Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2876420 C 2015/10/06

(11)(21) 2 876 420

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(22) Date de dépôt/Filing Date: 2010/08/13

(41) Mise à la disp. pub./Open to Public Insp.: 2011/02/17

(45) Date de délivrance/Issue Date: 2015/10/06

(62) Demande originale/Original Application: 2 770 991

(30) Priorité/Priority: 2009/08/14 (KR10-2009-0075335)

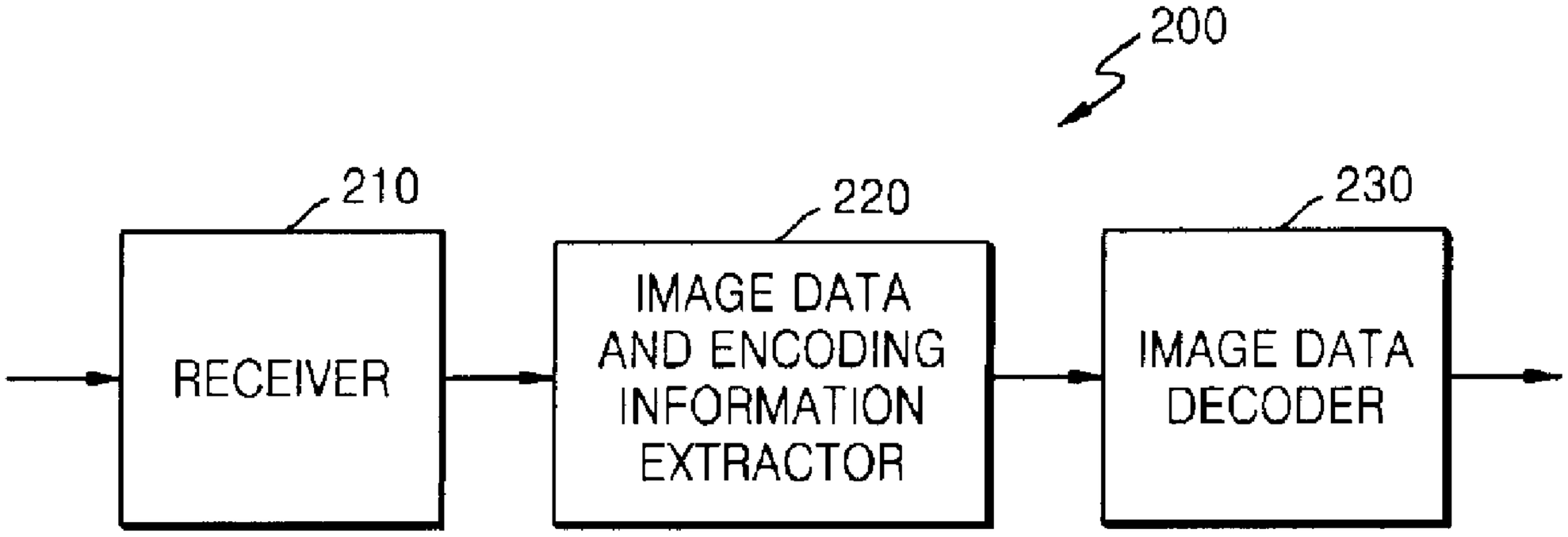
(51) Cl.Int./Int.Cl. *H04N 19/34* (2014.01), *H04N 19/184* (2014.01), *H04N 19/44* (2014.01)

(72) Inventeurs/Inventors:
CHEN, JIANLE, KR;
CHEON, MIN-SU, KR;
LEE, JAE-CHOOL, KR;
MIN, JUNG-HYE, KR;
JUNG, HAE-KYUNG, KR;
KIM, IL-KOO, KR;
LEE, SANG-RAE, KR;
LEE, KYO-HYUK, KR

(73) Propriétaire/Owner:
SAMSUNG ELECTRONICS CO., LTD., KR

(74) Agent: RIDOUT & MAYBEE LLP

(54) Titre : CODAGE ET DECODAGE VIDEO A L'AIDE D'UNITES DE CODAGE HIERARCHIQUES
(54) Title: ENCODING AND DECODING VIDEO USING HIERARCHICALLY-DIVIDED CODING UNITS



(57) Abrégé/Abstract:
Disclosed is a method of encoding a video, the method including; splitting a current picture into at least one maximum coding unit; determining a coded depth to output a final encoding result according to at least one split region obtained by splitting a region of the maximum coding unit according to depths, by encoding the at least one split region, based on a depth that deepens in proportion to the number of times the region of the maximum coding unit is split; and outputting image data constituting the final encoding result according to the at least one split region, and encoding information about the coded depth and a prediction mode, according to the at least one maximum coding unit.

Abstract

Disclosed is a method of encoding a video, the method including: splitting a current picture into at least one maximum coding unit; determining a coded depth to output a final encoding result according to at least one split region obtained by splitting a region of the maximum coding unit according to depths, by encoding the at least one split region, based on a depth that deepens in proportion to the number of times the region of the maximum coding unit is split; and outputting image data constituting the final encoding result according to the at least one split region, and encoding information about the coded depth and a prediction mode, according to the at least one maximum coding unit.

Description

ENCODING AND DECODING VIDEO USING HIERARCHICALLY-DIVIDED CODING UNITS

Technical Field

[1] The exemplary embodiments relate to encoding and decoding a video.

Background Art

[2] As hardware for reproducing and storing high resolution or high quality video content is being developed and supplied, a need for a video codec for effectively encoding or decoding the high resolution or high quality video content is increasing. In a conventional video codec, a video is encoded according to a limited encoding method based on a macroblock having a predetermined size.

Disclosure of Invention

Technical Problem

[3] The exemplary embodiments provide encoding and decoding of an image, based on a hierarchical encoding unit in various sizes.

Solution to Problem

[4] According to an aspect of an exemplary embodiment, there is provided a method of encoding a video, the method including: splitting a current picture into at least one maximum coding unit; determining a coded depth to output a final encoding result according to at least one split region obtained by splitting a region of each of the at least one maximum coding unit according to depths, by encoding the at least one split region, based on a depth that deepens in proportion to the number of times the region of the each maximum coding unit is split; and outputting image data constituting the final encoding result according to the at least one split region, and encoding information about the coded depth and a prediction mode, according to the each maximum coding unit.

Advantageous Effects of Invention

[5] When image data has high resolution and a large amount of data, the image data may be efficiently decoded and restored by using a size of a coding unit and an encoding mode, which are adaptively determined according to characteristics of the image data, by using information about an optimum encoding mode received from an encoder.

Brief Description of Drawings

[6] The above and other features of the exemplary embodiment will become more apparent by describing in detail exemplary embodiments thereof with reference to the

attached drawings in which:

[7] FIG. 1 is a block diagram of an apparatus for encoding a video, according to an exemplary embodiment;

[8] FIG. 2 is a block diagram of an apparatus for decoding a video, according to an exemplary embodiment;

[9] FIG. 3 is a diagram for describing a concept of coding units according to an exemplary embodiment;

[10] FIG. 4 is a block diagram of an image encoder based on coding units according to an exemplary embodiment;

[11] FIG. 5 is a block diagram of an image decoder based on coding units according to an exemplary embodiment;

[12] FIG. 6 is a diagram illustrating deeper coding units according to depths, and partitions according to an exemplary embodiment;

[13] FIG. 7 is a diagram for describing a relationship between a coding unit and transformation units, according to an exemplary embodiment;

[14] FIG. 8 is a diagram for describing encoding information of coding units corresponding to a coded depth, according to an exemplary embodiment;

[15] FIG. 9 is a diagram of deeper coding units according to depths, according to an exemplary embodiment;

[16] FIGS. 10 through 12 are diagrams for describing a relationship between coding units, prediction units, and transformation units, according to an exemplary embodiment;

[17] FIG. 13 is a diagram for describing a relationship between a coding unit, a prediction unit or a partition, and a transformation unit, according to encoding mode information of Table 1;

[18] FIG. 14 is a flowchart illustrating a method of encoding a video, according to an exemplary embodiment; and

[19] FIG. 15 is a flowchart illustrating a method of decoding a video, according to an exemplary embodiment.

Best Mode for Carrying out the Invention

[20] According to an aspect of an exemplary embodiment, there is provided a method of encoding a video, the method including: splitting a current picture into at least one maximum coding unit; determining a coded depth to output a final encoding result according to at least one split region obtained by splitting a region of each of the at least one maximum coding unit according to depths, by encoding the at least one split region, based on a depth that deepens in proportion to the number of times the region of the each maximum coding unit is split; and outputting image data constituting the final encoding result according to the at least one split region, and encoding in-

formation about the coded depth and a prediction mode, according to the each maximum coding unit.

[21] The coding unit may be characterized by a maximum size and a depth.

[22] The depth denotes the number of times a coding unit is hierarchically split, and as the depth deepens, deeper coding units according to depths may be split from the maximum coding unit to obtain minimum coding units. The depth is deepened from an upper depth to a lower depth. As the depth deepens, the number of times the maximum coding unit is split increases, and a total number of possible times the maximum coding unit is split corresponds to a maximum depth. The maximum size and the maximum depth of the coding unit may be predetermined.

[23] The method may further include predetermining a maximum depth indicating a total number of times a height and a width of the at least one maximum coding unit are hierarchically split, and a maximum size of the at least one maximum coding unit.

[24] The image data may be encoded based on the deeper coding units according to at least one depth, for each of the at least one maximum coding unit, and as the depth deepens, coding units may be hierarchically split from the each maximum coding unit.

[25] The determining of the coded depth may include performing prediction encoding according to coding units corresponding to a current depth and partitions obtained by splitting at least one of a height and a width of the coding units corresponding to the current depth, according to the deeper coding units according to depths for the each maximum coding unit.

[26] The determining of the coded depth may include performing transformation on a data unit having a smaller height or width than coding units corresponding to the current depth, according to the deeper coding units according to depths for the maximum coding unit.

[27] The determining of the coded depth may further include determining the coded depth having the minimum encoding error by comparing results of encoding according to depths for the each maximum coding unit, and determining at least one of a partition type and a prediction mode of a data unit that performed the prediction encoding in each coding unit corresponding to the coded depth, and a size of a data unit that performed transformation in the coding units corresponding to the coded depth.

[28] The prediction mode may be independently determined according to coding units corresponding to the coded depth, and include at least one of an intra mode, an inter mode, and a skip mode.

[29] The information about the encoding mode may include at least one of the coded depth for the each maximum coding unit, information about a partition type of a data unit that performed prediction encoding according to coding units corresponding to the coded depth, information about a prediction mode according to partitions of the coding

units corresponding to the coded depth, and information about a size of a data unit that performed transformation according to the coding units corresponding to the coded depth.

[30] The method may further include encoding split information indicating whether encoding is performed on coding units of a lower depth instead of a current depth, according to depths in the each maximum coding unit. The method may further include, if the encoding is performed on the coding units of the lower depth based on the split information of the current depth, repeatedly performing the encoding on each partial coding unit obtained by splitting the coding unit corresponding to the current depth.

[31] A depth of the maximum coding unit may be an uppermost depth, and deeper coding units according to depths constituting a split region according to depths may be coding units obtained by dividing a height and width of a coding unit of an upper depth by two. The partitions corresponding to the current depth may be data units obtained by dividing at least one of a height and a width of a data unit corresponding to the current depth by two.

[32] The encoding error may be measured by using a Rate-Distortion Optimization method based on Lagrangian multipliers.

[33] Prediction encoding in the intra mode and prediction encoding in the inter mode may be performed according to data units obtained by dividing the coding units corresponding to the current depth or at least one of the height and the width of the coding units corresponding to the current depth by two.

[34] According to another aspect of an exemplary embodiment, there is provided a method of decoding a video, the method including: receiving and parsing a bitstream of an encoded video; extracting image data of a current picture assigned to at least one maximum coding unit, and information about a coded depth and an encoding mode according to the at leas one maximum coding unit, from the bitstream; and decoding the image data in each of the at least one maximum coding unit based on the information about the coded depth and the encoding mode for the each maximum coding unit, wherein the depth deepens in proportion to the number of times a region of the each maximum coding unit is split, and a depth to which a final encoding result is output according to at least one split region, which is obtained by splitting a region of the each maximum coding unit according to depths, by encoding the at least one split region of the each maximum coding unit while encoding the current picture, is determined as at least one coded depth for the each maximum coding unit.

[35] The method may further include extracting information about a maximum depth indicating a total number of times a height and a width of the each maximum coding unit are hierarchically split, and information about a maximum size of the each maximum

coding unit from the bitstream.

[36] The decoding may include performing prediction encoding on the coding units corresponding to the coded depth for the each maximum coding unit, based on the information about the coded depth and information about a partition type and a prediction mode of a data unit to perform the prediction decoding, from among the extracted information about the encoding mode for the each maximum coding unit.

[37] The decoding may include performing inverse transformation on the each coding unit corresponding to the coded depth for the each maximum coding unit, based on information about the coded depth and a size of a data unit to perform the inverse transformation for the each maximum coding unit.

[38] The method may further include extracting split information indicating whether decoding is performed on coding units corresponding to a lower depth instead of a current depth according to the depths for the each maximum coding unit, from the bitstream.

[39] According to another aspect of an exemplary embodiment, there is provided an apparatus for encoding a video, the apparatus including: a maximum coding unit splitter for splitting a current picture into at least one maximum coding unit; an coding unit determiner for determining a coded depth to output a final encoding result according to at least one split region obtained by splitting a region of each of the at least one maximum coding unit according to depths, by encoding the at least one split region, based on a depth that deepens in proportion to the number of times the region of the each maximum coding unit is split; and an output unit for outputting image data constituting the final encoding result according to the at least one split region, and encoding information about the coded depth and a prediction mode, according to the each maximum coding unit.

[40] According to another aspect of an exemplary embodiment, there is provided an apparatus for decoding a video, the apparatus including: a receiver for receiving and parsing a bitstream of an encoded video; an image data and encoding information extractor for extracting image data of a current picture assigned to at least one maximum coding unit, and information about an coded depth and an encoding mode according to the at least one maximum coding unit, from the bitstream; and a decoder for decoding the image data in each maximum coding unit based on the information about the coded depth and the encoding mode for each of the at least one maximum coding unit, wherein the depth deepens in proportion to the number of times a region of the each maximum coding unit is split, and a depth to which a final encoding result is output according to at least one split region, which is obtained by splitting a region of the each maximum coding unit according to depths, by encoding the at least one split region of the each maximum coding unit while encoding the current picture, is de-

termined as at least one coded depth for the each maximum coding unit.

[41] According to another aspect of an exemplary embodiment, there is provided a computer readable recording medium having recorded thereon a program for executing the method of encoding a video.

[42] According to another aspect of an exemplary embodiment, there is provided a computer readable recording medium having recorded thereon a program for executing the method of decoding a video.

