enhance the encoding the image information. The pre-processors are selected based on the at least one extracted features. Each pre-processor of the multi pre-processors comprises a set of ratios each of which corresponds to each channel of the plurality of frames of each frame of the plurality of frames, and wherein the at least one selected pre-processor comprises an optimal set of ratios for encoding the image information from each channel. The method includes generating, by the electronic device, an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor. The method includes transmitting, by the electronic device, the encoded video and an index corresponding to the at least one selected pre-processor to a decoder of the electronic device.

[112] In an embodiment, encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames at AI based In Loop stage at an AI-based in loop stage. The encoding includes dividing the received video into various blocks to perform an intra prediction and an inter prediction. Further, performing transformation and quantization based on the intra predicted and the inter predicted video blocks and generating quantized coefficients for the divided video blocks which have undergone the intra prediction and the inter prediction. Further, reconstructing the divided video blocks based on the generated quantized coefficients. Further, sending the reconstructed video through the AI model to get an artefact removed reconstructed video and sending the artefact removed reconstructed video through entropy coding to encode.

[113] In an embodiment, the method includes receiving the encoded video, and the index corresponding to the at least one selected pre-processor for determining the at least one selected pre-processor for decoding the encoded video based on the index corresponding to the at least one selected pre-processor and decoding the encoded video using the optimal set of ratios of the at least one selected pre-processor.

[114] In an embodiment, encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor comprises embedding the optimal set of ratios of the at least one selected pre-processor into the encoded image information for use by the decoder for decoding of the encoded video and inputting the pre-processed image information from each channel and the frame to the AI in-loop filter for generating the encoded video.

[115] In an embodiment, the optimal set of ratios of the at least one selected pre-processor is embedded in at least one of a sequence header, a picture header, a slice header, and a Coding Tree Unit (CTU) of each frame of the plurality of frames.

- [116] In an embodiment, the plurality of channels comprises, a luma reconstruction, a prediction buffer, a boundary strength, a Quantization Parameter(QP) base, a QP Slice, and a block coding type.
- [117] Accordingly, the embodiment herein is to provide an electronic device for managing multi pre-processors for AILF in a video codec. The electronic device comprises the multi pre-processors connected to an encoder and a decoder, a memory comprising a video to be encoded, an assemblage block controller, coupled to the memory and the multi pre-processors. The assemblage block controller is configured to receive a video comprising a plurality of frames. Each of the frames includes a plurality of channels before applying the AILF and each of the channels includes image information to be encoded by an encoder of the electronic device. Further, extract at least one feature from each frame of the plurality of frames of the video. Further, the assemblage block controller selects at least one pre-processor from the multi pre-processors for applying the AILF based the at least one extracted features. Each of the multi pre-processors includes a set of ratios each of which corresponds to each channel of the plurality of frames of each frame of the plurality of frames. The selected pre-processor includes an optimal set of ratios for encoding the image information from each channel. The electronic device may the assemblage block controller generates an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor for the AILF. The electronic device may transmit the encoded video and an index corresponding to the at least one selected pre-processor to a decoder of the electronic device.
- [118] In an embodiment of the disclosure, a method for managing multi pre-processors (311) for an AI-based In Loop Filter(AILF) (318) in a video codec, may include obtaining, by an electronic device (310), a video comprising a plurality of frames, wherein each frame of the pluralities of frames comprises a plurality of channels before applying the AILF (318), and wherein each of the channels comprises image information. The method may include extracting, by the electronic device (310), at least one feature from each frame of the plurality of frames. The method may include selecting, by the electronic device (30), at least one pre-processor from the multi pre-processors (311) for applying the AILF (318) based on the at least one extracted feature, wherein each pre-processor of the multi pre-processors (311) comprises a set of ratios each of which corresponds to each channel of the plurality of channels of each frame of the plurality of frames. The method may include generating, by the electronic device (310), an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor for the AILF (318). The method may include transmitting,

by the electronic device (310), the encoded video and an index corresponding to the at least one selected pre-processor to a decoder.

[119] In an embodiment of the disclosure, the method may include dividing, by the electronic device (310), the received video into various blocks to perform an intra prediction and an inter prediction. The method may include performing, by the electronic device (310), a transformation (603a) of the video blocks based on the intra predicted and the inter predicted video blocks. The method may include performing, by the electronic device (310), a quantization (603b) on the video blocks based on the intra predicted (610) and the inter predicted (512) video blocks. The method may include generating, by the electronic device (310), quantized coefficients for the divided video blocks which have undergone the intra prediction (610) and the inter prediction (512), wherein the quantization divides the transformed video blocks and generates the quantized coefficients. The method may include reconstructing, by the electronic device (310), the divided video blocks based on the generated quantized coefficients. The method may include performing by the electronic device (310), in-loop filtering the reconstructed video with the AILF (318) to get an artefact removed reconstructed video, wherein the artefact removed reconstructed video is quality enhanced video. The method may include performing, by the electronic device (310), entropy coding (604) to encode the artefact removed reconstructed video.

[120] In an embodiment of the disclosure, the method may include embedding, by the electronic device (310), the set of ratios of the at least one selected pre-processor into the encoded image information. The method may include inputting, by the electronic device (310), the pre-processed image information from each channel and the frame to the AILF (318).

[121] In an embodiment of the disclosure, the set of ratios for the at least one selected pre-processor may be embedded in at least one of a sequence header (801), a picture header (810), a slice header (811), and a Coding Tree Unit (CTU) (815) of each frame of the plurality of frames.

[122] In an embodiment of the disclosure, the plurality of channels may comprise a luma reconstruction, a prediction buffer (105), a boundary strength (205), a Quantization Parameter (QP) base (402), a QP Slice (401), and a block coding type.

[123] In an embodiment of the disclosure, the method may include performing AILF after at least one of LMCS, deblocking filtering, SAO, ALF, and CC-ALF is performed.

[124] In an embodiment of the disclosure, the at least one feature from each frame of the plurality of frames is pre-defined.

[125] In an embodiment of the disclosure, a method for managing multi pre-processors (311) for an AI-based In Loop Filter(AILF) (318) in a video codec, may include obtaining, by an electronic device (310), encoded video and index corresponding

to at least one selected pre-processor. The method may include determining, by the electronic device (310), the at least one selected pre-processor for decoding the encoded video based on the index corresponding to the at least one selected pre-processor. The method may include decoding, by the electronic device (310), the encoded video using a set of ratios of the at least one selected pre-processor.

[126] In an embodiment of the disclosure, an electronic device (310) for managing multi pre-processors (311) for an AI-based In Loop Filter (AILF) (318) in a video codec, comprising: a memory (313) comprising a video to be encoded and storing one or more instructions; and at least one processor including the multi pre-processors (311), configured to execute the one or more instructions. The at least one processor may be configured to execute the one or more instructions to obtain a video comprising a plurality of frames, wherein each frame of the pluralities of frames comprises a plurality of channels before applying the AILF (318), and wherein each of the channels comprises image information. The at least one processor may be configured to extract at least one feature from each frame of the plurality of frames. The at least one processor may be configured to select at least one pre-processor from the multi pre-processors (311) for applying the AILF (318) based on the at least one extracted feature, wherein each pre-processor of the multi pre-processors (311) comprises a set of ratios each of which corresponds to each channel of the plurality of frames of each frame of the plurality of frames. The at least one processor may be configured to generate an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor for the AILF (318). The at least one processor may be configured to transmit the encoded video and an index corresponding to the at least one selected pre-processor to a decoder.

[127] In an embodiment of the disclosure, the at least one processor is configured to divide the received video into various blocks to perform an intra prediction and an inter prediction. The at least one processor is configured to perform a transformation (603a) of the video blocks based on the intra predicted and the inter predicted video blocks. The at least one processor is configured to perform a quantization (603b) on the video blocks based on the intra predicted (610) and the inter predicted (512) video blocks. The at least one processor is configured to generate quantized coefficients for the divided video blocks which have undergone the intra prediction (610) and the inter prediction (512), wherein the quantization divides the transformed video blocks and generates the quantized coefficients. The at least one processor is configured to reconstruct the divided video blocks based on the generated quantized coefficients. The at least one processor is configured to perform in-loop filtering the reconstructed video with the AILF (318) to get an artefact removed reconstructed video, wherein

the artefact removed reconstructed video is quality enhanced video. The at least one processor is configured to perform entropy coding (604) to encode the artefact removed reconstructed video through entropy coding (604) to encode.

