DESCRIPTION

Title of Invention

IMAGE PREDICTIVE ENCODING DEVICE, IMAGE PREDICTIVE ENCODING METHOD, IMAGE PREDICTIVE ENCODING PROGRAM, IMAGE PREDICTIVE DECODING DEVICE, IMAGE PREDICTIVE DECODING METHOD, AND IMAGE PREDICTIVE DECODING PROGRAM

Technical Field

[0001] Embodiments of the present invention relate to an image predictive encoding device, an image predictive encoding method, an image predictive encoding program, an image predictive decoding device, an image predictive decoding method, and an image predictive decoding program and, more particularly, to an image predictive encoding device, an image predictive encoding method, an image predictive encoding program, an image predictive decoding device, an image predictive decoding method, and an image predictive decoding program to generate a predicted signal of a target block, using motion information of neighborhood blocks.

Background Art

[0002] The compression encoding technologies are used for efficient transmission and storage of still pictures and video data. The techniques defined in MPEG-1 to 4 and ITU (International Telecommunication Union) H.261 to H.264 are commonly used for video data.

[0003] In these encoding techniques, a picture as an encoding target is partitioned into a plurality of blocks and then an encoding process and a

decoding process are carried out on a block basis. In intra-picture predictive encoding, a predicted signal is generated using a reconstructed picture signal (or a restored signal of compressed picture data) of a neighboring region present in the same picture as a target block and then a difference signal of a difference between the predicted signal and a pixel signal of the target block is encoded. In inter-picture predictive encoding, a predicted signal is generated with compensation for motion with reference to a reconstructed picture signal of a neighboring region present in a picture different from a target block, and a difference signal of a difference between the predicted signal and a signal of the target block is encoded.

[0004] For example, the intra-picture predictive encoding of H.264 employs a method of generating the predicted signal by extrapolating previously-reproduced pixel values of a region adjacent to a block serving as an encoding target, in predetermined directions. Fig. 17 is a schematic diagram for explaining the intra-picture prediction method used in ITU H.264. In (A) of Fig. 17, a target block 802 is a block serving as an encoding target, and a pixel group 801 consisting of pixels A to M adjacent to a boundary of the target block 802 is a neighboring region, which is a picture signal previously reproduced in past processing.

[0005] In the prediction shown in (A) of Fig. 17, the predicted signal is generated by downwardly extending the pixel group 801 as neighboring pixels located immediately above the target block 802. In the prediction shown in (B) of Fig. 17, the predicted signal is generated by rightwardly extending previously-reproduced pixels (I to L) located on

the left side of the target block 804. For specific methods for generation of the predicted signal, reference is made, for example, to Patent Literature 1. A difference is calculated between each of nine predicted signals generated by the methods shown in (A) to (I) of Fig. 17 in the above-described manner, and the pixel signal of the target block, and these obtained differences are compared to select a candidate to provide a minimum difference, as an optimum predicted signal. As described above, the predicted signal can be generated by extrapolation of pixels. The contents above are described in Patent Literature 1 below.

[0006] In ordinary inter-picture predictive encoding, the predicted signal is generated by a method of searching reconstructed pictures for a signal similar to the pixel signal of the target block. Then encoding is performed of a motion vector as a spatial displacement amount between the target block and a region composed of the detected signal, and a residual signal of the difference between the pixel signal of the target block and the predicted signal. The searching technique for the motion vector for each block as described above is called block matching.

[0007] Fig. 16 is a schematic diagram for explaining a block matching process. The below will describe a procedure of generating a predicted signal for an example of a target block 702 on an encoding target picture 701. A picture 703 is a reconstructed picture and a region 704 in the reconstructed picture is a region located spatially at the same position as the target block 702. In the block matching, a search range 705 that includes the region 704 is set. A region 706 with a signal to minimize the sum of absolute differences from the pixel signal of the target block

702 is detected from this search range. The signal of this region 706 is determined to be a predicted signal, and a displacement amount from the region 704 to the region 706 is detected as a motion vector 707. Also often employed is a method of preparing a plurality of reference pictures, selecting a reference picture to be applied to block matching for each target block, and detecting reference picture selection information. In some cases in H.264 a plurality of prediction types of different block sizes are prepared for encoding of motion vector, in order to adapt for local feature changes of pictures. The prediction types of H.264 are described, for example, in Patent Literature 2.

[0008] In compression encoding of video data, an encoding order of pictures (frames or fields) may be optional. For this reason, there are three types of techniques regarding the encoding order in inter- picture prediction to generate the predicted signal with reference to reconstructed pictures. The first technique is forward prediction to generate the predicted signal with reference to a past reconstructed picture in a display order, the second technique is backward prediction to generate the predicted signal with reference to a future reconstructed picture in the display order, and the third technique is bidirectional prediction to perform both forward prediction and backward prediction and average two predicted signals. The types of inter- picture predictions are described, for example, in Patent Literature 3.

[0009] In H.264, the third technique is performed using two reference picture lists each consisting of a plurality of reconstructed pictures as candidates for a reference picture. More specifically, the block matching is carried out using the plurality of reference pictures

registered in each reference picture list to detect two regions corresponding to the region 706, and two predicted signals thus detected are averaged.

[0010] An example of reference picture lists will be described with Figs. 5 and 6. In (A) of Fig. 5, a picture 505 indicates an encoding target picture and pictures 501 to 504 indicate reconstructed pictures. Each image (picture) is identified by a frame number (frame_num). Reference picture list identification List0 and List1 in (A) of Fig. 6 indicate two reference picture lists and in this example, four reference pictures are registered in each of the two reference picture lists. Each reference picture is identified by a reference index (ref_idx).