Mode for the Invention

[43] Hereinafter, the exemplary embodiments will be described more fully with reference to the accompanying drawings, in which exemplary embodiments are shown. In the exemplary embodiments, "unit" may or may not refer to a unit of size, depending on its context.

[44] Hereinafter, a 'coding unit' is an encoding data unit in which the image data is encoded at an encoder side and an encoded data unit in which the encoded image data is decoded at a decoder side, according to exemplary embodiments. Also, a 'coded depth' means a depth where a coding unit is encoded.

[45] Hereinafter, an 'image' may denote a still image for a video or a moving image, that is, the video itself.

[46] FIG. 1 is a block diagram of a video encoding apparatus 100, according to an exemplary embodiment.

[47] The video encoding apparatus 100 includes a maximum coding unit splitter 110, a coding unit determiner 120, and an output unit 130.

[48] The maximum coding unit splitter 110 may split a current picture based on a maximum coding unit for the current picture of an image. If the current picture is larger than the maximum coding unit, image data of the current picture may be split into the at least one maximum coding unit. The maximum coding unit according to an exemplary embodiment may be a data unit having a size of 32x32, 64x64, 128x128, 256x256, etc., wherein a shape of the data unit is a square having a width and height in squares of 2. The image data may be output to the coding unit determiner 120 according to the at least one maximum coding unit.

[49] A coding unit according to an exemplary embodiment may be characterized by a maximum size and a depth. The depth denotes a number of times the coding unit is spatially split from the maximum coding unit, and as the depth deepens or increases, deeper encoding units according to depths may be split from the maximum coding unit to a minimum coding unit. A depth of the maximum coding unit is an uppermost depth and a depth of the minimum coding unit is a lowermost depth. Since a size of a coding unit corresponding to each depth decreases as the depth of the maximum coding unit

deepens, a coding unit corresponding to an upper depth may include a plurality of coding units corresponding to lower depths.

[50] As described above, the image data of the current picture is split into the maximum coding units according to a maximum size of the coding unit, and each of the maximum coding units may include deeper coding units that are split according to depths. Since the maximum coding unit according to an exemplary embodiment is split according to depths, the image data of a spatial domain included in the maximum coding unit may be hierarchically classified according to depths.

[51] A maximum depth and a maximum size of a coding unit, which limit the total number of times a height and a width of the maximum coding unit are hierarchically split may be predetermined.

[52] The coding unit determiner 120 encodes at least one split region obtained by splitting a region of the maximum coding unit according to depths, and determines a depth to output a finally encoded image data according to the at least one split region. In other words, the coding unit determiner 120 determines a coded depth by encoding the image data in the deeper coding units according to depths, according to the maximum coding unit of the current picture, and selecting a depth having the least encoding error. Thus, the encoded image data of the coding unit corresponding to the determined coded depth is finally output. Also, the coding units corresponding to the coded depth may be regarded as encoded coding units.

[53] The determined coded depth and the encoded image data according to the determined coded depth are output to the output unit 130.

[54] The image data in the maximum coding unit is encoded based on the deeper coding units corresponding to at least one depth equal to or below the maximum depth, and results of encoding the image data are compared based on each of the deeper coding units. A depth having the least encoding error may be selected after comparing encoding errors of the deeper coding units. At least one coded depth may be selected for each maximum coding unit.

[55] The size of the maximum coding unit is split as a coding unit is hierarchically split according to depths, and as the number of coding units increases. Also, even if coding units correspond to same depth in one maximum coding unit, it is determined whether to split each of the coding units corresponding to the same depth to a lower depth by measuring an encoding error of the image data of the each coding unit, separately. Accordingly, even when image data is included in one maximum coding unit, the image data is split to regions according to the depths and the encoding errors may differ according to regions in the one maximum coding unit, and thus the coded depths may differ according to regions in the image data. Thus, one or more coded depths may be determined in one maximum coding unit, and the image data of the maximum coding

unit may be divided according to coding units of at least one coded depth.

[56] Accordingly, the coding unit determiner 120 may determine coding units having a tree structure included in the maximum coding unit. The 'coding units having a tree structure' according to an exemplary embodiment include coding units corresponding to a depth determined to be the coded depth, from among all deeper coding units included in the maximum coding unit. A coding unit of a coded depth may be hierarchically determined according to depths in the same region of the maximum coding unit, and may be independently determined in different regions. Similarly, a coded depth in a current region may be independently determined from a coded depth in another region.

[57] A maximum depth according to an exemplary embodiment is an index related to the number of splitting times from a maximum coding unit to a minimum coding unit. A first maximum depth according to an exemplary embodiment may denote the total number of splitting times from the maximum coding unit to the minimum coding unit. A second maximum depth according to an exemplary embodiment may denote the total number of depth levels from the maximum coding unit to the minimum coding unit. For example, when a depth of the maximum coding unit is 0, a depth of a coding unit, in which the maximum coding unit is split once, may be set to 1, and a depth of a coding unit, in which the maximum coding unit is split twice, may be set to 2. Here, if the minimum coding unit is a coding unit in which the maximum coding unit is split four times, 5 depth levels of depths 0, 1, 2, 3 and 4 exist, and thus the first maximum depth may be set to 4, and the second maximum depth may be set to 5.

[58] Prediction encoding and transformation may be performed according to the maximum coding unit. The prediction encoding and the transformation are also performed based on the deeper coding units according to a depth equal to or depths less than the maximum depth, according to the maximum coding unit. Transformation may be performed according to method of orthogonal transformation or integer transformation.

[59] Since the number of deeper coding units increases whenever the maximum coding unit is split according to depths, encoding including the prediction encoding and the transformation is performed on all of the deeper coding units generated as the depth deepens. For convenience of description, the prediction encoding and the transformation will now be described based on a coding unit of a current depth, in a maximum coding unit.

[60] The video encoding apparatus 100 may variably select a size or shape of a data unit for encoding the image data. In order to encode the image data, operations, such as prediction encoding, transformation, and entropy encoding, are performed, and at this time, the same data unit may be used for all operations or different data units may be

used for each operation.

[61] For example, the video encoding apparatus 100 may select not only a coding unit for encoding the image data, but also a data unit different from the coding unit so as to perform the prediction encoding on the image data in the coding unit.

[62] In order to perform prediction encoding in the maximum coding unit, the prediction encoding may be performed based on a coding unit corresponding to a coded depth, i.e., based on a coding unit that is no longer split to coding units corresponding to a lower depth. Hereinafter, the coding unit that is no longer split and becomes a basis unit for prediction encoding will now be referred to as a 'prediction unit'. A partition obtained by splitting the prediction unit may include a prediction unit or a data unit obtained by splitting at least one of a height and a width of the prediction unit.

[63] For example, when a coding unit of 2Nx2N (where N is a positive integer) is no longer split and becomes a prediction unit of 2Nx2N, and a size of a partition may be 2Nx2N, 2NxN, Nx2N, or NxN. Examples of a partition type include symmetrical partitions that are obtained by symmetrically splitting a height or width of the prediction unit, partitions obtained by asymmetrically splitting the height or width of the prediction unit, such as 1:n or n:1, partitions that are obtained by geometrically splitting the prediction unit, and partitions having arbitrary shapes.

[64] A prediction mode of the prediction unit may be at least one of an intra mode, a inter mode, and a skip mode. For example, the intra mode or the inter mode may be performed on the partition of 2Nx2N, 2NxN, Nx2N, or NxN. Also, the skip mode may be performed only on the partition of 2Nx2N. The encoding is independently performed on one prediction unit in a coding unit, thereby selecting a prediction mode having a least encoding error.

[65] The video encoding apparatus 100 may also perform the transformation on the image data in a coding unit based not only on the coding unit for encoding the image data, but also based on a data unit that is different from the coding unit.

[66] In order to perform the transformation in the coding unit, the transformation may be performed based on a data unit having a size smaller than or equal to the coding unit. For example, the data unit for the transformation may include a data unit for an intra mode and a data unit for an inter mode.

[67] A data unit used as a base of the transformation will now be referred to as a 'transformation unit'. A transformation depth indicating the number of splitting times to reach the transformation unit by splitting the height and width of the coding unit may also be set in the transformation unit. For example, in a current coding unit of 2Nx2N, a transformation depth may be 0 when the size of a transformation unit is also 2Nx2N, may be 1 when each of the height and width of the current coding unit is split into two equal parts, totally split into 4^1 transformation units, and the size of the trans-

formation unit is thus NxN, and may be 2 when each of the height and width of the current coding unit is split into four equal parts, totally split into 4^2 transformation units and the size of the transformation unit is thus N/2xN/2. For example, the transformation unit may be set according to a hierarchical tree structure, in which a transformation unit of an upper transformation depth is split into four transformation units of a lower transformation depth according to the hierarchical characteristics of a transformation depth.

[68] Similarly to the coding unit, the transformation unit in the coding unit may be recursively split into smaller sized regions, so that the transformation unit may be determined independently in units of regions. Thus, residual data in the coding unit may be divided according to the transformation having the tree structure according to transformation depths.

[69] Encoding information according to coding units corresponding to a coded depth requires not only information about the coded depth, but also about information related to prediction encoding and transformation. Accordingly, the coding unit determiner 120 not only determines a coded depth having a least encoding error, but also determines a partition type in a prediction unit, a prediction mode according to prediction units, and a size of a transformation unit for transformation.

[70] Coding units according to a tree structure in a maximum coding unit and a method of determining a partition, according to exemplary embodiments, will be described in detail later with reference to FIGS. 3 through 12.

[71] The coding unit determiner 120 may measure an encoding error of deeper coding units according to depths by using Rate-Distortion Optimization based on Lagrangian multipliers.

[72] The output unit 130 outputs the image data of the maximum coding unit, which is encoded based on the at least one coded depth determined by the coding unit determiner 120, and information about the encoding mode according to the coded depth, in bitstreams.

[73] The encoded image data may be obtained by encoding residual data of an image.

[74] The information about the encoding mode according to coded depth may include information about the coded depth, about the partition type in the prediction unit, the prediction mode, and the size of the transformation unit.

[75] The information about the coded depth may be defined by using split information according to depths, which indicates whether encoding is performed on coding units of a lower depth instead of a current depth. If the current depth of the current coding unit is the coded depth, image data in the current coding unit is encoded and output, and thus the split information may be defined not to split the current coding unit to a lower depth. Alternatively, if the current depth of the current coding unit is not the coded

depth, the encoding is performed on the coding unit of the lower depth, and thus the split information may be defined to split the current coding unit to obtain the coding units of the lower depth.

[76] If the current depth is not the coded depth, encoding is performed on the coding unit that is split into the coding unit of the lower depth. Since at least one coding unit of the lower depth exists in one coding unit of the current depth, the encoding is repeatedly performed on each coding unit of the lower depth, and thus the encoding may be re-cursively performed for the coding units having the same depth.

[77] Since the coding units having a tree structure are determined for one maximum coding unit, and information about at least one encoding mode is determined for a coding unit of a coded depth, information about at least one encoding mode may be de-termined for one maximum coding unit. Also, a coded depth of the image data of the maximum coding unit may be different according to locations since the image data is hierarchically split according to depths, and thus information about the coded depth and the encoding mode may be set for the image data.

[78] Accordingly, the output unit 130 may assign encoding information about a corre-sponding coded depth and an encoding mode to at least one of the coding unit, the prediction unit, and a minimum unit included in the maximum coding unit.

[79] The minimum unit according to an exemplary embodiment is a rectangular data unit obtained by splitting the minimum coding unit constituting the lowermost depth by 4. Alternatively, the minimum unit may be a maximum rectangular data unit that may be included in all of the coding units, prediction units, partition units, and transformation units included in the maximum coding unit.

[80] For example, the encoding information output through the output unit 130 may be classified into encoding information according to coding units, and encoding in-formation according to prediction units. The encoding information according to the coding units may include the information about the prediction mode and about the size of the partitions. The encoding information according to the prediction units may include information about an estimated direction of an inter mode, about a reference image index of the inter mode, about a motion vector, about a chroma component of an intra mode, and about an interpolation method of the intra mode. Also, information about a maximum size of the coding unit defined according to pictures, slices, or GOPs, and information about a maximum depth may be inserted into SPS (Sequence Parameter Set) or a header of a bitstream.

[81] In the video encoding apparatus 100, the deeper coding unit may be a coding unit obtained by dividing a height or width of a coding unit of an upper depth, which is one layer above, by two. In other words, when the size of the coding unit of the current depth is 2Nx2N, the size of the coding unit of the lower depth is NxN. Also, the coding

unit of the current depth having the size of 2Nx2N may include a maximum of 4 coding units of the lower depth.

[82] Accordingly, the video encoding apparatus 100 may form the coding units having the tree structure by determining coding units having an optimum shape and an optimum size for each maximum coding unit, based on the size of the maximum coding unit and the maximum depth determined considering characteristics of the current picture. Also, since encoding may be performed on each maximum coding unit by using any one of various prediction modes and transformations, an optimum encoding mode may be determined considering characteristics of the coding unit of various image sizes.

[83] Thus, if an image having high resolution or large data amount is encoded in a conventional macroblock, a number of macroblocks per picture excessively increases. Accordingly, a number of pieces of compressed information generated for each macroblock increases, and thus it is difficult to transmit the compressed information and data compression efficiency decreases. However, by using the video encoding apparatus 100, image compression efficiency may be increased since a coding unit is adjusted while considering characteristics of an image while increasing a maximum size of a coding unit while considering a size of the image.

[84] FIG. 2 is a block diagram of a video decoding apparatus 200, according to an exemplary embodiment.

[85] The video decoding apparatus 200 includes a receiver 210, an image data and encoding information extractor 220, and an image data decoder 230. Definitions of various terms, such as a coding unit, a depth, a prediction unit, a transformation unit, and information about various encoding modes, for various operations of the video decoding apparatus 200 are identical to those described with reference to FIG. 1 and the video encoding apparatus 100.

[86] The receiver 210 receives and parses a bitstream of an encoded video. The image data and encoding information extractor 220 extracts encoded image data for each coding unit from the parsed bitstream, wherein the coding units have a tree structure according to each maximum coding unit, and outputs the extracted image data to the image data decoder 230. The image data and encoding information extractor 220 may extract information about a maximum size of a coding unit of a current picture, from a header about the current picture or SPS.

[87] Also, the image data and encoding information extractor 220 extracts information about a coded depth and an encoding mode for the coding units having a tree structure according to each maximum coding unit, from the parsed bitstream. The extracted information about the coded depth and the encoding mode is output to the image data decoder 230. In other words, the image data in a bit stream is split into the maximum coding unit so that the image data decoder 230 decodes the image data for each

maximum coding unit.

[88] The information about the coded depth and the encoding mode according to the maximum coding unit may be set for information about at least one coding unit corresponding to the coded depth, and information about an encoding mode may include information about a partition type of a corresponding coding unit corresponding to the coded depth, about a prediction mode, and a size of a transformation unit. Also, splitting information according to depths may be extracted as the information about the coded depth.