- [128] In an embodiment of the disclosure, the at least one processor is configured to embed the set of ratios of the at least one selected pre-processor into the encoded image information. The at least one processor is configured to input the pre-processed image information from each channel and the frame to an AILF (318).
- [129] In an embodiment of the disclosure, the set of ratios for the at least one selected pre-processor is embedded in at least one of a sequence header (801), a picture header (810), a slice header (811), and a Coding Tree Unit (CTU) (815) of each frame of the plurality of frames.
- [130] In an embodiment of the disclosure, the plurality of channels comprises a luma reconstruction, a prediction buffer (105), a boundary strength (205), a Quantization Parameter (QP) base (402), a QP Slice (401), and a block coding type.
- [131] In an embodiment of the disclosure, the at least one processor is configured to perform AILF after at least one of LMCS, deblocking filtering, SAO, ALF, and CC-ALF is performed.
- [132] In an embodiment of the disclosure, the at least one feature from each frame of the plurality of frames is pre-defined.
- [133] The foregoing description of the specific embodiments will so fully reveal the general nature of the embodiments herein that others can, by applying current knowledge, readily modify and or adapt for various applications such specific embodiments without departing from the generic concept, and, therefore, such adaptations and modifications are intended to be comprehended within the meaning and range of equivalents of the disclosed embodiments. It is to be understood that the phraseology or terminology employed herein is for the purpose of description and not of limitation. Therefore, while the embodiments herein have been described in terms of preferred embodiments, those skilled in the art will recognize that the embodiments herein can be practiced with modification within the scope of the embodiments as described herein.

Claims

[Claim 1]

A method for managing multi pre-processors (311) for an AI-based In Loop Filter(AILF) (318) in a video codec, comprising: obtaining, by an electronic device (310), a video comprising a plurality of frames, wherein each frame of the pluralities of frames comprises a plurality of channels before applying the AILF (318), and wherein each of the channels comprises image information; extracting, by the electronic device (310), at least one feature from each frame of the plurality of frames; selecting, by the electronic device (30), at least one pre-processor from the multi pre-processors (311) for applying the AILF (318) based on the at least one extracted feature, wherein each pre-processor of the multi pre-processors (311) comprises a set of ratios each of which corresponds to each channel of the plurality of channels of each frame of the plurality of frames; generating, by the electronic device (310), an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor for the AILF (318); and transmitting, by the electronic device (310), the encoded video and an index corresponding to the at least one selected pre-processor to a decoder.

[Claim 2]

The method of claim 1, wherein encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames at an AI-based In Loop stage comprises: dividing, by the electronic device (310), the received video into various blocks to perform an intra prediction and an inter prediction; performing, by the electronic device (310), a transformation (603a) of the video blocks based on the intra predicted and the inter predicted video blocks; performing, by the electronic device (310), a quantization (603b) on the video blocks based on the intra predicted (610) and the inter predicted (512) video blocks; generating, by the electronic device (310), quantized coefficients for the divided video blocks which have undergone the intra prediction (610) and the inter prediction (512), wherein the

quantization divides the transformed video blocks and generates the quantized coefficients;
reconstructing, by the electronic device (310), the divided video blocks based on the generated quantized coefficients;
performing by the electronic device (310), in-loop filtering the reconstructed video with the AILF (318) to get an artefact removed reconstructed video, wherein the artefact removed reconstructed video is quality enhanced video; and
performing, by the electronic device (310), entropy coding (604) to encode the artefact removed reconstructed video .

- [Claim 3] The method of any one of claims 1 to 2, wherein encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor comprises:
embedding, by the electronic device (310), the set of ratios of the at least one selected pre-processor into the encoded image information; and
inputting, by the electronic device (310), the pre-processed image information from each channel and the frame to the AILF (318).
- [Claim 4] The method of any one of claims 1 to 3, wherein the set of ratios for the at least one selected pre-processor is embedded in at least one of a sequence header (801), a picture header (810), a slice header (811), and a Coding Tree Unit (CTU) (815) of each frame of the plurality of frames.
- [Claim 5] The method of any one of claims 1 to 4, wherein the plurality of channels comprises a luma reconstruction, a prediction buffer (105), a boundary strength (205), a Quantization Parameter (QP) base (402), a QP Slice (401), and a block coding type.
- [Claim 6] The method of any one of claims 1 to 5, wherein the performing of in-loop filtering the reconstructed video with AILF comprises: performing AILF after at least one of LMCS, deblocking filtering, SAO, ALF, and CC-ALF is performed.
- [Claim 7] The method of any one of claims 1 to 6, wherein the at least one feature from each frame of the plurality of frames is pre-defined.
- [Claim 8] A method for managing multi pre-processors (311) for an AILF (318) in a video codec, comprising:
obtaining, by an electronic device (310), encoded video and index corresponding to at least one selected pre-processor;

determining, by the electronic device (310), the at least one selected pre-processor for decoding the encoded video based on the index corresponding to the at least one selected pre-processor; and

decoding, by the electronic device (310), the encoded video using a set of ratios of the at least one selected pre-processor.

[Claim 9]

An electronic device (310) for managing multi pre-processors (311) for an AI-based In Loop Filter (AILF) (318) in a video codec, comprising: a memory (313) comprising a video to be encoded and storing one or more instructions; and

at least one processor including the multi pre-processors (311), configured to execute the one or more instructions to:

obtain a video comprising a plurality of frames, wherein each frame of the pluralities of frames comprises a plurality of channels before applying the AILF (318), and wherein each of the channels comprises image information;

extract at least one feature from each frame of the plurality of frames;

select at least one pre-processor from the multi pre-processors (311) for applying the AILF (318) based on the at least one extracted feature, wherein each pre-processor of the multi pre-processors (311) comprises a set of ratios each of which corresponds to each channel of the plurality of frames of each frame of the plurality of frames;

generate an encoded video by encoding the image information from each channel of the plurality of channels of each frame of the plurality of frames using the at least one selected pre-processor for the AILF (318); and

transmit the encoded video and an index corresponding to the at least one selected pre-processor to a decoder.

[Claim 10]

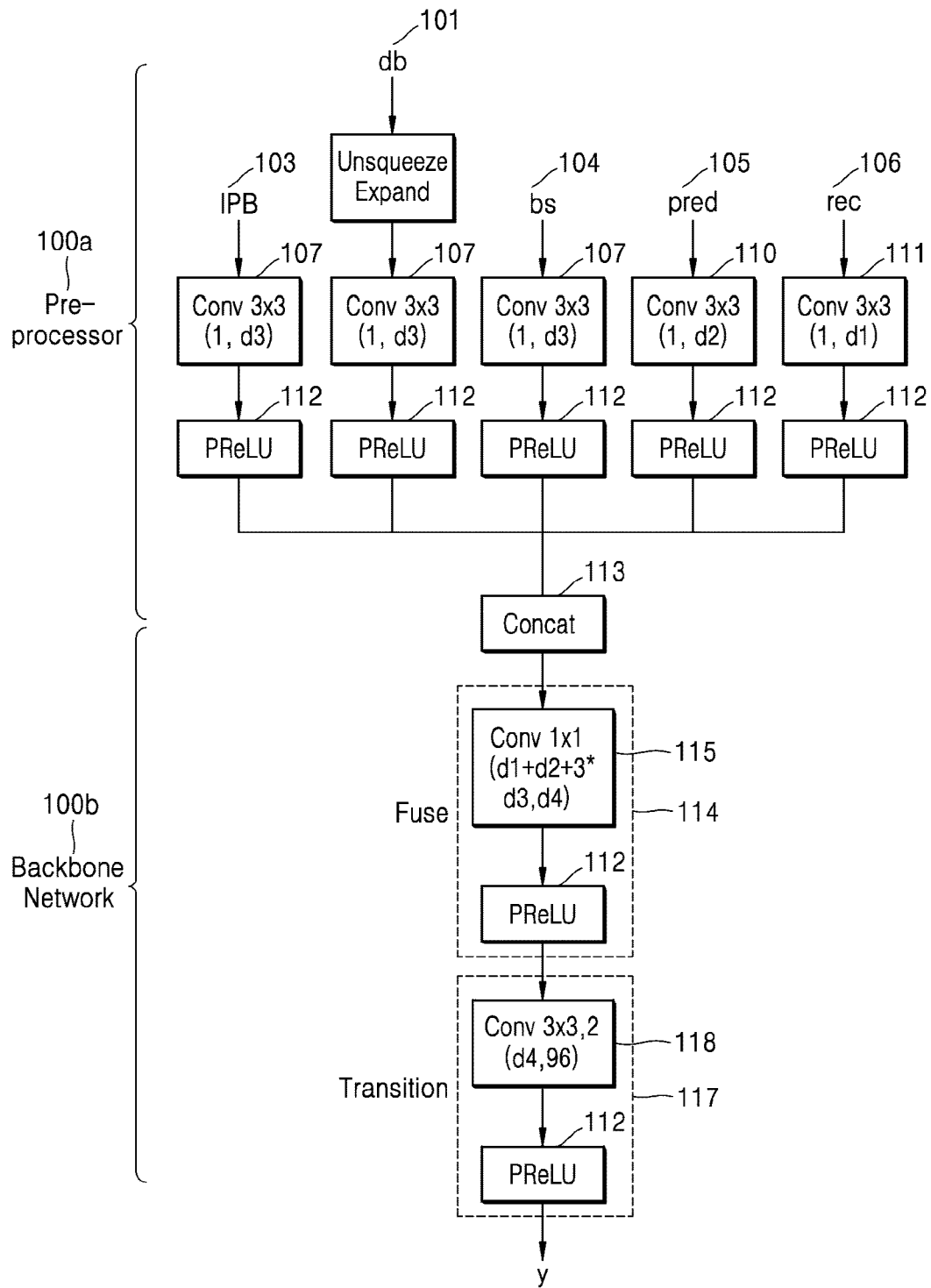
The electronic device (310) of claim 9, wherein the at least one processor is configured to execute the one or more instructions to: divide the received video into various blocks to perform an intra prediction and an inter prediction;

perform a transformation (603a) of the video blocks based on the intra predicted and the inter predicted video blocks, and

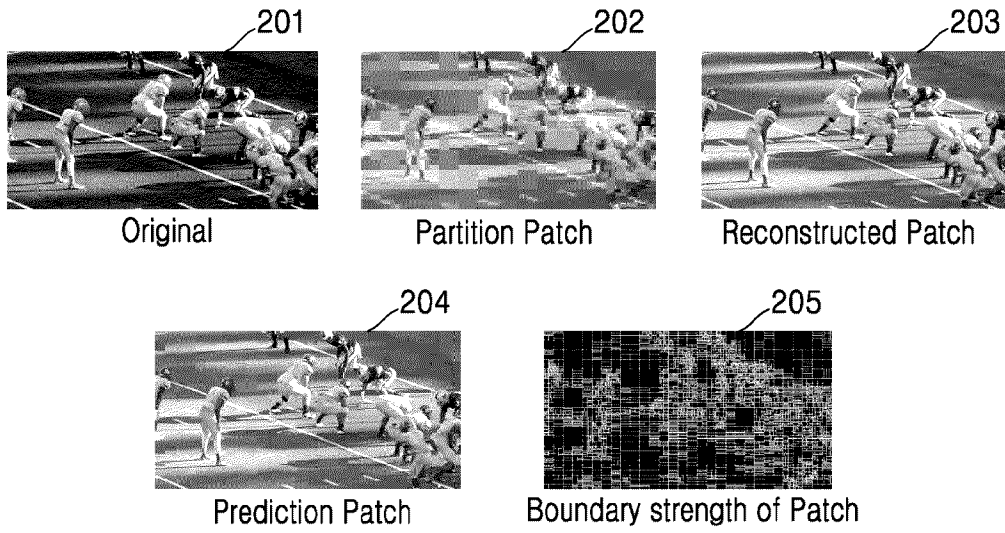
perform a quantization (603b) on the video blocks based on the intra predicted (610) and the inter predicted (512) video blocks;

- generate quantized coefficients for the divided video blocks which have undergone the intra prediction (610) and the inter prediction (512), wherein the quantization divides the transformed video blocks and the generates the quantized coefficients;
reconstruct the divided video blocks based on the generated quantized coefficients;
perform in-loop filtering the reconstructed video with the AILF (318) to get an artefact removed reconstructed video, wherein the artefact removed reconstructed video is quality enhanced video;
and
perform entropy coding (604) to encode the artefact removed reconstructed video through entropy coding (604) to encode.
- [Claim 11] The electronic device (310) of any one of claims 9 to 10, wherein the at least one processor is configured to execute the one or more instructions to:
embed the set of ratios of the at least one selected pre-processor into the encoded image information; and
input the pre-processed image information from each channel and the frame to an AILF (318).
- [Claim 12] The electronic device (310) of any one of claims 9 to 11, wherein the set of ratios for the at least one selected pre-processor is embedded in at least one of a sequence header (801), a picture header (810), a slice header (811), and a Coding Tree Unit (CTU) (815) of each frame of the plurality of frames.
- [Claim 13] The electronic device (310) of any one of claims 9 to 12, wherein the plurality of channels comprises a luma reconstruction, a prediction buffer (105), a boundary strength (205), a Quantization Parameter (QP) base (402), a QP Slice (401), and a block coding type.
- [Claim 14] The electronic device (310) of any one of claims 9 to 13, wherein the at least one processor is configured to execute the one or more instructions to:
perform AILF after at least one of LMCS, deblocking filtering, SAO, ALF, and CC-ALF is performed.
- [Claim 15] The method of any one of claims 9 to 14, wherein the at least one feature from each frame of the plurality of frames is pre-defined.

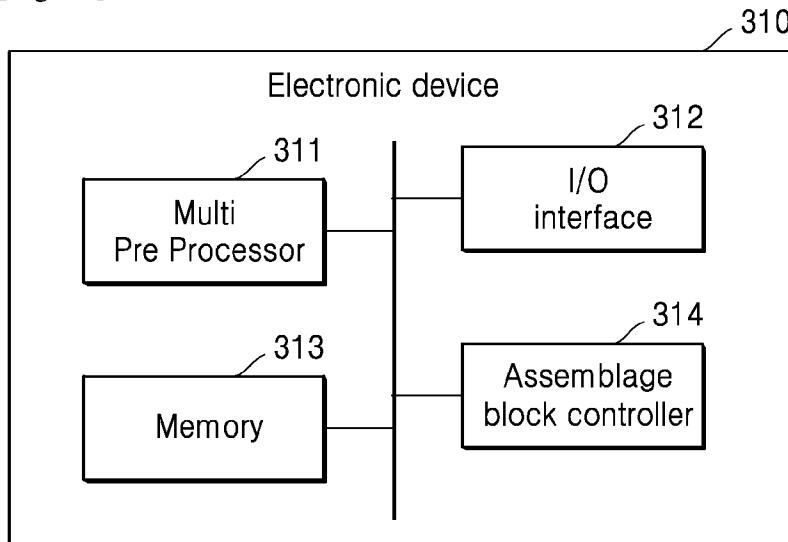
[Fig. 1]



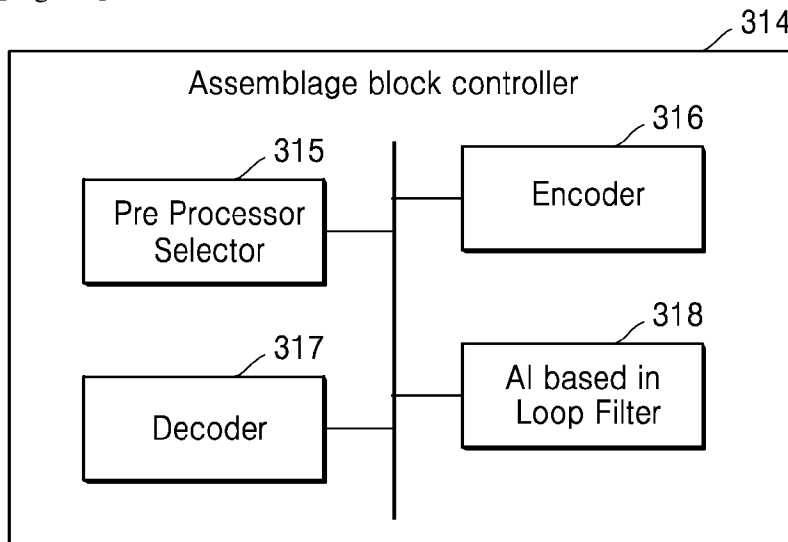
[Fig. 2]



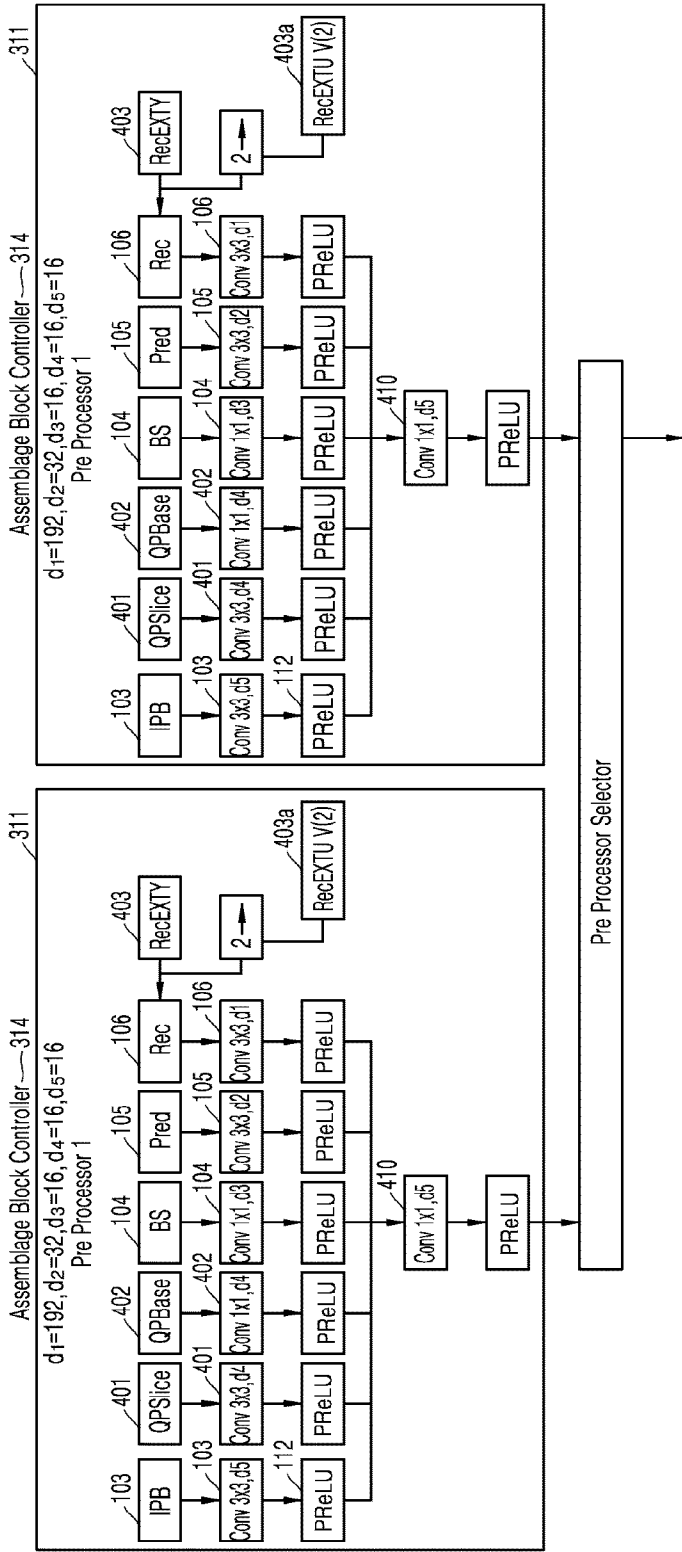
[Fig. 3a]