[89] The information about the coded depth and the encoding mode according to each maximum coding unit extracted by the image data and encoding information extractor 220 is information about a coded depth and an encoding mode determined to generate a minimum encoding error when an encoder, such as the video encoding apparatus 100, repeatedly performs encoding for each deeper coding unit according to depths according to each maximum coding unit. Accordingly, the video decoding apparatus 200 may restore an image by decoding the image data according to a coded depth and an encoding mode that generates the minimum encoding error.

[90] Since encoding information about the coded depth and the encoding mode may be assigned to a predetermined data unit from among a corresponding coding unit, a prediction unit, and a minimum unit, the image data and encoding information extractor 220 may extract the information about the coded depth and the encoding mode according to the predetermined data units. The predetermined data units to which the same information about the coded depth and the encoding mode is assigned may be inferred to be the data units included in the same maximum coding unit.

[91] The image data decoder 230 restores the current picture by decoding the image data in each maximum coding unit based on the information about the coded depth and the encoding mode according to the maximum coding units. In other words, the image data decoder 230 may decode the encoded image data based on the extracted information about the partition type, the prediction mode, and the transformation unit for each coding unit from among the coding units having the tree structure included in each maximum coding unit. A decoding process may include a prediction including intra prediction and motion compensation, and an inverse transformation. Inverse transformation may be performed according to method of inverse orthogonal transformation or inverse integer transformation.

[92] The image data decoder 230 may perform intra prediction or motion compensation according to a partition and a prediction mode of each coding unit, based on the information about the partition type and the prediction mode of the prediction unit of the coding unit according to coded depths.

[93] Also, the image data decoder 230 may perform inverse transformation according to

each transformation unit in the coding unit, based on the information about the size of the transformation unit of the coding unit according to coded depths, so as to perform the inverse transformation according to maximum coding units.

[94] The image data decoder 230 may determine at least one coded depth of a current maximum coding unit by using split information according to depths. If the split information indicates that image data is no longer split in the current depth, the current depth is a coded depth. Accordingly, the image data decoder 230 may decode encoded data of at least one coding unit corresponding to the each coded depth in the current maximum coding unit by using the information about the partition type of the prediction unit, the prediction mode, and the size of the transformation unit for each coding unit corresponding to the coded depth, and output the image data of the current maximum coding unit.

[95] In other words, data units containing the encoding information including the same split information may be gathered by observing the encoding information set assigned for the predetermined data unit from among the coding unit, the prediction unit, and the minimum unit, and the gathered data units may be considered to be one data unit to be decoded by the image data decoder 230 in the same encoding mode.

[96] The video decoding apparatus 200 may obtain information about at least one coding unit that generates the minimum encoding error when encoding is recursively performed for each maximum coding unit, and may use the information to decode the current picture. In other words, the coding units having the tree structure determined to be the optimum coding units in each maximum coding unit may be decoded. Also, the maximum size of coding unit is determined considering resolution and an amount of image data.

[97] Accordingly, even if image data has high resolution and a large amount of data, the image data may be efficiently decoded and restored by using a size of a coding unit and an encoding mode, which are adaptively determined according to characteristics of the image data, by using information about an optimum encoding mode received from an encoder.

[98] A method of determining coding units having a tree structure, a prediction unit, and a transformation unit, according to an exemplary embodiment, will now be described with reference to FIGS. 3 through 13.

[99] FIG. 3 is a diagram for describing a concept of coding units according to an exemplary embodiment.

[100] A size of a coding unit may be expressed in width x height, and may be 64x64, 32x32, 16x16, and 8x8. A coding unit of 64x64 may be split into partitions of 64x64, 64x32, 32x64, or 32x32, and a coding unit of 32x32 may be split into partitions of 32x32, 32x16, 16x32, or 16x16, a coding unit of 16x16 may be split into partitions of

16x16, 16x8, 8x16, or 8x8, and a coding unit of 8x8 may be split into partitions of 8x8, 8x4, 4x8, or 4x4.

[101] In video data 310, a resolution is 1920x1080, a maximum size of a coding unit is 64, and a maximum depth is 2. In video data 320, a resolution is 1920x1080, a maximum size of a coding unit is 64, and a maximum depth is 3. In video data 330, a resolution is 352x288, a maximum size of a coding unit is 16, and a maximum depth is 1. The maximum depth shown in FIG. 3 denotes a total number of splits from a maximum coding unit to a minimum decoding unit.

[102] If a resolution is high or a data amount is large, a maximum size of a coding unit may be large so as to not only increase encoding efficiency but also to accurately reflect characteristics of an image. Accordingly, the maximum size of the coding unit of the video data 310 and 320 having the higher resolution than the video data 330 may be 64.

[103] Since the maximum depth of the video data 310 is 2, coding units 315 of the video data 310 may include a maximum coding unit having a long axis size of 64, and coding units having long axis sizes of 32 and 16 since depths are deepened to two layers by splitting the maximum coding unit twice. Meanwhile, since the maximum depth of the video data 330 is 1, coding units 335 of the video data 330 may include a maximum coding unit having a long axis size of 16, and coding units having a long axis size of 8 since depths are deepened to one layer by splitting the maximum coding unit once.

[104] Since the maximum depth of the video data 320 is 3, coding units 325 of the video data 320 may include a maximum coding unit having a long axis size of 64, and coding units having long axis sizes of 32, 16, and 8 since the depths are deepened to 3 layers by splitting the maximum coding unit three times. As a depth deepens, detailed information may be precisely expressed.

[105] FIG. 4 is a block diagram of an image encoder 400 based on coding units, according to an exemplary embodiment.

[106] The image encoder 400 performs operations of the coding unit determiner 120 of the video encoding apparatus 100 to encode image data. In other words, an intra predictor 410 performs intra prediction on coding units in an intra mode, from among a current frame 405, and a motion estimator 420 and a motion compensator 425 performs inter estimation and motion compensation on coding units in an inter mode from among the current frame 405 by using the current frame 405, and a reference frame 495.

[107] Data output from the intra predictor 410, the motion estimator 420, and the motion compensator 425 is output as a quantized transformation coefficient through a transformer 430 and a quantizer 440. The quantized transformation coefficient is restored as data in a spatial domain through an inverse quantizer 460 and an inverse transformer 470, and the restored data in the spatial domain is output as the reference

frame 495 after being post-processed through a deblocking unit 480 and a loop filtering unit 490. The quantized transformation coefficient may be output as a bitstream 455 through an entropy encoder 450.

[108] In order for the image encoder 400 to be applied in the video encoding apparatus 100, all elements of the image encoder 400, i.e., the intra predictor 410, the motion estimator 420, the motion compensator 425, the transformer 430, the quantizer 440, the entropy encoder 450, the inverse quantizer 460, the inverse transformer 470, the de-blocking unit 480, and the loop filtering unit 490 perform operations based on each coding unit from among coding units having a tree structure while considering the maximum depth of each maximum coding unit.

[109] Specifically, the intra predictor 410, the motion estimator 420, and the motion compensator 425 determines partitions and a prediction mode of each coding unit from among the coding units having a tree structure while considering the maximum size and the maximum depth of a current maximum coding unit, and the transformer 430 determines the size of the transformation unit in each coding unit from among the coding units having a tree structure.

[110] FIG. 5 is a block diagram of an image decoder 500 based on coding units, according to an exemplary embodiment.

[111] A parser 510 parses encoded image data to be decoded and information about encoding required for decoding from a bitstream 505. The encoded image data is output as inverse quantized data through an entropy decoder 520 and an inverse quantizer 530, and the inverse quantized data is restored to image data in a spatial domain through an inverse transformer 540.

[112] An intra predictor 550 performs intra prediction on coding units in an intra mode with respect to the image data in the spatial domain, and a motion compensator 560 performs motion compensation on coding units in an inter mode by using a reference frame 585.

[113] The image data in the spatial domain, which passed through the intra predictor 550 and the motion compensator 560, may be output as a restored frame 595 after being post-processed through a deblocking unit 570 and a loop filtering unit 580. Also, the image data that is post-processed through the deblocking unit 570 and the loop filtering unit 580 may be output as the reference frame 585.

[114] In order to decode the image data in the image data decoder 230 of the video decoding apparatus 200, the image decoder 500 may perform operations that are performed after the parser 510.

[115] In order for the image decoder 500 to be applied in the video decoding apparatus 200, all elements of the image decoder 500, i.e., the parser 510, the entropy decoder 520, the inverse quantizer 530, the inverse transformer 540, the intra predictor 550, the

motion compensator 560, the deblocking unit 570, and the loop filtering unit 580 perform operations based on coding units having a tree structure for each maximum coding unit.

[116] Specifically, the intra predictor 550 and the motion compensator 560 perform operations based on partitions and a prediction mode for each of the coding units having a tree structure, and the inverse transformer 540 perform operations based on a size of a transformation unit for each coding unit.

[117] FIG. 6 is a diagram illustrating deeper coding units according to depths, and partitions, according to an exemplary embodiment.

[118] The video encoding apparatus 100 and the video decoding apparatus 200 use hierarchical coding units so as to consider characteristics of an image. A maximum height, a maximum width, and a maximum depth of coding units may be adaptively determined according to the characteristics of the image, or may be differently set by a user. Sizes of deeper coding units according to depths may be determined according to the predetermined maximum size of the coding unit.

[119] In a hierarchical structure 600 of coding units, according to an exemplary embodiment, the maximum height and the maximum width of the coding units are each 64, and the maximum depth is 4. Since a depth deepens along a vertical axis of the hierarchical structure 600, a height and a width of the deeper coding unit are each split. Also, a prediction unit and partitions, which are bases for prediction encoding of each deeper coding unit, are shown along a horizontal axis of the hierarchical structure 600.

[120] In other words, a coding unit 610 is a maximum coding unit in the hierarchical structure 600, wherein a depth is 0 and a size, i.e., a height by width, is 64x64. The depth deepens along the vertical axis, and a coding unit 620 having a size of 32x32 and a depth of 1, a coding unit 630 having a size of 16x16 and a depth of 2, a coding unit 640 having a size of 8x8 and a depth of 3, and a coding unit 650 having a size of 4x4 and a depth of 4 exist. The coding unit 650 having the size of 4x4 and the depth of 4 is a minimum coding unit.

[121] The prediction unit and the partitions of a coding unit are arranged along the horizontal axis according to each depth. In other words, if the coding unit 610 having the size of 64x64 and the depth of 0 is a prediction unit, the prediction unit may be split into partitions include in the encoding unit 610, i.e. a partition 610 having a size of 64x64, partitions 612 having the size of 64x32, partitions 614 having the size of 32x64, or partitions 616 having the size of 32x32.

[122] Similarly, a prediction unit of the coding unit 620 having the size of 32x32 and the depth of 1 may be split into partitions included in the coding unit 620, i.e. a partition 620 having a size of 32x32, partitions 622 having a size of 32x16, partitions 624 having a size of 16x32, and partitions 626 having a size of 16x16.

[123] Similarly, a prediction unit of the coding unit 630 having the size of 16x16 and the depth of 2 may be split into partitions included in the coding unit 630, i.e., a partition having a size of 16x16 included in the coding unit 630, partitions 632 having a size of 16x8, partitions 634 having a size of 8x16, and partitions 636 having a size of 8x8.

[124] Similarly, a prediction unit of the coding unit 640 having the size of 8x8 and the depth of 3 may be split into partitions included in the coding unit 640, i.e. a partition having a size of 8x8 included in the coding unit 640, partitions 642 having a size of 8x4, partitions 644 having a size of 4x8, and partitions 646 having a size of 4x4.

[125] The coding unit 650 having the size of 4x4 and the depth of 4 is the minimum coding unit and a coding unit of the lowermost depth. A prediction unit of the coding unit 650 is only assigned to a partition having a size of 4x4.

[126] In order to determine the at least one coded depth of the coding units constituting the maximum coding unit 610, the coding unit determiner 120 of the video encoding apparatus 100 performs encoding for coding units corresponding to each depth included in the maximum coding unit 610.

[127] A number of deeper coding units according to depths including data in the same range and the same size increases as the depth deepens. For example, four coding units corresponding to a depth of 2 are required to cover data that is included in one coding unit corresponding to a depth of 1. Accordingly, in order to compare encoding results of the same data according to depths, the coding unit corresponding to the depth of 1 and four coding units corresponding to the depth of 2 are each encoded.

[128] In order to perform encoding for a current depth from among the depths, a least encoding error may be selected for the current depth by performing encoding for each prediction unit in the coding units corresponding to the current depth, along the horizontal axis of the hierarchical structure 600. Alternatively, the minimum encoding error may be searched for by comparing the least encoding errors according to depths, by performing encoding for each depth as the depth deepens along the vertical axis of the hierarchical structure 600. A depth and a partition having the minimum encoding error in the coding unit 610 may be selected as the coded depth and a partition type of the coding unit 610.

[129] FIG. 7 is a diagram for describing a relationship between a coding unit 710 and transformation units 720, according to an exemplary embodiment.

[130] The video encoding apparatus 100 or 200 encodes or decodes an image according to coding units having sizes smaller than or equal to a maximum coding unit for each maximum coding unit. Sizes of transformation units for transformation during encoding may be selected based on data units that are not larger than a corresponding coding unit.

[131] For example, in the video encoding apparatus 100 or 200, if a size of the coding unit

710 is 64x64, transformation may be performed by using the transformation units 720 having a size of 32x32.

[132] Also, data of the coding unit 710 having the size of 64x64 may be encoded by performing the transformation on each of the transformation units having the size of 32x32, 16x16, 8x8, and 4x4, which are smaller than 64x64, and then a transformation unit having the least coding error may be selected.

[133] FIG. 8 is a diagram for describing encoding information of coding units corresponding to a coded depth, according to an exemplary embodiment.

[134] The output unit 130 of the video encoding apparatus 100 may encode and transmit information 800 about a partition type, information 810 about a prediction mode, and information 820 about a size of a transformation unit for each coding unit corresponding to a coded depth, as information about an encoding mode.

[135] The information 800 indicates information about a shape of a partition obtained by splitting a prediction unit of a current coding unit, wherein the partition is a data unit for prediction encoding the current coding unit. For example, a current coding unit CU_0 having a size of 2Nx2N may be split into any one of a partition 802 having a size of 2Nx2N, a partition 804 having a size of 2NxN, a partition 806 having a size of Nx2N, and a partition 808 having a size of NxN. Here, the information 800 about a partition type is set to indicate one of the partition 804 having a size of 2NxN, the partition 806 having a size of Nx2N, and the partition 808 having a size of NxN

[136] The information 810 indicates a prediction mode of each partition. For example, the information 810 may indicate a mode of prediction encoding performed on a partition indicated by the information 800, i.e., an intra mode 812, an inter mode 814, or a skip mode 816.

[137] The information 820 indicates a transformation unit to be based on when transformation is performed on a current coding unit. For example, the transformation unit may be a first intra transformation unit 822, a second intra transformation unit 824, a first inter transformation unit 826, or a second intra transformation unit 828.

[138] The image data and encoding information extractor 220 of the video decoding apparatus 200 may extract and use the information 800, 810, and 820 for decoding, according to each deeper coding unit

[139] FIG. 9 is a diagram of deeper coding units according to depths, according to an exemplary embodiment.

[140] Split information may be used to indicate a change of a depth. The spilt information indicates whether a coding unit of a current depth is split into coding units of a lower depth.

[141] A prediction unit 910 for prediction encoding a coding unit 900 having a depth of 0 and a size of 2N_0x2N_0 may include partitions of a partition type 912 having a size

of 2N_0x2N_0, a partition type 914 having a size of 2N_0xN_0, a partition type 916 having a size of N_0x2N_0, and a partition type 918 having a size of N_0xN_0. FIG. 9 only illustrates the partition types 912 through 918 which are obtained by symmetrically splitting the prediction unit 910, but a partition type is not limited thereto, and the partitions of the prediction unit 910 may include asymmetrical partitions, partitions having a predetermined shape, and partitions having a geometrical shape.

[142] Prediction encoding is repeatedly performed on one partition having a size of 2N_0x2N_0, two partitions having a size of 2N_0xN_0, two partitions having a size of N_0x2N_0, and four partitions having a size of N_0xN_0, according to each partition type. The prediction encoding in an intra mode and an inter mode may be performed on the partitions having the sizes of 2N_0x2N_0, N_0x2N_0, 2N_0xN_0, and N_0xN_0. The prediction encoding in a skip mode is performed only on the partition having the size of 2N_0x2N_0.

[143] Errors of encoding including the prediction encoding in the partition types 912 through 918 are compared, and the least encoding error is determined among the partition types. If an encoding error is smallest in one of the partition types 912 through 916, the prediction unit 910 may not be split into a lower depth.

[144] If the encoding error is the smallest in the partition type 918, a depth is changed from 0 to 1 to split the partition type 918 in operation 920, and encoding is repeatedly performed on coding units 930 having a depth of 2 and a size of N_0xN_0 to search for a minimum encoding error.

[145] A prediction unit 940 for prediction encoding the coding unit 930 having a depth of 1 and a size of 2N_1x2N_1 (=N_0xN_0) may include partitions of a partition type 942 having a size of 2N_1x2N_1, a partition type 944 having a size of 2N_1xN_1, a partition type 946 having a size of N_1x2N_1, and a partition type 948 having a size of N_1xN_1.

[146] If an encoding error is the smallest in the partition type 948, a depth is changed from 1 to 2 to split the partition type 948 in operation 950, and encoding is repeatedly performed on coding units 960, which have a depth of 2 and a size of N_2xN_2 to search for a minimum encoding error.

[147] When a maximum depth is d, split operation according to each depth may be performed up to when a depth becomes d-1, and split information may be encoded as up to when a depth is one of 0 to d-2. In other words, when encoding is performed up to when the depth is d-1 after a coding unit corresponding to a depth of d-2 is split in operation 970, a prediction unit 990 for prediction encoding a coding unit 980 having a depth of d-1 and a size of 2N_(d-1)x2N_(d-1) may include partitions of a partition type 992 having a size of 2N_(d-1)x2N_(d-1), a partition type 994 having a size of 2N_(d-1)xN_(d-1), a partition type 996 having a size of N_(d-1)x2N_(d-1), and a

partition type 998 having a size of N_(d-1)xN_(d-1).

[148] Prediction encoding may be repeatedly performed on one partition having a size of 2N_(d-1)x2N_(d-1), two partitions having a size of 2N_(d-1)xN_(d-1), two partitions having a size of N_(d-1)x2N_(d-1), four partitions having a size of N_(d-1)xN_(d-1) from among the partition types 992 through 998 to search for a partition type having a minimum encoding error.

[149] Even when the partition type 998 has the minimum encoding error, since a maximum depth is d, a coding unit CU_(d-1) having a depth of d-1 is no longer split to a lower depth, and a coded depth for the coding units constituting a current maximum coding unit 900 is determined to be d-1 and a partition type of the current maximum coding unit 900 may be determined to be N_(d-1)xN_(d-1). Also, since the maximum depth is d and a minimum coding unit 980 having a lowermost depth of d-1 is no longer split to a lower depth, split information for the minimum coding unit 980 is not set.

[150] A data unit 999 may be a 'minimum unit' for the current maximum coding unit. A minimum unit according to an exemplary embodiment may be a rectangular data unit obtained by splitting a minimum coding unit 980 by 4. By performing the encoding repeatedly, the video encoding apparatus 100 may select a depth having the least encoding error by comparing encoding errors according to depths of the coding unit 900 to determine a coded depth, and set a corresponding partition type and a prediction mode as an encoding mode of the coded depth.

[151] As such, the minimum encoding errors according to depths are compared in all of the depths of 1 through d, and a depth having the least encoding error may be determined as a coded depth. The coded depth, the partition type of the prediction unit, and the prediction mode may be encoded and transmitted as information about an encoding mode. Also, since a coding unit is split from a depth of 0 to a coded depth, only split information of the coded depth is set to 0, and split information of depths excluding the coded depth is set to 1.

[152] The image data and encoding information extractor 220 of the video decoding apparatus 200 may extract and use the information about the coded depth and the prediction unit of the coding unit 900 to decode the partition 912. The video decoding apparatus 200 may determine a depth, in which split information is 0, as a coded depth by using split information according to depths, and use information about an encoding mode of the corresponding depth for decoding.

[153] FIGS. 10 through 12 are diagrams for describing a relationship between coding units 1010, prediction units 1060, and transformation units 1070, according to an exemplary embodiment.

[154] The coding units 1010 are coding units having a tree structure, corresponding to coded depths determined by the video encoding apparatus 100, in a maximum coding

unit. The prediction units 1060 are partitions of prediction units of each of the coding units 1010, and the transformation units 1070 are transformation units of each of the coding units 1010.

[155] When a depth of a maximum coding unit is 0 in the coding units 1010, depths of coding units 1012 and 1054 are 1, depths of coding units 1014, 1016, 1018, 1028, 1050, and 1052 are 2, depths of coding units 1020, 1022, 1024, 1026, 1030, 1032, and 1048 are 3, and depths of coding units 1040, 1042, 1044, and 1046 are 4.

[156] In the prediction units 1060, some encoding units 1014, 1016, 1022, 1032, 1048, 1050, 1052, and 1054 are obtained by splitting the coding units in the encoding units 1010. In other words, partition types in the coding units 1014, 1022, 1050, and 1054 have a size of 2NxN, partition types in the coding units 1016, 1048, and 1052 have a size of Nx2N, and a partition type of the coding unit 1032 has a size of NxN. Prediction units and partitions of the coding units 1010 are smaller than or equal to each coding unit.

[157] Transformation or inverse transformation is performed on image data of the coding unit 1052 in the transformation units 1070 in a data unit that is smaller than the coding unit 1052. Also, the coding units 1014, 1016, 1022, 1032, 1048, 1050, and 1052 in the transformation units 1070 are different from those in the prediction units 1060 in terms of sizes and shapes. In other words, the video encoding and decoding apparatuses 100 and 200 may perform intra prediction, motion estimation, motion compensation, transformation, and inverse transformation individually on a data unit in the same coding unit.

[158] Accordingly, encoding is recursively performed on each of coding units having a hierarchical structure in each region of a maximum coding unit to determine an optimum coding unit, and thus coding units having a recursive tree structure may be obtained. Encoding information may include split information about a coding unit, information about a partition type, information about a prediction mode, and information about a size of a transformation unit. Table 1 shows the encoding information that may be set by the video encoding and decoding apparatuses 100 and 200.

[159] Table 1

[Table 1]

[Table]

Split Information 0 (Encoding on Coding Unit having Size of 2Nx2N and Current Depth of d)					Split Information 1
Prediction Mode	Partition Type		Size of Transformation Unit		Repeatedly Encode Coding Units having Lower Depth of d+1
IntraInter Skip (Only 2Nx2N)	Symmetrical Partition Type	Asymmetrical Partition Type	Split Information 0 of Transformation Unit	Split Information 1 of Transformation Unit	
	2Nx2N2NxNNx2NNxN	2NxnU2NxnDnLx2NnRx2N	2Nx2N	NxN(Symmetrical Type)N/2xN/2(Asymmetrical Type)	

[160] The output unit 130 of the video encoding apparatus 100 may output the encoding information about the coding units having a tree structure, and the image data and encoding information extractor 220 of the video decoding apparatus 200 may extract the encoding information about the coding units having a tree structure from a received bitstream.

[161] Split information indicates whether a current coding unit is split into coding units of a lower depth. If split information of a current depth d is 0, a depth, in which a current coding unit is no longer split into a lower depth, is a coded depth, and thus information about a partition type, prediction mode, and a size of a transformation unit may be defined for the coded depth. If the current coding unit is further split according to the split information, encoding is independently performed on four split coding units of a lower depth.

[162] A prediction mode may be one of an intra mode, an inter mode, and a skip mode. The intra mode and the inter mode may be defined in all partition types, and the skip mode is defined only in a partition type having a size of 2Nx2N.

[163] The information about the partition type may indicate symmetrical partition types having sizes of 2Nx2N, 2NxN, Nx2N, and NxN, which are obtained by symmetrically splitting a height or a width of a prediction unit, and asymmetrical partition types having sizes of 2NxnU, 2NxnD, nLx2N, and nRx2N, which are obtained by asymmetrically splitting the height or width of the prediction unit. The asymmetrical partition types having the sizes of 2NxnU and 2NxnD may be respectively obtained by

splitting the height of the prediction unit in 1:3 and 3:1, and the asymmetrical partition types having the sizes of nLx2N and nRx2N may be respectively obtained by splitting the width of the prediction unit in 1:3 and 3:1

[164] The size of the transformation unit may be set to be two types in the intra mode and two types in the inter mode. In other words, if split information of the transformation unit is 0, the size of the transformation unit may be 2Nx2N, which is the size of the current coding unit. If split information of the transformation unit is 1, the trans-formation units may be obtained by splitting the current coding unit. Also, if a partition type of the current coding unit having the size of 2Nx2N is a symmetrical partition type, a size of a transformation unit may be NxN, and if the partition type of the current coding unit is an asymmetrical partition type, the size of the transformation unit may be N/2xN/2.

[165] The encoding information about coding units having a tree structure may include at least one of a coding unit corresponding to a coded depth, a prediction unit, and a minimum unit. The coding unit corresponding to the coded depth may include at least one of a prediction unit and a minimum unit containing the same encoding in-formation.

[166] Accordingly, it is determined whether adjacent data units are included in the same coding unit corresponding to the coded depth by comparing encoding information of the adjacent data units. Also, a corresponding coding unit corresponding to a coded depth is determined by using encoding information of a data unit, and thus a dis-tribution of coded depths in a maximum coding unit may be determined.

[167] Accordingly, if a current coding unit is predicted based on encoding information of adjacent data units, encoding information of data units in deeper coding units adjacent to the current coding unit may be directly referred to and used.

[168] Alternatively, if a current coding unit is predicted based on encoding information of adjacent data units, data units adjacent to the current coding unit are searched using encoded information of the data units, and the searched adjacent coding units may be referred for predicting the current coding unit.

[169] FIG. 13 is a diagram for describing a relationship between a coding unit, a prediction unit or a partition, and a transformation unit, according to encoding mode information of Table 1.

[170] A maximum coding unit 1300 includes coding units 1302, 1304, 1306, 1312, 1314, 1316, and 1318 of coded depths. Here, since the coding unit 1318 is a coding unit of a coded depth, split information may be set to 0. Information about a partition type of the coding unit 1318 having a size of 2Nx2N may be set to be one of a partition type 1322 having a size of 2Nx2N, a partition type 1324 having a size of 2NxN, a partition type 1326 having a size of Nx2N, a partition type 1328 having a size of NxN, a partition

type 1332 having a size of 2NxnU, a partition type 1334 having a size of 2NxnD, a partition type 1336 having a size of nLx2N, and a partition type 1338 having a size of nRx2N.

[171] When the partition type is set to be symmetrical, i.e. the partition type 1322, 1324, 1326, or 1328, a transformation unit 1342 having a size of 2Nx2N is set if split information (TU size flag) of a transformation unit is 0, and a transformation unit 1344 having a size of NxN is set if a TU size flag is 1.

[172] When the partition type is set to be asymmetrical, i.e., the partition type 1332, 1334, 1336, or 1338, a transformation unit 1352 having a size of 2Nx2N is set if a TU size flag is 0, and a transformation unit 1354 having a size of N/2xN/2 is set if a TU size flag is 1.

[173] Referring to FIG. 13, the TU size flag is a flag having a value or 0 or 1, but the TU size flag is not limited to 1 bit, and a transformation unit may be hierarchically split having a tree structure while the TU size flag increases from 0.

[174] In this case, the size of a transformation unit that has been actually used may be expressed by using a TU size flag of a transformation unit, according to an exemplary embodiment, together with a maximum size and minimum size of the transformation unit. According to an exemplary embodiment, the video encoding apparatus 100 is capable of encoding maximum transformation unit size information, minimum transformation unit size information, and a maximum TU size flag. The result of encoding the maximum transformation unit size information, the minimum transformation unit size information, and the maximum TU size flag may be inserted into an SPS. According to an exemplary embodiment, the video decoding apparatus 200 may decode video by using the maximum transformation unit size information, the minimum transformation unit size information, and the maximum TU size flag.

[175] For example, if the size of a current coding unit is 64x64 and a maximum transformation unit size is 32x32, then the size of a transformation unit may be 32x32 when a TU size flag is 0, may be 16x16 when the TU size flag is 1, and may be 8x8 when the TU size flag is 2.

[176] As another example, if the size of the current coding unit is 32x32 and a minimum transformation unit size is 32x32, then the size of the transformation unit may be 32x32 when the TU size flag is 0. Here, the TU size flag cannot be set to a value other than 0, since the size of the transformation unit cannot be less than 32x32.

[177] As another example, if the size of the current coding unit is 64x64 and a maximum TU size flag is 1, then the TU size flag may be 0 or 1. Here, the TU size flag cannot be set to a value other than 0 or 1.

[178] Thus, if it is defined that the maximum TU size flag is 'MaxTransformSizeIndex', a minimum transformation unit size is 'MinTransformSize', and a transformation unit

size is 'RootTuSize' when the TU size flag is 0, then a current minimum transformation unit size 'CurrMinTuSize' that can be determined in a current coding unit, may be defined by Equation (1):

[179] CurrMinTuSize = max(MinTransformSize, RootTuSize/(2^MaxTransformSizeIndex))..... (1)

[180] Compared to the current minimum transformation unit size 'CurrMinTuSize' that can be determined in the current coding unit, a transformation unit size 'RootTuSize' when the TU size flag is 0 may denote a maximum transformation unit size that can be selected in the system. In Equation (1), 'RootTuSize/(2^MaxTransformSizeIndex)' denotes a transformation unit size when the transformation unit size 'RootTuSize', when the TU size flag is 0, is split a number of times corresponding to the maximum TU size flag, and 'MinTransformSize' denotes a minimum transformation size. Thus, a smaller value from among 'RootTuSize/(2^MaxTransformSizeIndex)' and 'MinTransformSize' may be the current minimum transformation unit size 'CurrMinTuSize' that can be determined in the current coding unit.