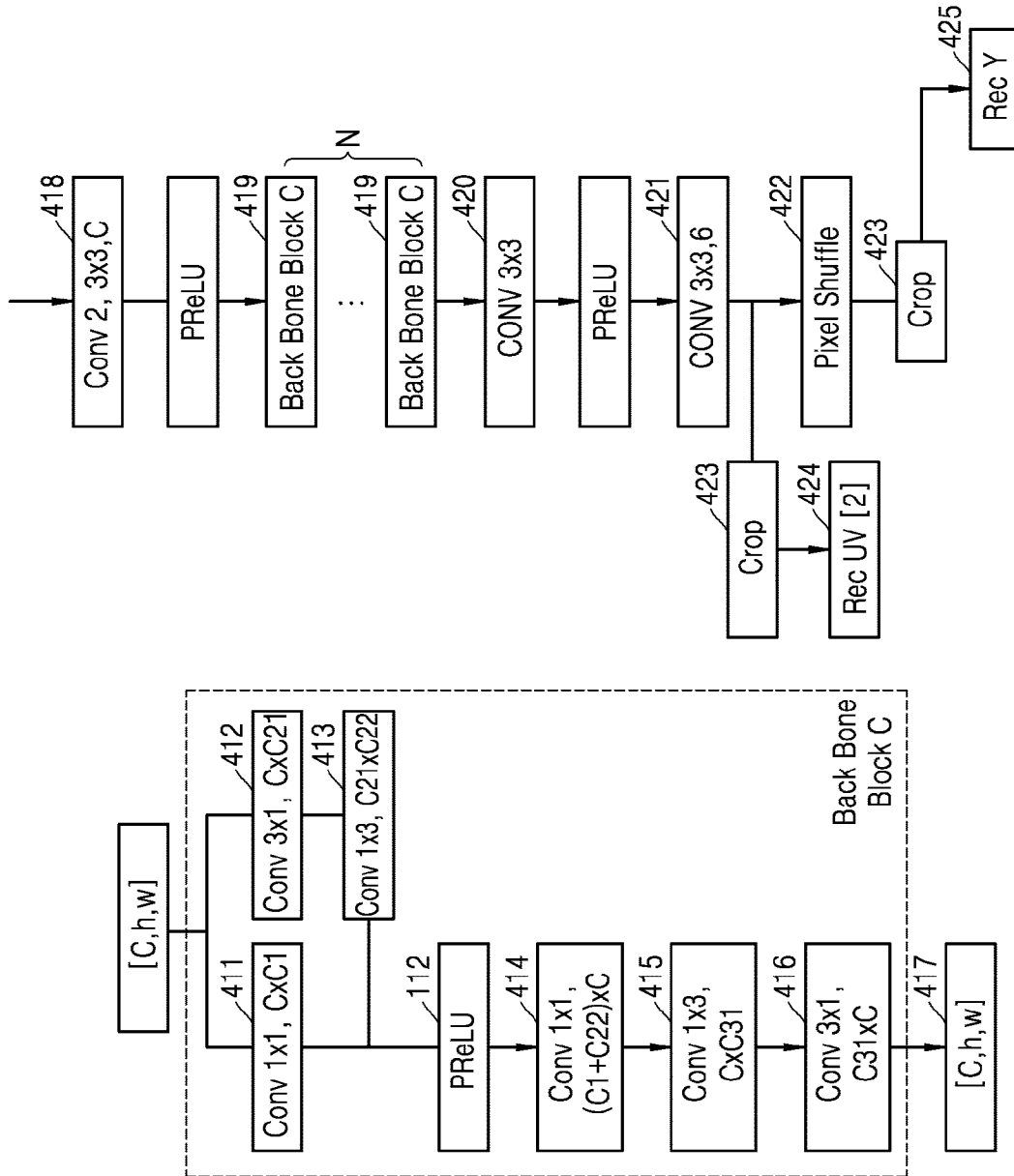
[Fig. 3b]



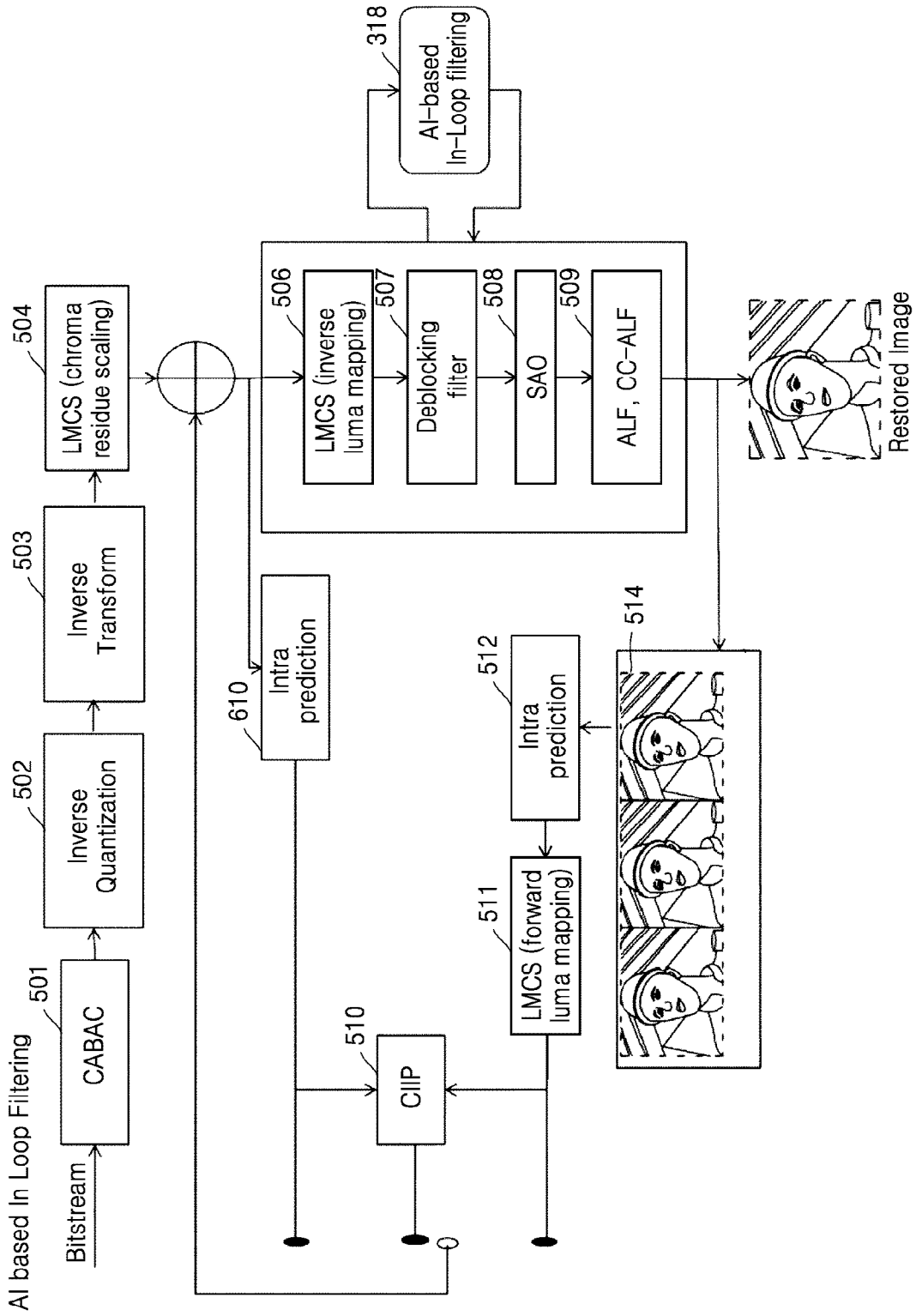
[Fig. 4a]