[181] According to an exemplary embodiment, the maximum transformation unit size RootTuSize may vary according to the type of a prediction mode.

[182] For example, if a current prediction mode is an inter mode, then 'RootTuSize' may be determined by using Equation (2) below. In Equation (2), 'MaxTransformSize' denotes a maximum transformation unit size, and 'PUSize' denotes a current prediction unit size.

[183] RootTuSize = min(MaxTransformSize, PUSize) ..... (2)

[184] That is, if the current prediction mode is the inter mode, the transformation unit size 'RootTuSize' when the TU size flag is 0, may be a smaller value from among the maximum transformation unit size and the current prediction unit size.

[185] If a prediction mode of a current partition unit is an intra mode, 'RootTuSize' may be determined by using Equation (3) below. In Equation (3), 'PartitionSize' denotes the size of the current partition unit.

[186] RootTuSize = min(MaxTransformSize, PartitionSize) .....(3)

[187] That is, if the current prediction mode is the intra mode, the transformation unit size 'RootTuSize' when the TU size flag is 0 may be a smaller value from among the maximum transformation unit size and the size of the current partition unit.

[188] However, the current maximum transformation unit size 'RootTuSize' that varies according to the type of a prediction mode in a partition unit is just an example and is not limited thereto.

[189] FIG. 14 is a flowchart illustrating a method of encoding a video, according to an exemplary embodiment.

[190] In operation 1210, a current picture is split into at least one maximum coding unit. A

maximum depth indicating the total number of possible splitting times may be predetermined.

[191] In operation 1220, a coded depth to output a final encoding result according to at least one split region, which is obtained by splitting a region of each maximum coding unit according to depths, is determined by encoding the at least one split region, and a coding unit according to a tree structure is determined.

[192] The maximum coding unit is spatially split whenever the depth deepens, and thus is split into coding units of a lower depth. Each coding unit may be split into coding units of another lower depth by being spatially split independently from adjacent coding units. Encoding is repeatedly performed on each coding unit according to depths.

[193] Also, a transformation unit according to partition types having the least encoding error is determined for each deeper coding unit. In order to determine a coded depth having a minimum encoding error in each maximum coding unit, encoding errors may be measured and compared in all deeper coding units according to depths.

[194] In operation 1230, encoded image data constituting the final encoding result according to the coded depth is output for each maximum coding unit, with encoding information about the coded depth and an encoding mode. The information about the encoding mode may include information about a coded depth or split information, information about a partition type of a prediction unit, a prediction mode, and a size of a transformation unit. The encoded information about the encoding mode may be transmitted to a decoder with the encoded image data.

[195] FIG. 15 is a flowchart illustrating a method of decoding a video, according to an exemplary embodiment.

[196] In operation 1310, a bitstream of an encoded video is received and parsed.

[197] In operation 1320, encoded image data of a current picture assigned to a maximum coding unit, and information about a coded depth and an encoding mode according to maximum coding units are extracted from the parsed bitstream. The coded depth of each maximum coding unit is a depth having the least encoding error in each maximum coding unit. In encoding each maximum coding unit, the image data is encoded based on at least one data unit obtained by hierarchically splitting the each maximum coding unit according to depths.

[198] According to the information about the coded depth and the encoding mode, the maximum coding unit may be split into coding units having a tree structure. Each of the coding units having the tree structure is determined as a coding unit corresponding to a coded depth, and is optimally encoded as to output the least encoding error. Accordingly, encoding and decoding efficiency of an image may be improved by decoding each piece of encoded image data in the coding units after determining at least one coded depth according to coding units.

[199] In operation 1330, the image data of each maximum coding unit is decoded based on the information about the coded depth and the encoding mode according to the maximum coding units. The decoded image data may be reproduced by a reproducing apparatus, stored in a storage medium, or transmitted through a network.

[200] The exemplary embodiments can be written as computer programs and can be implemented in general-use digital computers that execute the programs using a computer readable recording medium. Examples of the computer readable recording medium include magnetic storage media (e.g., ROM, floppy disks, hard disks, etc.) and optical recording media (e.g., CD-ROMs, or DVDs).

[201] While this invention has been particularly shown and described with reference to exemplary embodiments thereof, it will be understood by those of ordinary skill in the art that various changes in form and details may be made therein without departing from the scope of the invention as defined by the appended claims. The exemplary embodiments should be considered in descriptive sense only and not for purposes of limitation. Therefore, the scope of the invention is defined not by the detailed description of the invention but by the appended claims, and all differences within the scope will be construed as being included in the present invention.

Claims

1. A method of decoding a video, the method comprising:

extracting, from a bitstream, maximum depth information indicating a total number of possible times the maximum coding unit is split;

extracting, from the bitstream, split information for a coding unit indicating whether a coding unit of a current depth is split into coding units of a lower depth;

determining, based on the maximum depth information and the split information for a coding unit, coding units of a hierarchical structure comprising coding units which are not split any more among coding units split from the maximum coding unit;

extracting, from the bitstream, split information for a transformation unit indicating whether transformation units are determined from the coding units;

determining, based on the split information for a transformation unit, transformation units in the coding units; and

decoding the coding units of the hierarchical structure in the maximum coding unit,

wherein:

an image is divided into a plurality of maximum coding units using a size of the maximum coding unit,

the maximum coding unit is hierarchically split into coding units of depths including at least one of a current depth and a lower depth according to the split information for a coding unit,

when the split information indicates a split for the current depth, the coding unit of a current depth is split into four coding units of a lower depth, independently from neighboring coding units, and

when the split information indicates a non-split for the current depth, one or more transformation units are obtained from the coding unit of the current depth using the split information for a transformation unit.

2. An apparatus for decoding a video, the apparatus comprising:

a receiver configured to extract, from a bitstream, maximum depth information indicating a total number of possible times the maximum coding unit is split, split information for a coding unit indicating whether a coding unit of a current depth is

split into coding units of a lower depth, and split information for a transformation unit indicating whether transformation units are determined from the coding units; and

a decoder configured to determine, based on the maximum depth information and the split information, coding units of a hierarchical structure comprising coding units which are not split any more among coding units split from the maximum coding unit, determine, based on the split information for a transformation unit, transformation units in the coding units, and decode the coding units of the hierarchical structure in the maximum coding unit,

wherein:

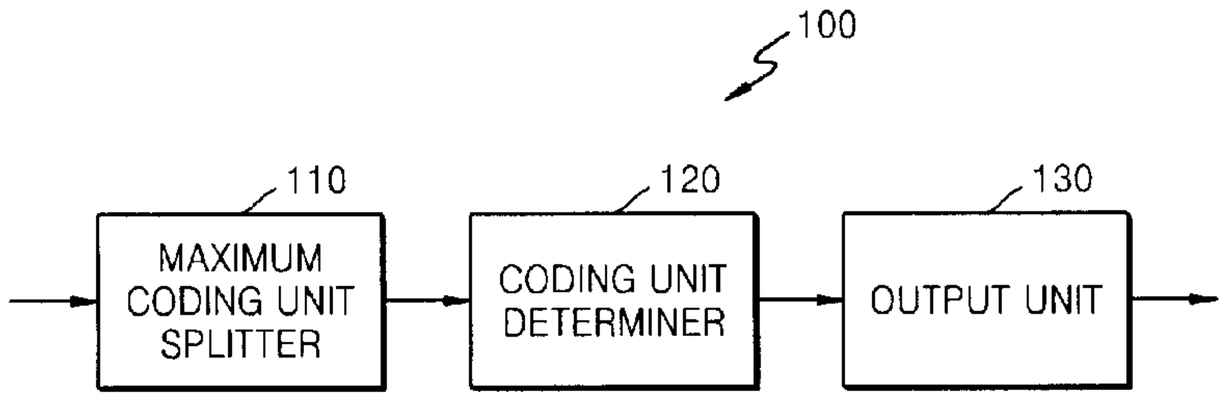
an image is divided into a plurality of maximum coding units using a size of the maximum coding unit,

the maximum coding unit is hierarchically split into coding units of depths including at least one of a current depth and a lower depth according to the split information for a coding unit,

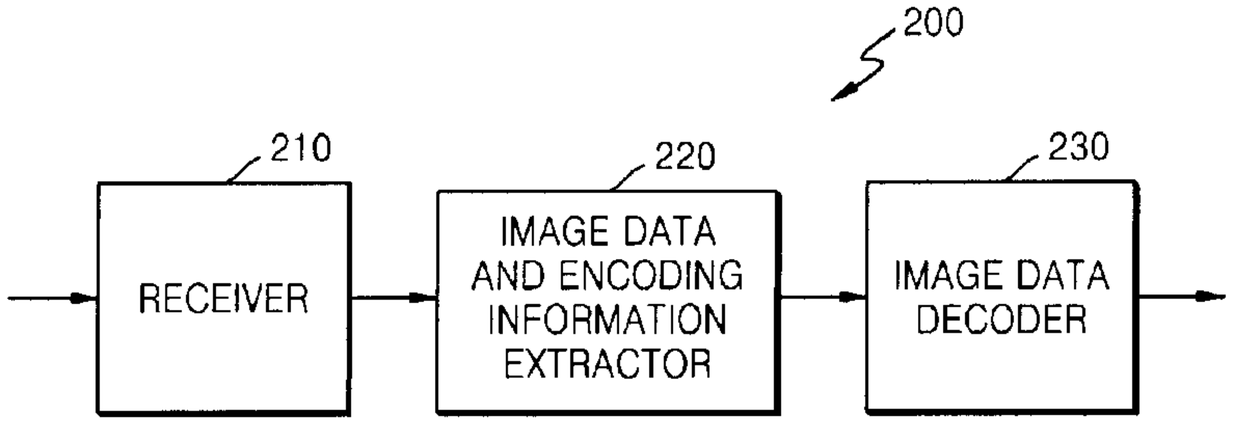
when the split information indicates a split for the current depth, the coding unit of a current depth is split into four coding units of a lower depth, independently from neighboring coding units, and

when the split information indicates a non-split for the current depth, one or more prediction units are obtained from the coding unit of the current depth using the split information for a transformation unit.

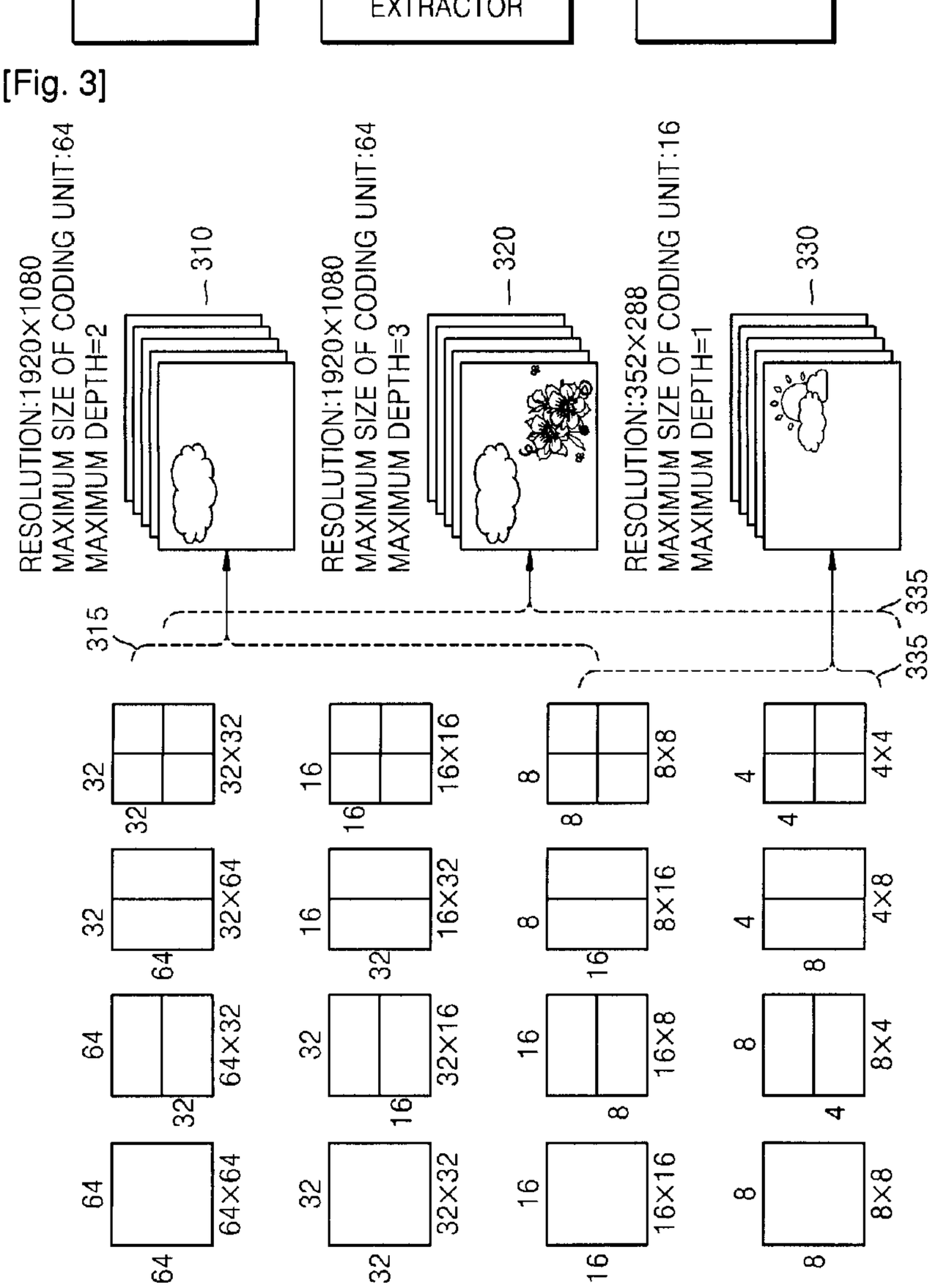
[Fig. 1]



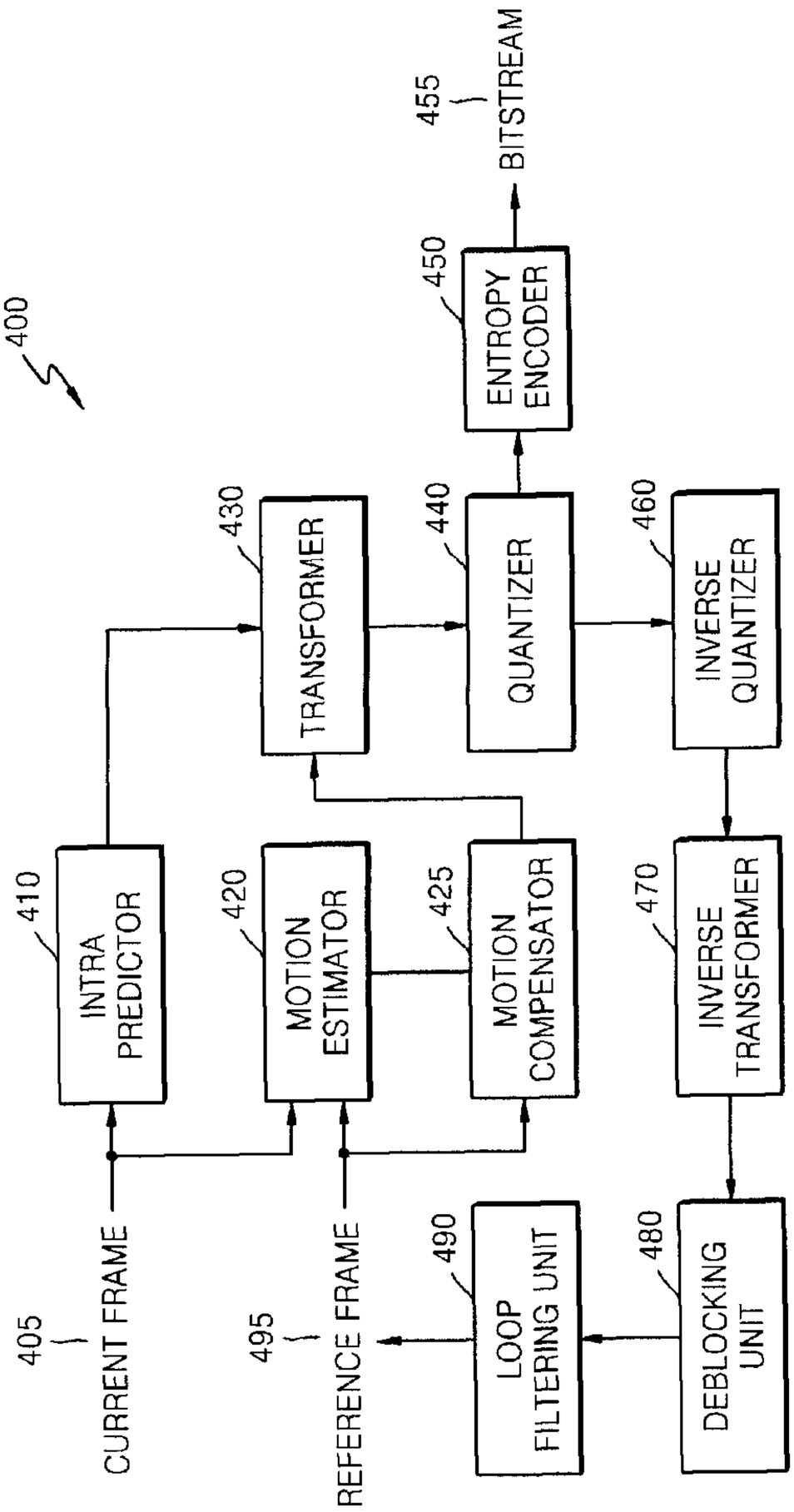
[Fig. 2]



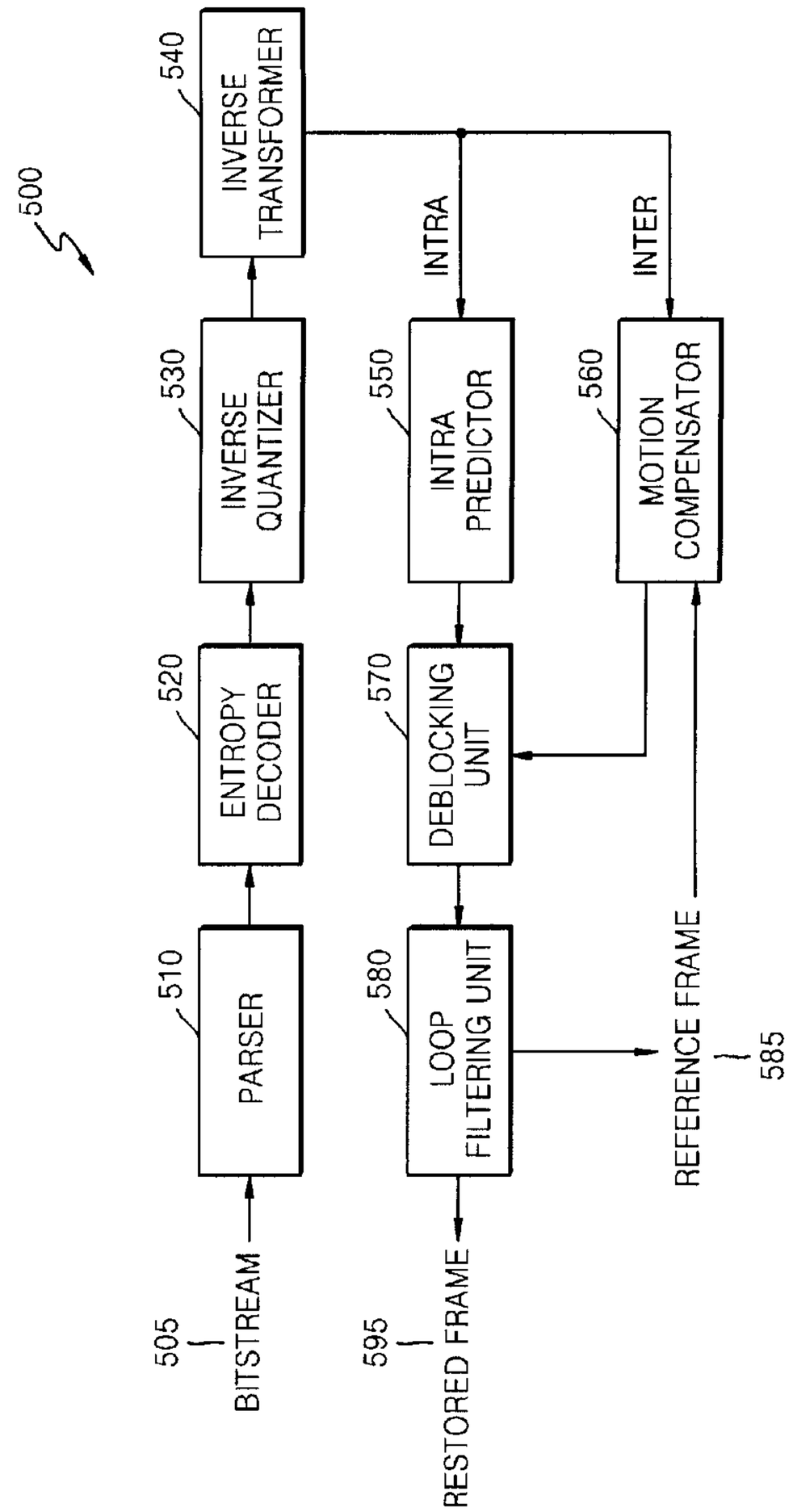
[Fig. 3]



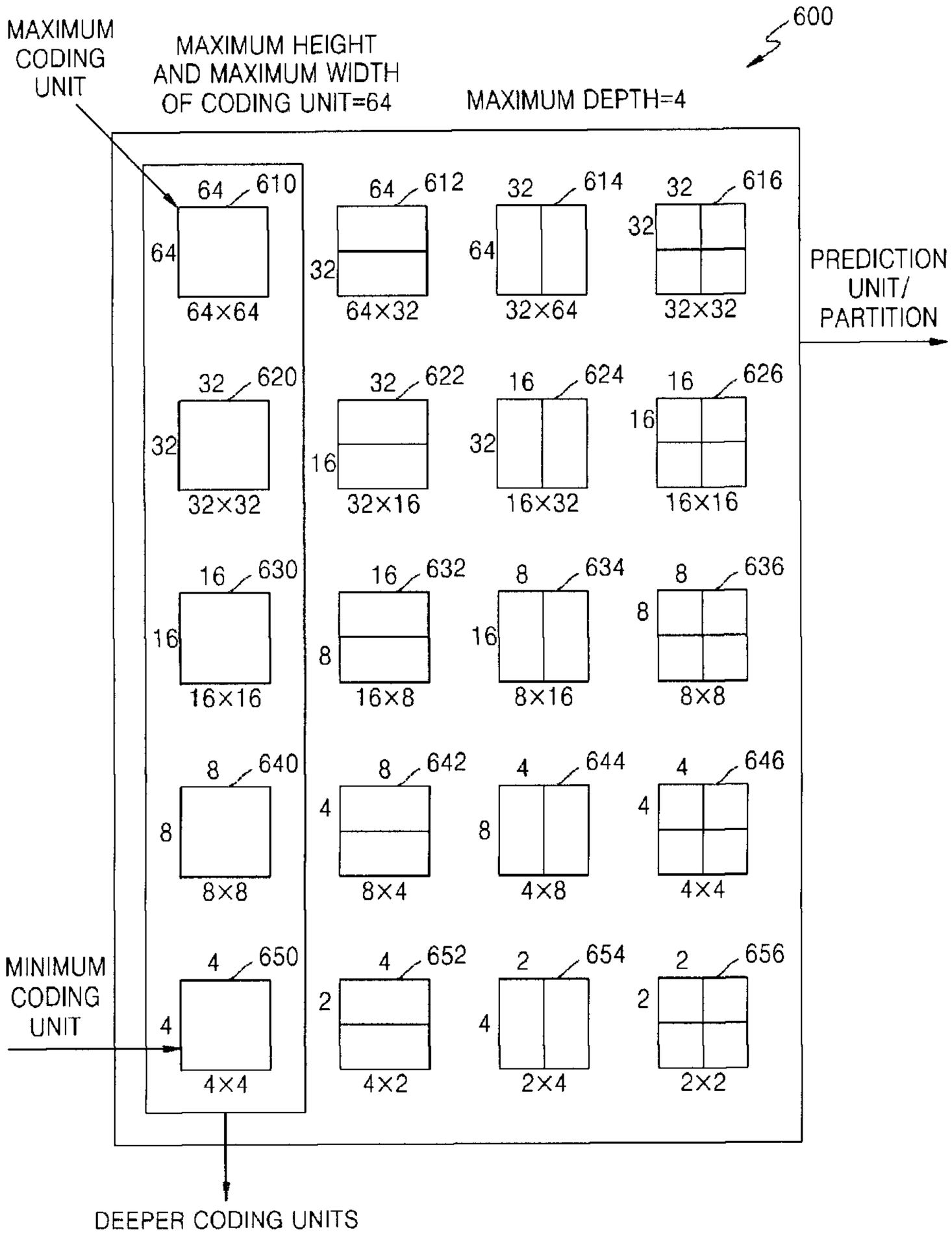
[Fig. 4]


400
455
BITSTREAM
450
ENTROPY ENCODER
430
TRANSFORMER
440
QUANTIZER
460
INVERSE QUANTIZER
410
INTRA PREDICTOR
420
MOTION ESTIMATOR
425
MOTION COMPENSATOR
470
INVERSE TRANSFORMER
405
CURRENT FRAME
495
REFERENCE FRAME
490
LOOP FILTERING UNIT
480
DEBLOCKING UNIT

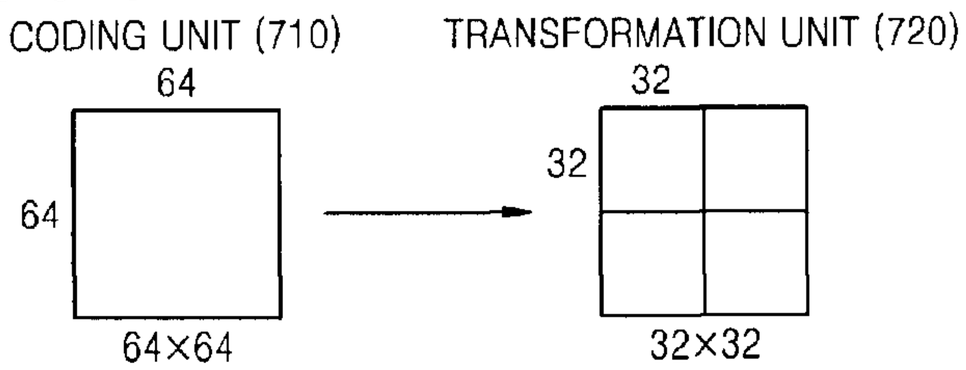
[Fig. 5]


500
505
BITSTREAM
510
PARSER
520
ENTROPY DECODER
530
INVERSE QUANTIZER
540
INVERSE TRANSFORMER
550
INTRA PREDICTOR
560
MOTION COMPENSATOR
570
DEBLOCKING UNIT
580
LOOP FILTERING UNIT
585
REFERENCE FRAME
595
RESTORED FRAME
INTRA
INTER

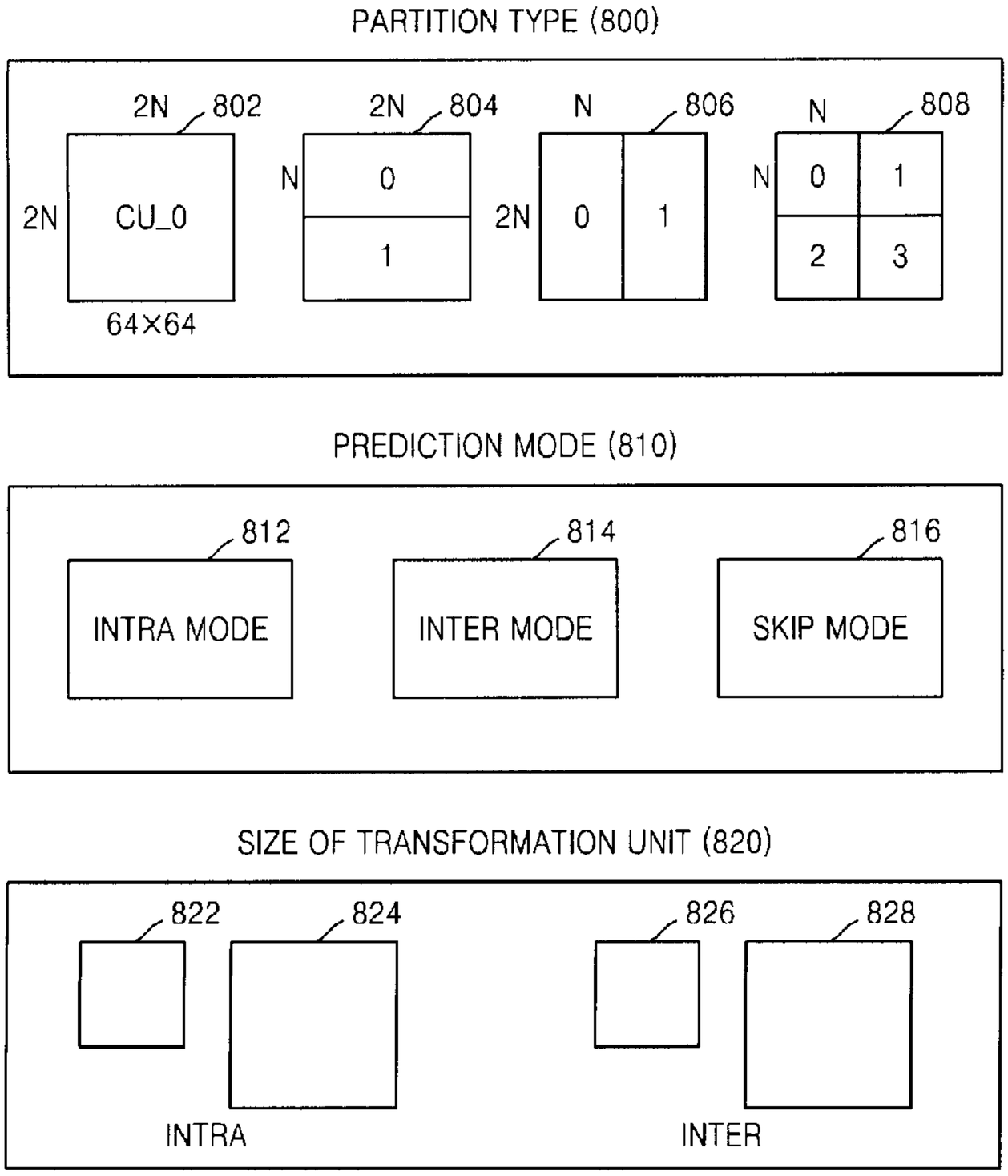
[Fig. 6]