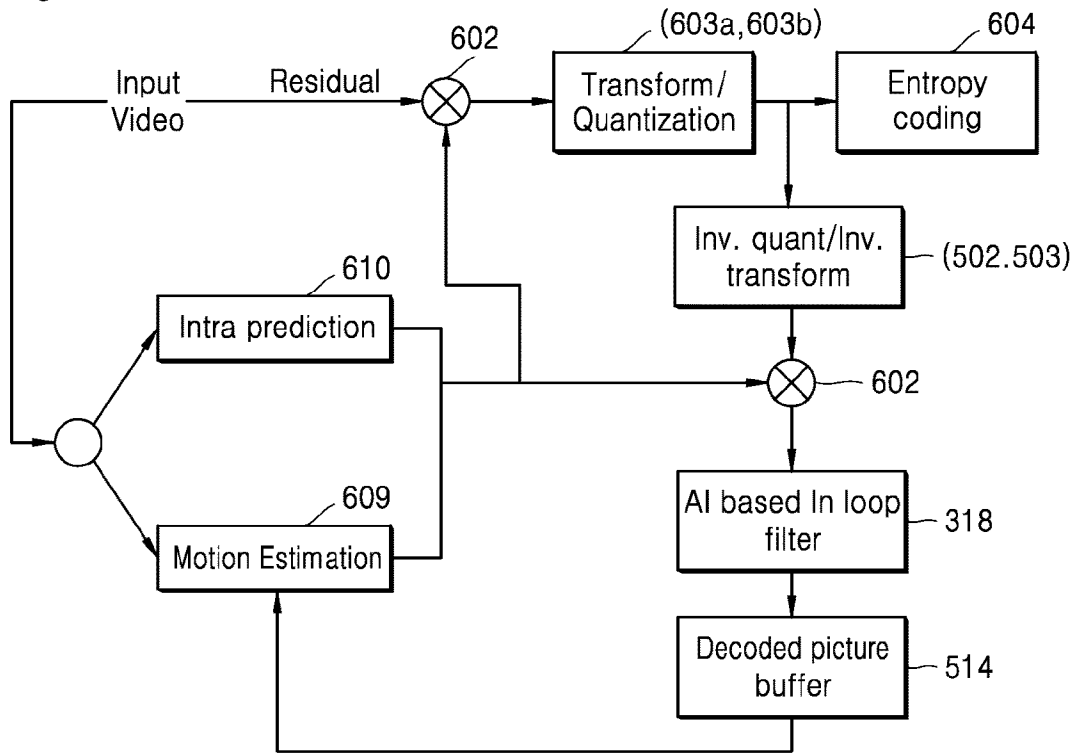
[Fig. 4b]



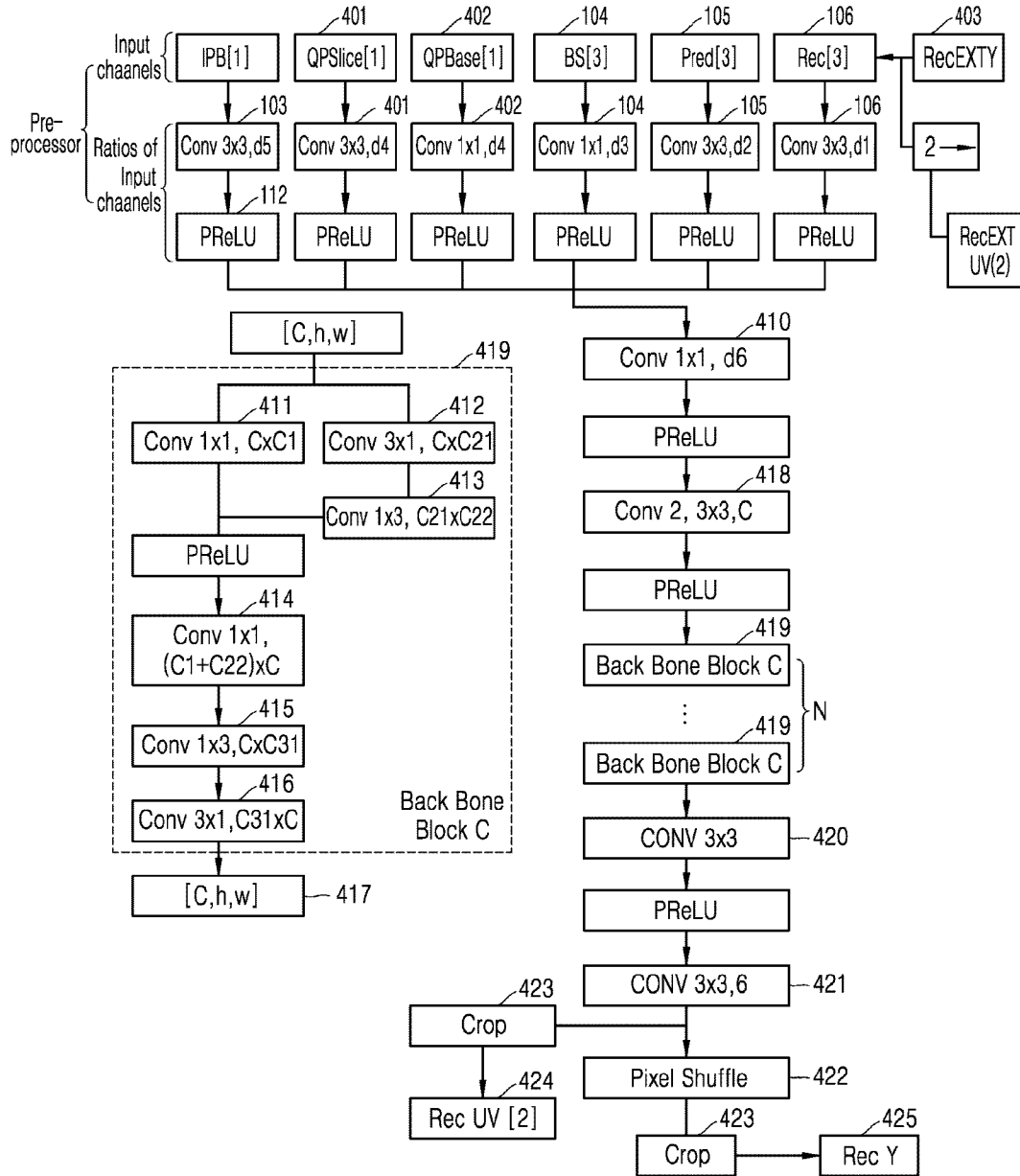
[Fig. 5]



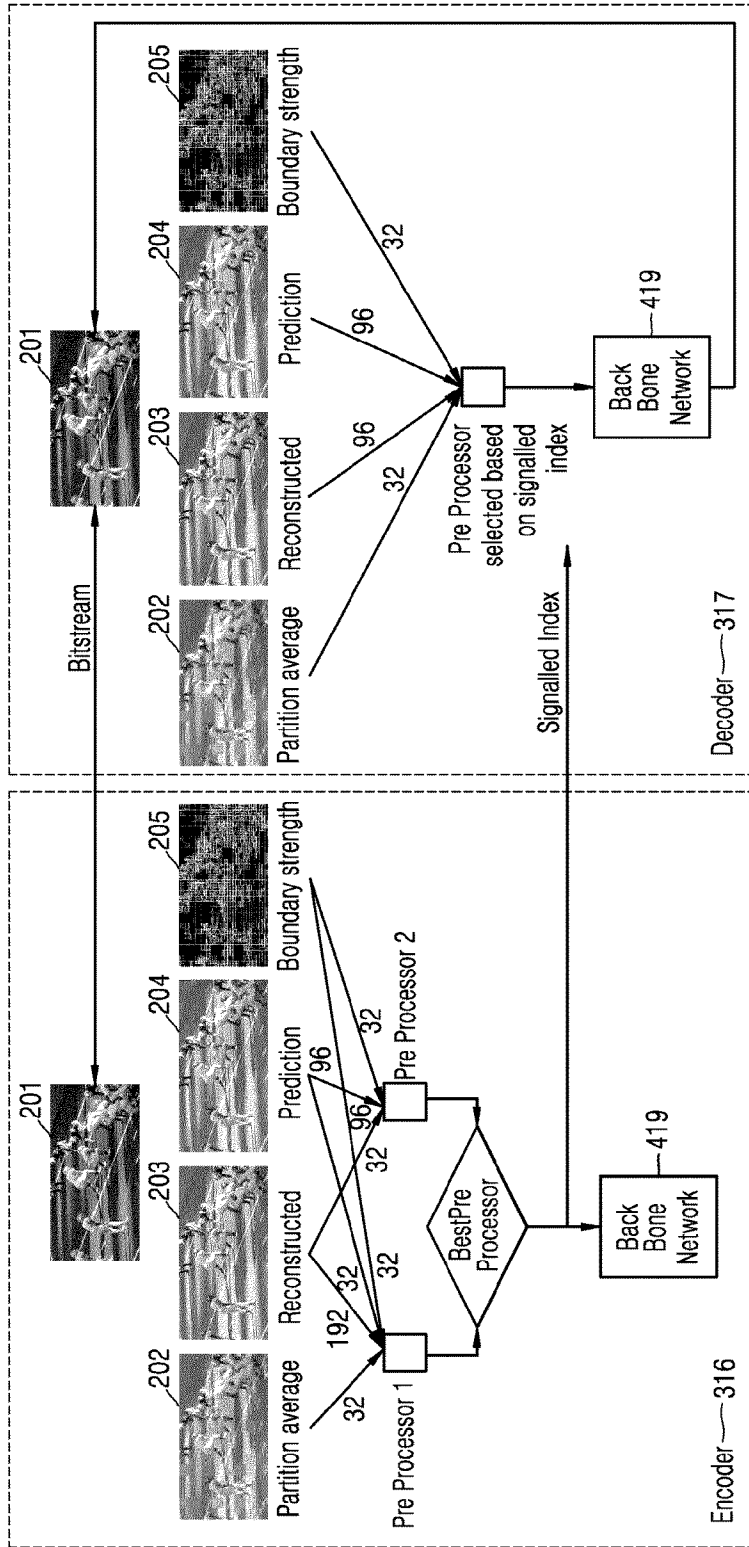
[Fig. 6]