600
MAXIMUM CODING UNIT
MAXIMUM HEIGHT AND MAXIMUM WIDTH OF CODING UNIT=64
MAXIMUM DEPTH=4
PREDICTION UNIT/ PARTITION
64
610
64×64
64
612
64×32
32
614
32×64
32
616
32×32
32
620
32×32
32
622
32×16
16
624
16×32
16
626
16×16
16
630
16×16
16
632
16×8
8
634
8×16
8
636
8×8
8
640
8×8
8
642
8×4
4
644
4×8
4
646
4×4
MINIMUM CODING UNIT
4
650
4×4
4
652
4×2
2
654
2×4
2
656
2×2
DEEPER CODING UNITS

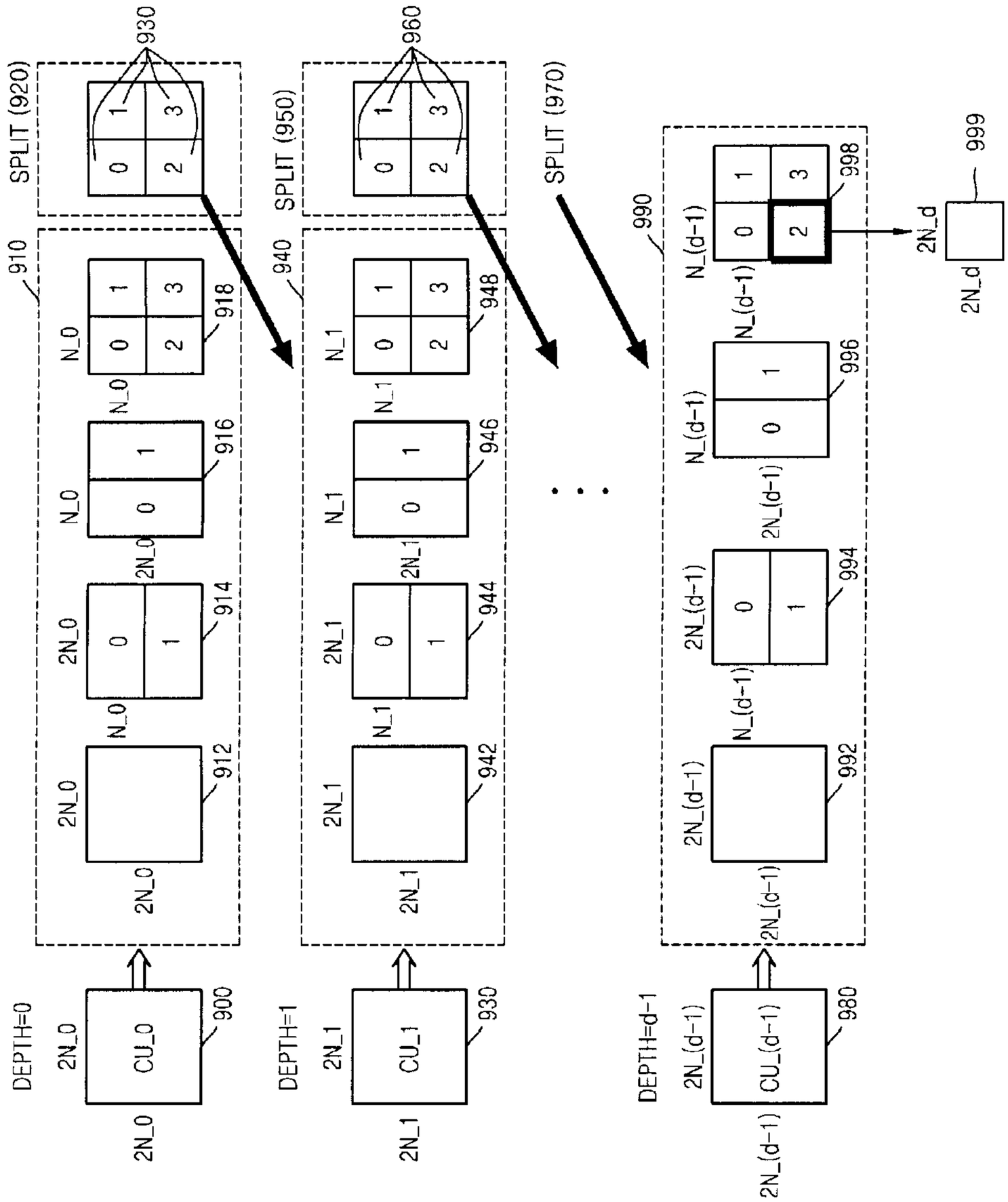
[Fig. 7]


CODING UNIT (710)
TRANSFORMATION UNIT (720)
64
64
64×64
32
32
32×32

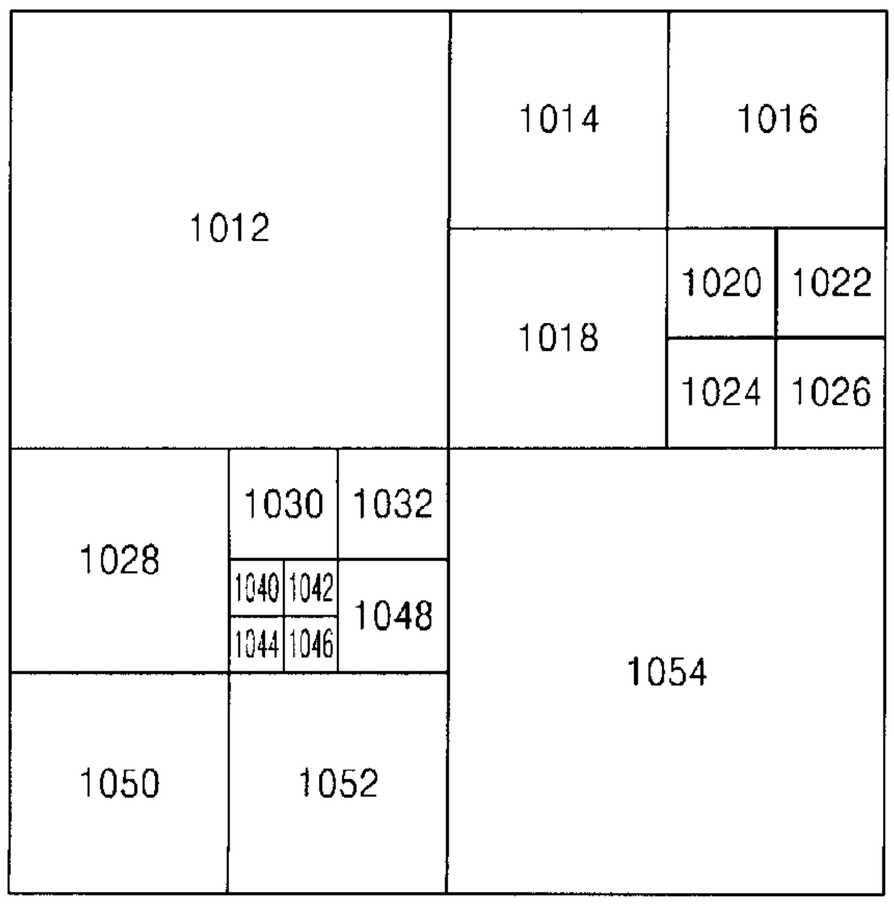
[Fig. 8]


PARTITION TYPE (800)
2N
802
CU_0
64×64
2N
804
N
0
1
N
806
2N
0
1
N
808
N
0
1
2
3
PREDICTION MODE (810)
812
INTRA MODE
814
INTER MODE
816
SKIP MODE
SIZE OF TRANSFORMATION UNIT (820)
822
824
826
828
INTRA
INTER

[Fig. 9]

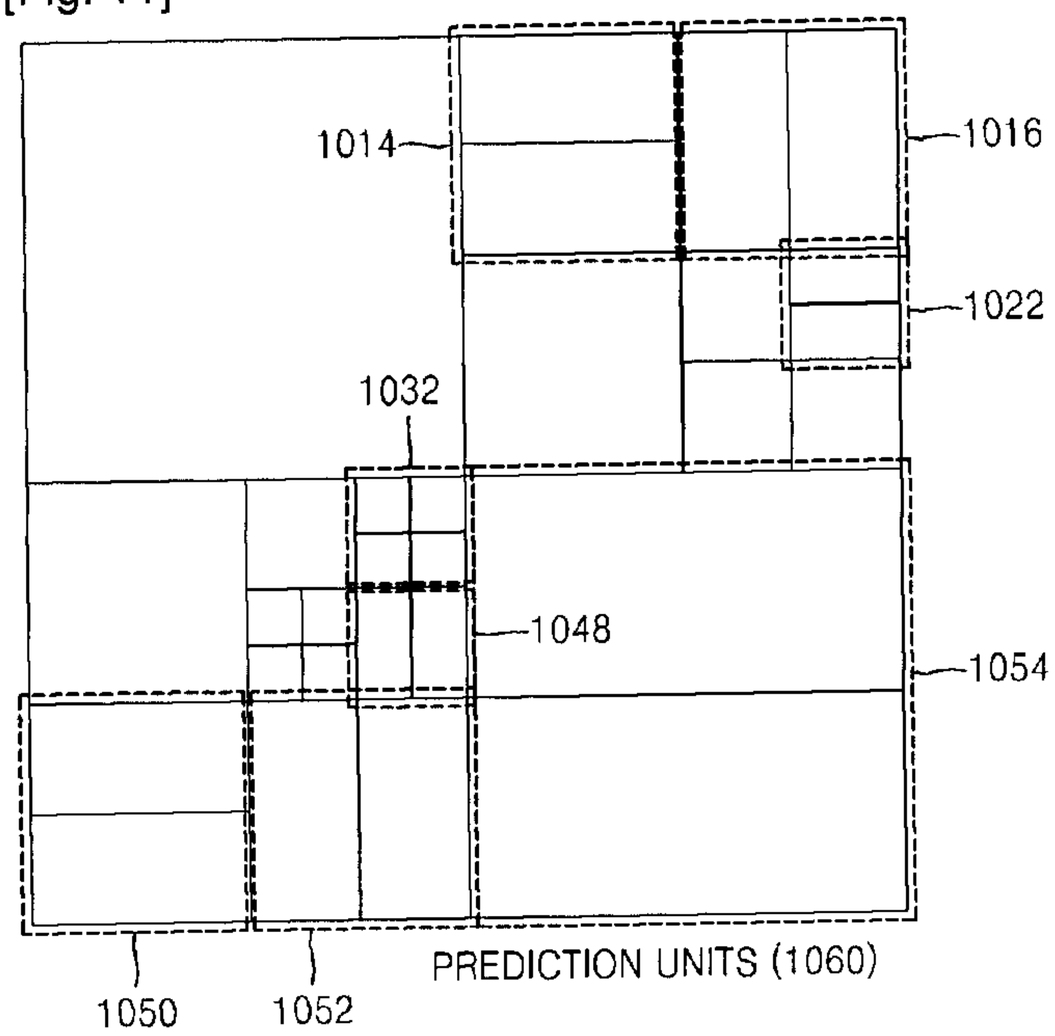


[Fig. 10]

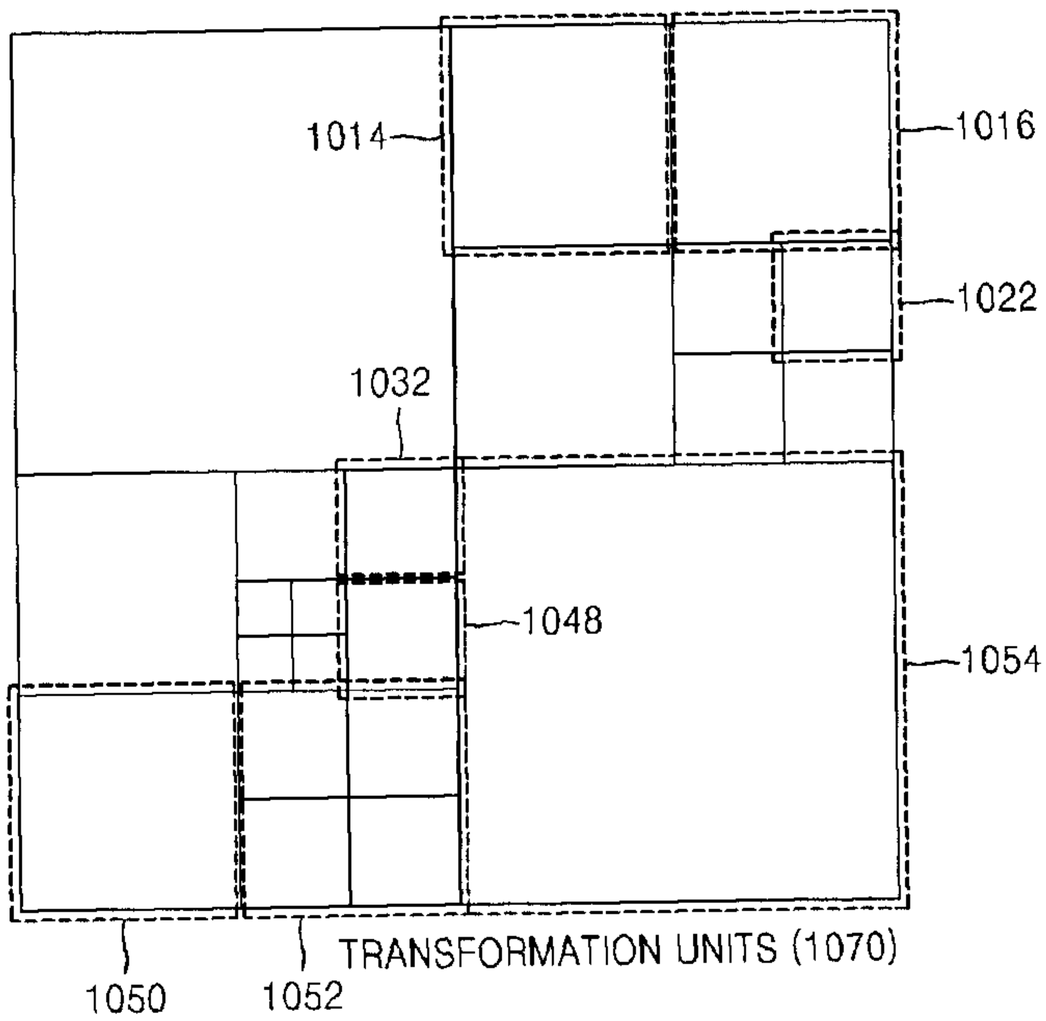


CODING UNITS (1010)