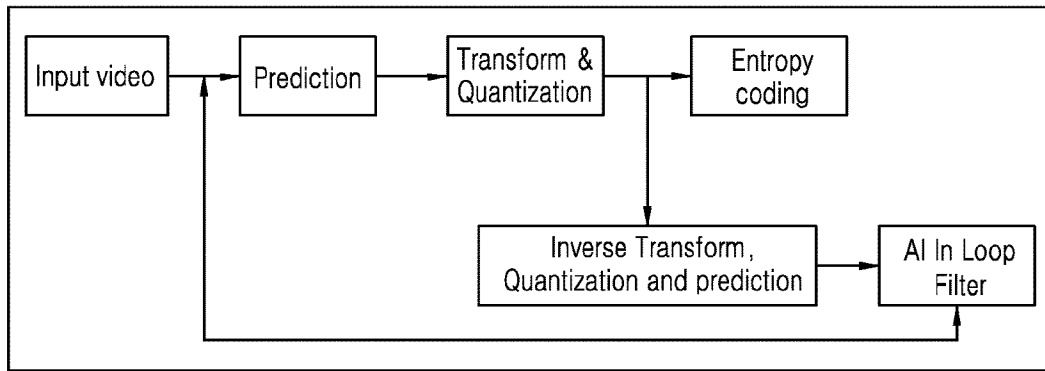
[Fig. 7]



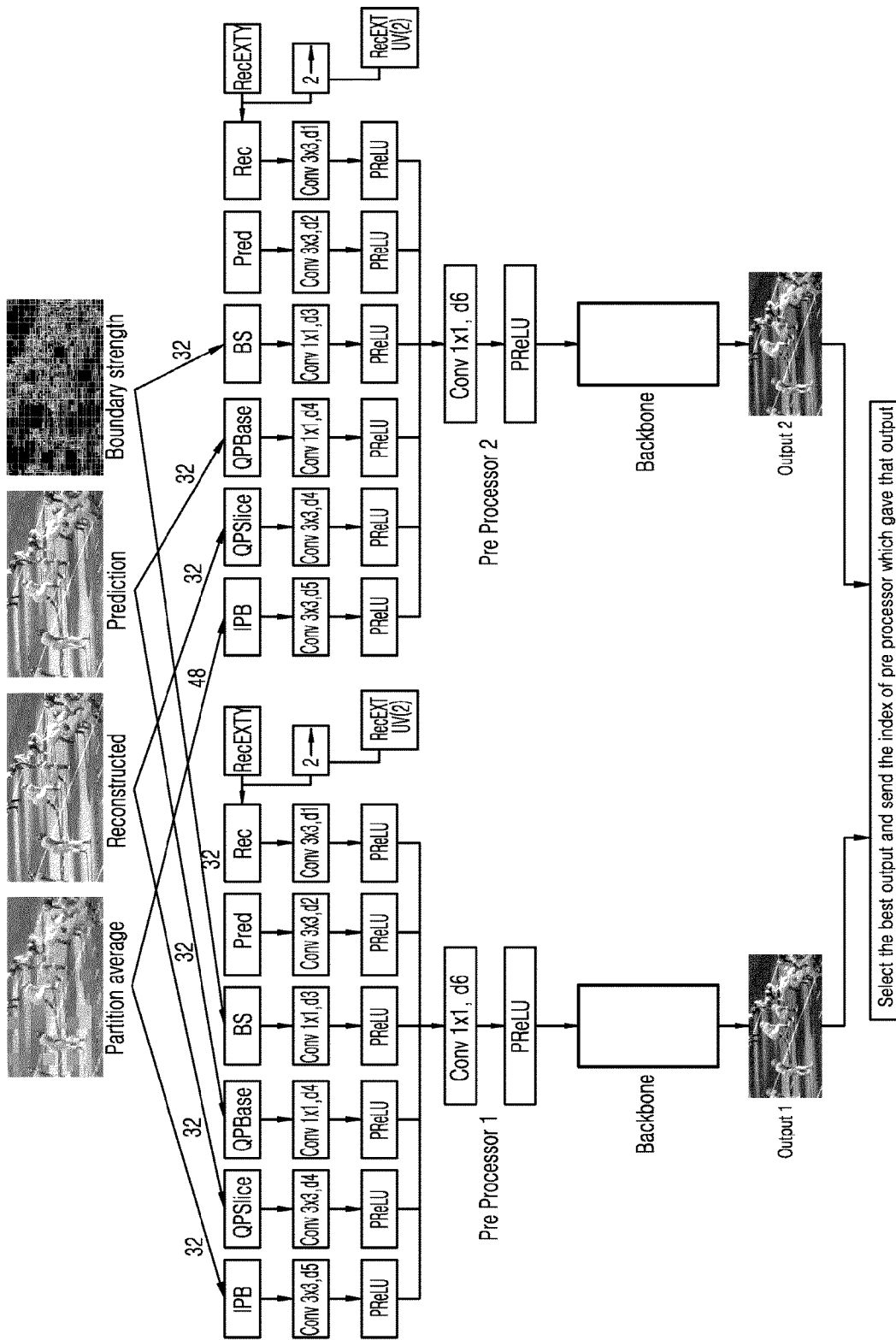
[Fig. 8]



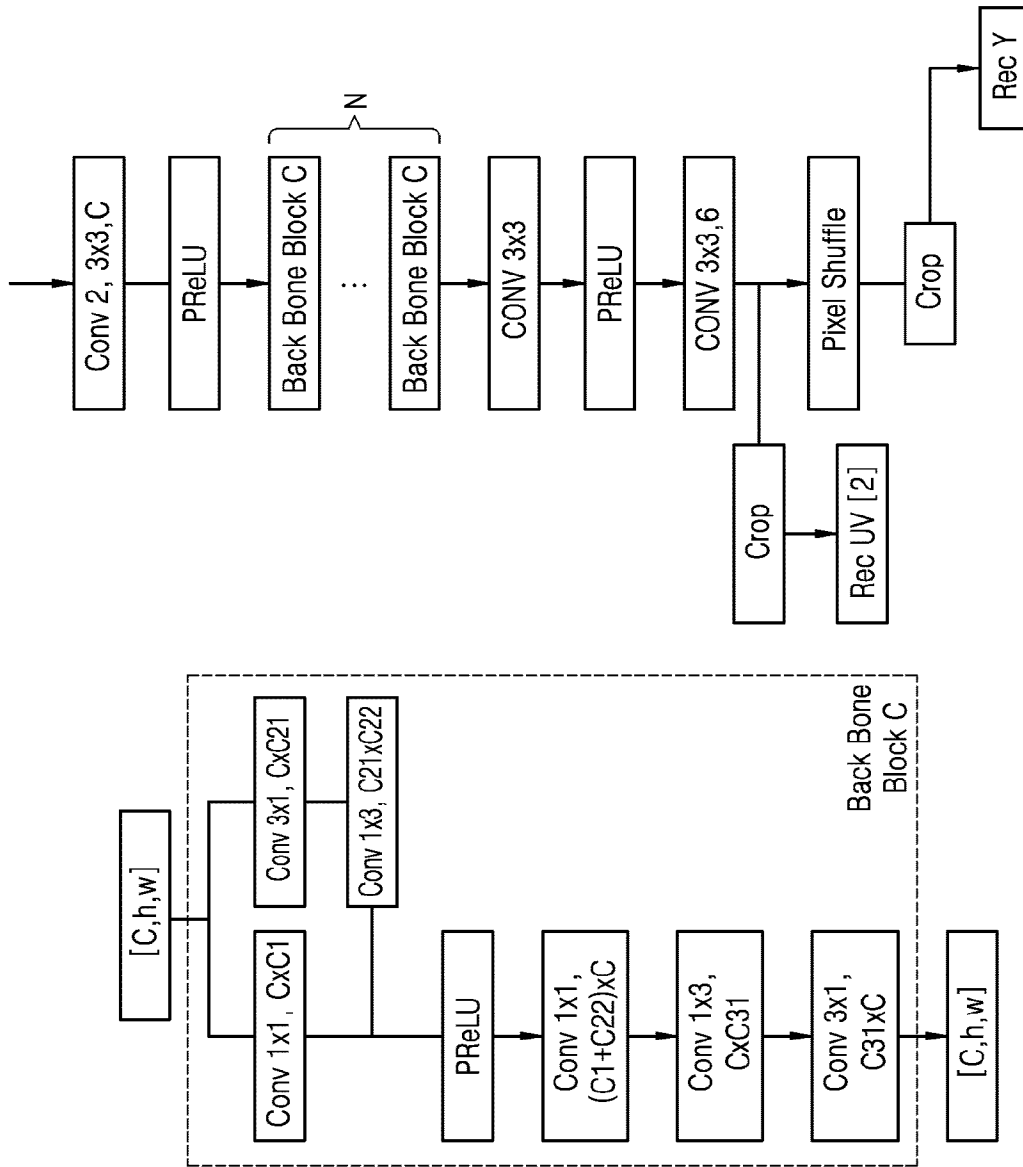
[Fig. 9a]



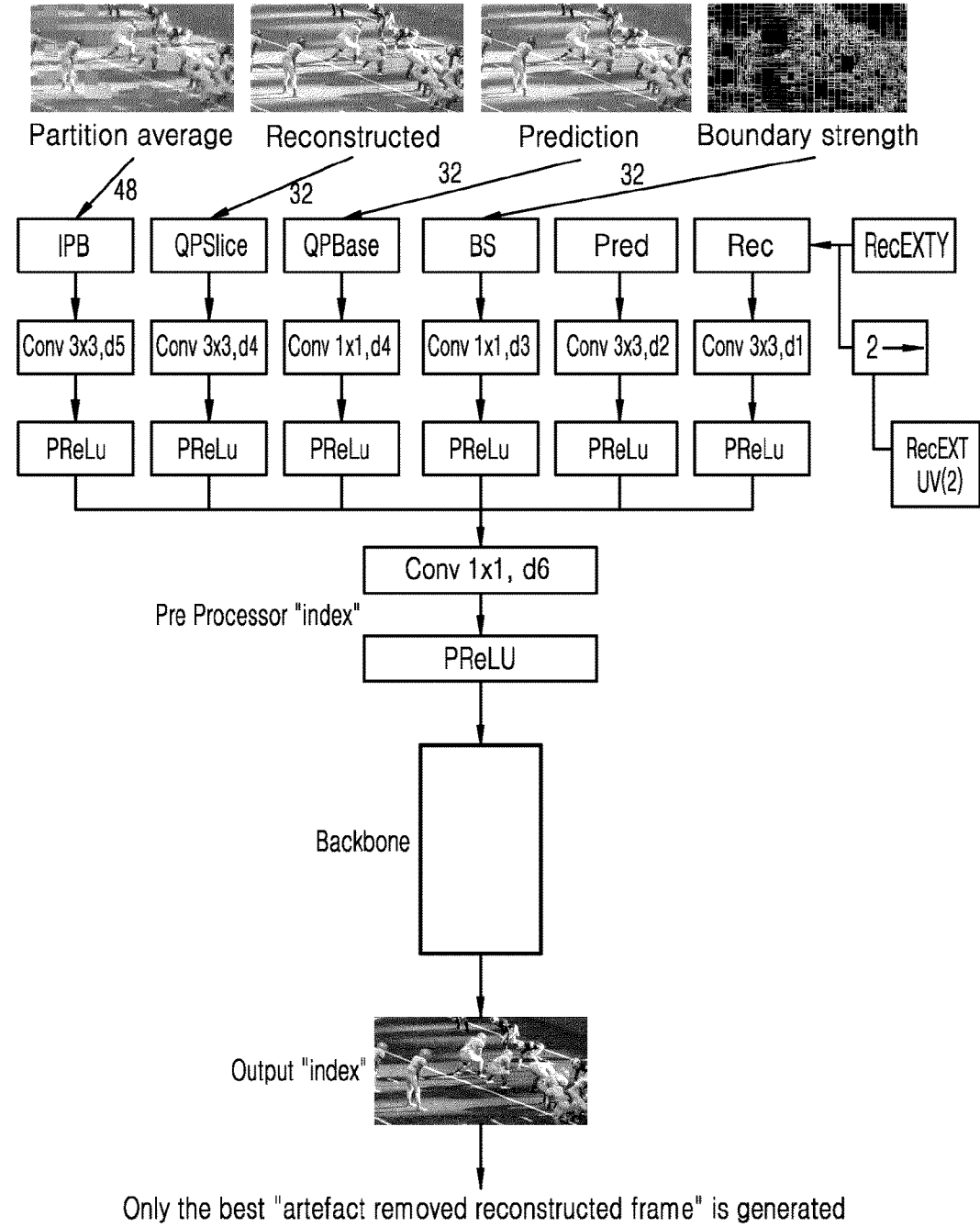
[Fig. 9b]



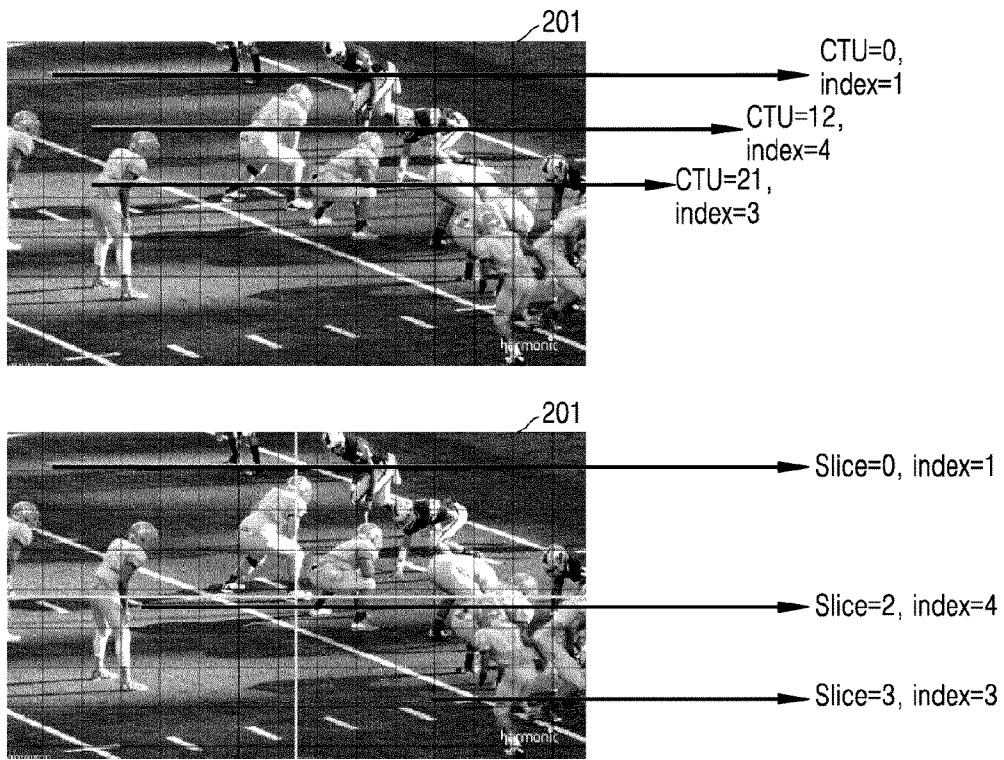
[Fig. 9c]