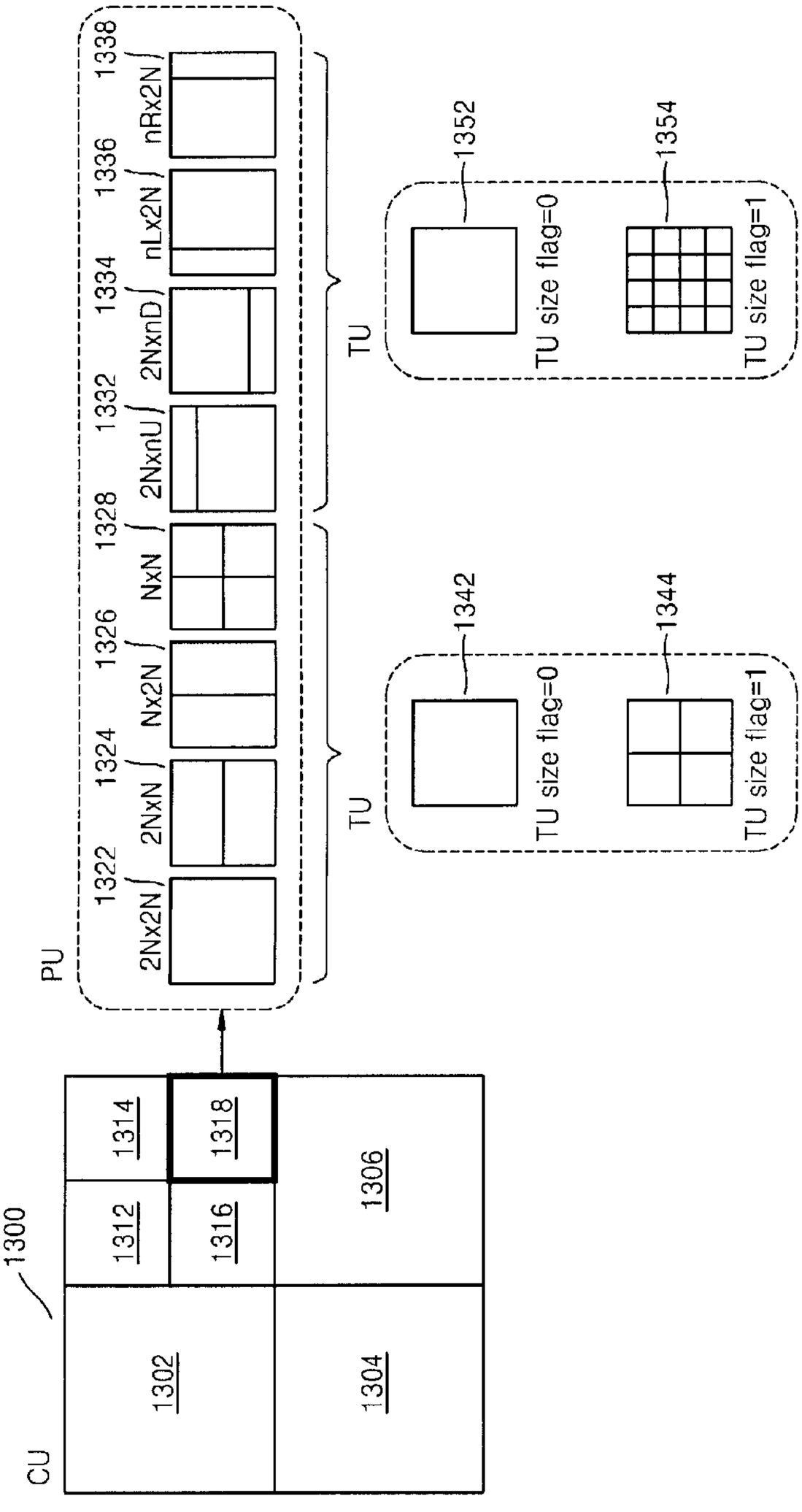
[Fig. 11]


1014
1016
1022
1032
1048
1054
1050
1052
PREDICTION UNITS (1060)

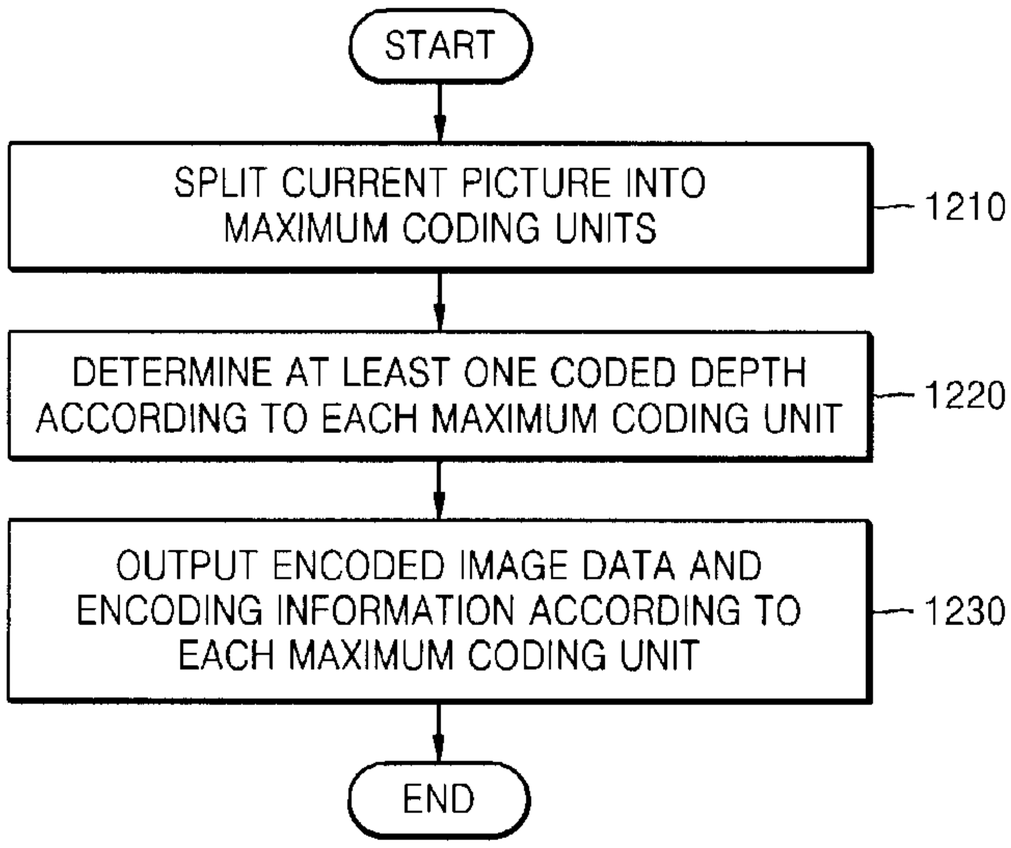
[Fig. 12]


1014
1016
1022
1032
1048
1054
1050
1052
TRANSFORMATION UNITS (1070)

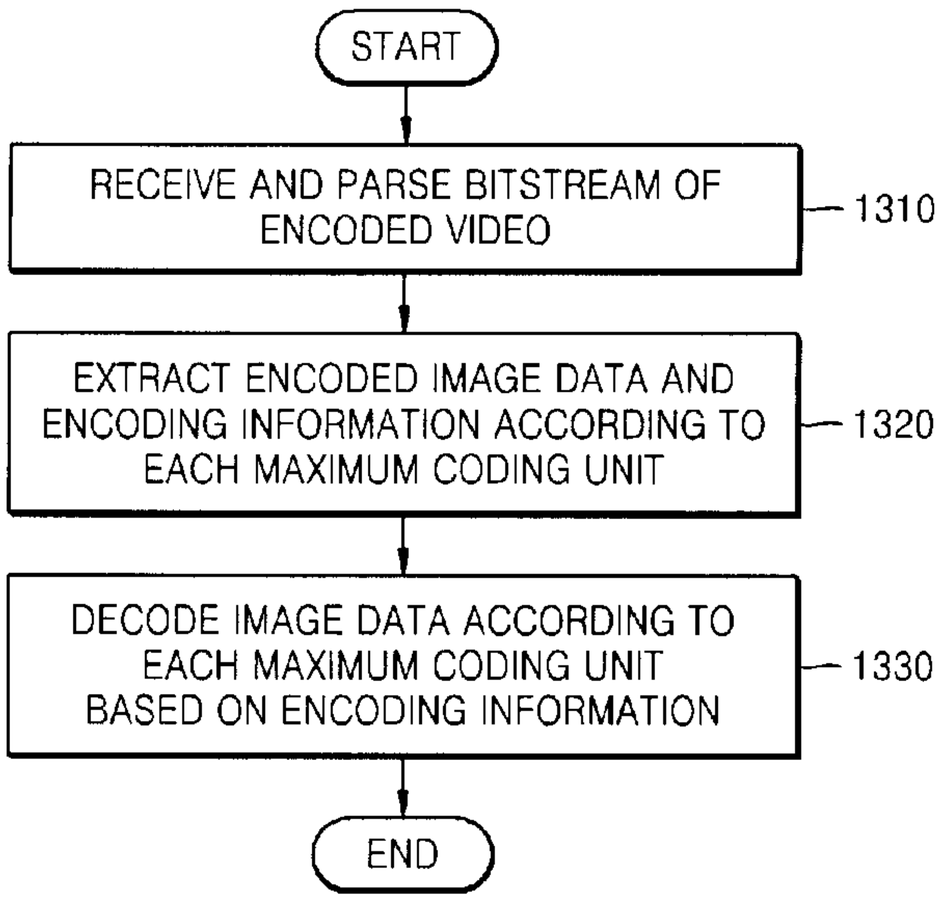
[Fig. 13]

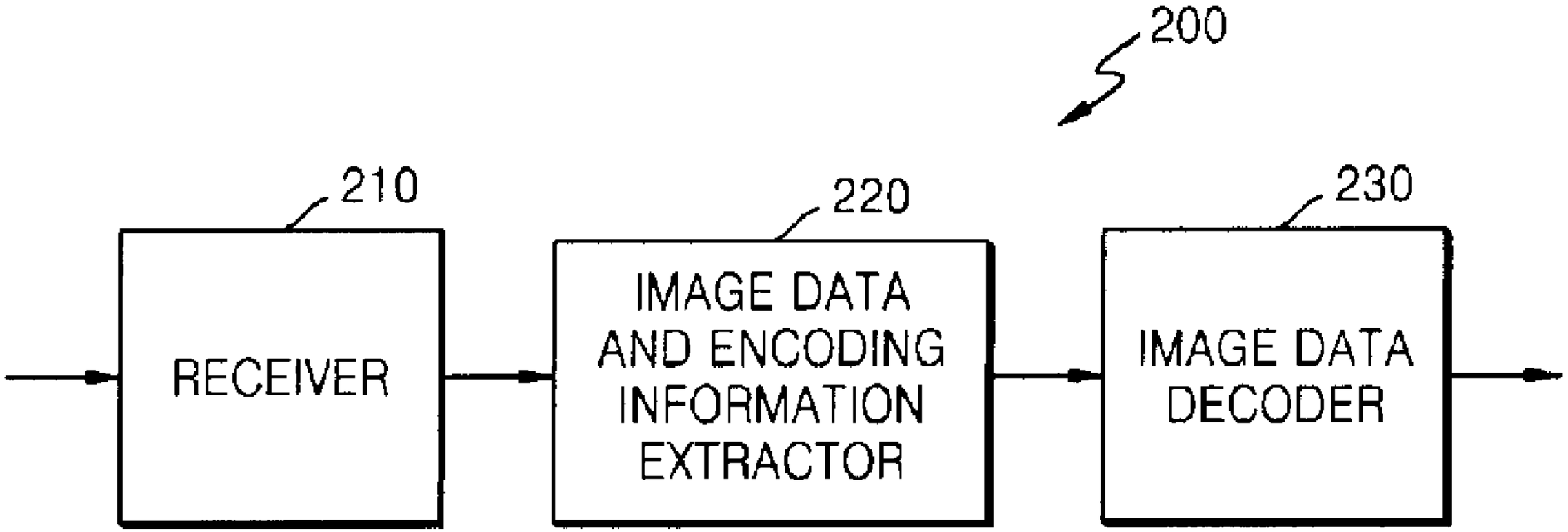

CU
1300
1302
1304
1306
1312
1314
1316
1318
PU
2Nx2N
1322
2NxN
1324
Nx2N
1326
NxN
1328
2NxnU
1332
2NxnD
1334
nLx2N
1336
nRx2N
1338
TU
1342
TU size flag=0
1344
TU size flag=1
TU
1352
TU size flag=0
1354
TU size flag=1

[Fig. 14]


START
SPLIT CURRENT PICTURE INTO MAXIMUM CODING UNITS
1210
DETERMINE AT LEAST ONE CODED DEPTH ACCORDING TO EACH MAXIMUM CODING UNIT
1220
OUTPUT ENCODED IMAGE DATA AND ENCODING INFORMATION ACCORDING TO EACH MAXIMUM CODING UNIT
1230
END

[Fig. 15]


START
RECEIVE AND PARSE BITSTREAM OF ENCODED VIDEO
1310
EXTRACT ENCODED IMAGE DATA AND ENCODING INFORMATION ACCORDING TO EACH MAXIMUM CODING UNIT
1320
DECODE IMAGE DATA ACCORDING TO EACH MAXIMUM CODING UNIT BASED ON ENCODING INFORMATION
1330
END


200
210
220
230
RECEIVER
IMAGE DATA AND ENCODING INFORMATION EXTRACTOR
IMAGE DATA DECODER