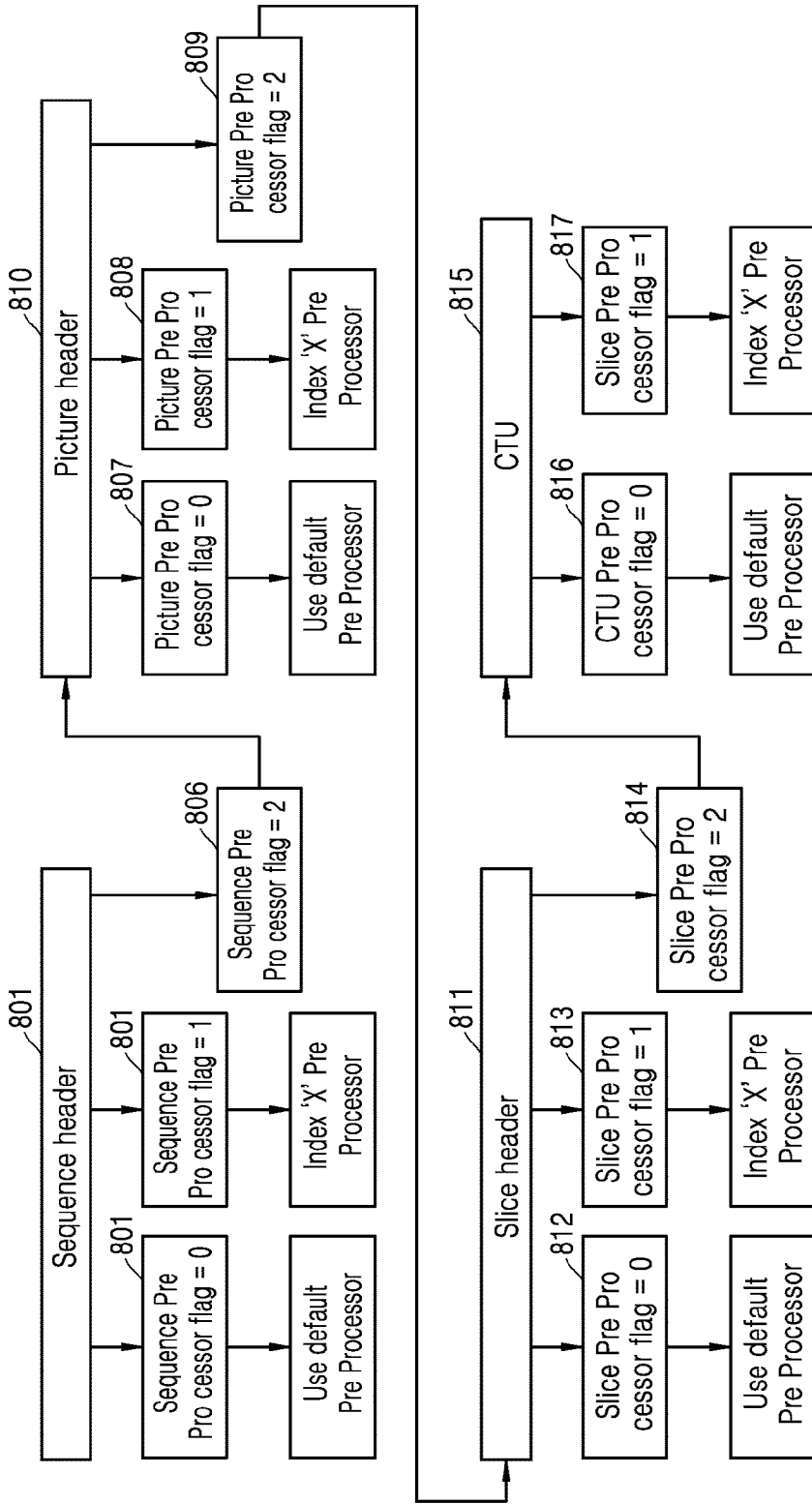
[Fig. 10]



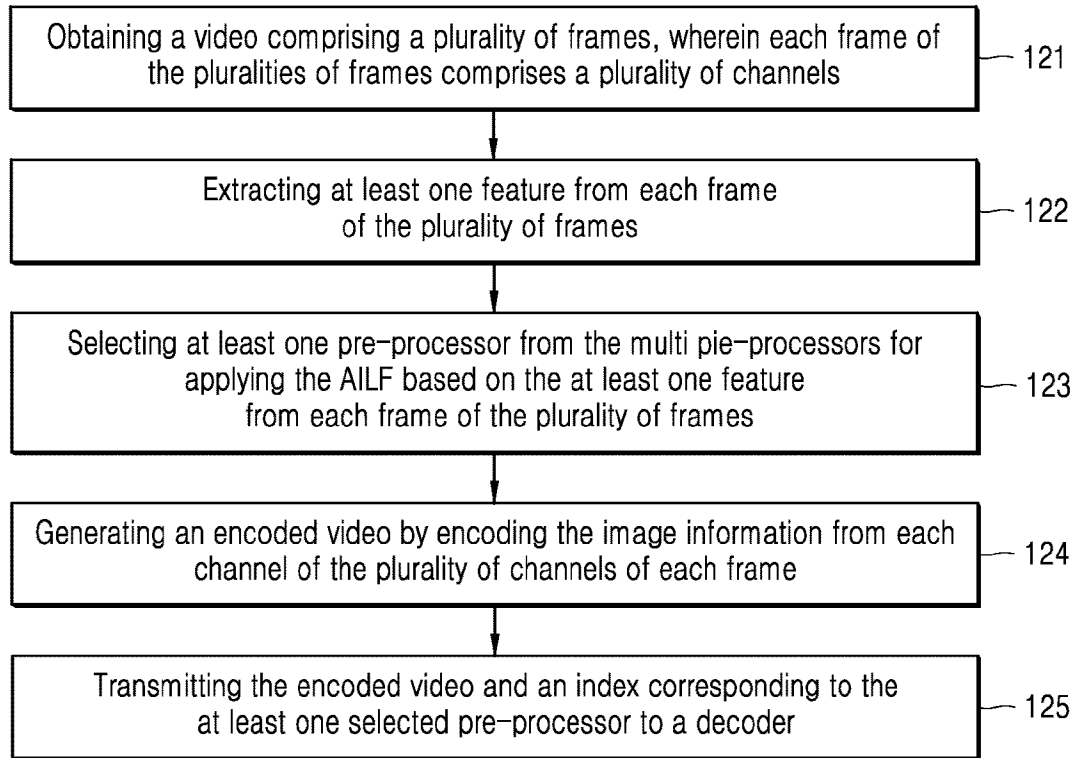
[Fig. 11a]



[Fig. 11b]



[Fig. 12]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2024/007733

A. CLASSIFICATION OF SUBJECT MATTER		
H04N 19/82(2014.01)i; H04N 19/86(2014.01)i; H04N 19/61(2014.01)i; H04N 19/124(2014.01)i; H04N 19/119(2014.01)i; H04N 19/13(2014.01)i		
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 19/82(2014.01); G06N 3/08(2006.01); G06T 5/00(2006.01); G06T 9/00(2006.01); H04N 19/117(2014.01)		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models Japanese utility models and applications for utility models		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS(KIPO internal) & Keywords: AI based In Loop Filter (AILF), a plurality of frames, feature extract, a plurality of channels, image information, embedding, a set of ratios, index, selected pre-processor, reconstructed video		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2022-0025689 A (HYUNDAI MOTOR COMPANY et al.) 03 March 2022 (2022-03-03) paragraph [0011]; and claims 1, 12	1-15
A	CHI DO-KIM PHAM et al., 'Deep Learning based Spatial-Temporal In-loop filtering for Versatile Video Coding', 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 19-25 June 2021 pages 1-5	1-15
A	KR 10-2018-0052651 A (MEDIATEK INC.) 18 May 2018 (2018-05-18) paragraphs [0024], [0029]; and claim 1	1-15
A	KR 10-2022-0136163 A (HYUNDAI MOTOR COMPANY et al.) 07 October 2022 (2022-10-07) claims 1-3	1-15
A	KR 10-2023-0026944 A (HYUNDAI MOTOR COMPANY et al.) 27 February 2023 (2023-02-27) paragraph [0010]	1-15
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 06 September 2024		Date of mailing of the international search report 06 September 2024
Name and mailing address of the ISA/KR Korean Intellectual Property Office 189 Cheongsa-ro, Seo-gu, Daejeon 35208, Republic of Korea Facsimile No. +82-42-481-8578		Authorized officer YANG, Jeong Rok Telephone No. +82-42-481-5709

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2024/007733

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
KR 10-2022-0025689	A	03 March 2022	CN	116113985	A	12 May 2023	
			US	2023-0269399	A1	24 August 2023	
			WO	2022-045738	A1	03 March 2022	

KR 10-2018-0052651	A	18 May 2018	CN	107925762	A	17 April 2018	
			CN	107925762	B	27 November 2020	
			EP	3342164	A1	04 July 2018	
			EP	3342164	A4	27 February 2019	
			EP	3342164	B1	15 April 2020	
			KR	10-2124714	B1	19 June 2020	
			US	11196992	B2	07 December 2021	
			US	11589041	B2	21 February 2023	
			US	2022-0078418	A1	10 March 2022	
			WO	2017-036370	A1	09 March 2017	

KR 10-2022-0136163	A	07 October 2022	CN	117099372	A	21 November 2023	
			JP	2024-511636	A	14 March 2024	
			US	2024-0031580	A1	25 January 2024	
			WO	2022-211375	A1	06 October 2022	

KR 10-2023-0026944	A	27 February 2023	CN	117837149	A	05 April 2024	
			US	2024-0179324	A1	30 May 2024	
			WO	2023-022376	A1	23 February 2023	