[0011] Reconstructed pictures that can be registered in the reference picture lists are basically optional. As in table 521 shown in (A) of Fig. 6, all the reference pictures registered in the two reference picture lists may be past reconstructed pictures. Since the two predicted signals both are forward predicted signals in this case, the prediction technique to average the two predicted signals is called bi-predictive prediction, instead of the bidirectional prediction.

[0012] In bi-predictive prediction, it is necessary to encode two combinations (motion information) of motion vectors and reference indices. Therefore, in order to decrease the bit count, there is a technique of performing the bi-predictive prediction using the motion information identified by List0 and the motion information identified by List1, out of the motion information of neighboring blocks.

Citation List

Patent Literature

[0013] Patent Literature 1: U.S. Pat. No. 6,765,964

Patent Literature 2: U.S. Pat. No. 7,003,035

Patent Literature 3: U.S. Pat. No. 6,259,739

Summary of Invention

Technical Problem

[0014] Bi-predictive prediction involves averaging of two similar predicted signals, whereby noise in the predicted signals can be more effectively removed by a smoothing effect. For this purpose, it is effective to acquire two signals with slightly different motion vectors from an identical reference picture. In the conventional methods of bi-predictive prediction using the motion information of neighboring blocks, however, the selection of motion information to be used in generation of two predicted signals is restricted by reference picture lists. Because of this restriction, even if a set of two pieces of motion information of neighboring blocks include motion vectors of close values based on the same reference picture, they cannot be selected.

[0015] It is assumed herein that when a block 400 in Fig. 7 is an encoding target block and blocks 401 to 403 are neighboring blocks to the target block, each of the three neighboring blocks has two pieces of motion information identified by List0 and List1. According to the technique of the background art, the two predicted signals are generated by selecting one of the three pieces of motion information identified by List0 and one of the three pieces of motion information identified by List1, which are the motion information of the three neighboring blocks. In general, a bit count necessary for reference index ref_idx becomes smaller when the reference index ref_idx is "0"; therefore, it is often the

case that all the reference indices ref_idx in the motion information of neighboring blocks are 0. In the case where the reference index is "0," use of the reference picture lists (521) in (A) of Fig. 6 results in acquiring the two predicted signals from a reference picture with the frame number (frame_num) of "3" and from a reference picture with the frame number of "2," respectively. In this case, it is not feasible to achieve a high smoothing effect.

[0016] As another example, let us assume a situation in which two reference picture lists are composed of different reference pictures. In the case, as shown in (B) of Fig. 5, where a picture 510 identified by the frame number of "3" is an encoding target picture, where pictures 507, 508, 509, and 511 identified by the frame numbers of "0," "1," "2," and "4", respectively, are reconstructed pictures, and where the reference picture lists are lists 522 shown in (B) of Fig. 6, the two predicted signals are generated from different reference pictures. In this case, it is also not feasible to achieve a high smoothing effect.

[0017] Therefore, there are needs for an image predictive encoding device, an image predictive encoding method, an image predictive encoding program, an image predictive decoding device, an image predictive decoding method, and an image predictive decoding program capable of suppressing the noise of predicted signals.

Solution to Problem

[0018] An aspect of the present invention relates to image predictive encoding.

[0019] An image predictive encoding device according to an aspect of the present invention is an image predictive encoding device

comprising: region partition means which partitions an input picture into a plurality of regions; predicted signal generation means which determines motion information for acquisition of a pixel signal having a high correlation with a pixel signal of a target region of an encoding target out of the plurality of regions partitioned by the region partition means, from a reconstructed picture, and which generates a predicted signal of the target region from the reconstructed picture on the basis of the motion information; motion information recording means which stores the motion information; residual signal generation means which generates a residual signal of a difference between the predicted signal of the target region and the pixel signal of the target region; residual signal encoding means which encodes the residual signal generated by the residual signal generation means; residual signal reconstruction means which decodes encoded date generated by the residual signal encoding means, to generate a reconstructed residual signal; addition means which adds the predicted signal to the reconstructed residual signal to generate a reconstructed pixel signal of the target region; and picture recording means which stores the reconstructed pixel signal generated by the addition means, as a part of a reconstructed picture, wherein the predicted signal generation means comprises: motion information derivation means which derives one or more motion information sets from plural pieces of motion information stored in the motion information recording means, each of the one or more motion information sets consisting of two pieces of motion information between which any one of constituent elements differs in value; motion prediction information selection means which selects a motion

information set to be used in generation of the predicted signal of the target region, from the one or more motion information sets and which directs storage of two pieces of motion information included in the selected set in the motion information recording means; and motion compensation means which combines two signals acquired from the reconstructed picture based on the two pieces of motion information in the selected set, to generate the predicted signal of the target region.

[0020] An image predictive encoding method according to an aspect of the present invention is an image predictive encoding method comprising: partitioning an input picture into a plurality of regions; generating a predicted signal of a target region of an encoding target out of the plurality of regions, said generating including determining motion information for acquisition of a pixel signal having a high correlation with a pixel signal of the target region from a reconstructed picture and generating the predicted signal from the reconstructed picture on the basis of the motion information; storing the motion information with motion information recording means; generating a residual signal of a difference between the predicted signal of the target region and the pixel signal of the target region; encoding the residual signal; decoding encoded data generated during said encoding of the residual signal, to generate a reconstructed residual signal; adding the predicted signal to the reconstructed residual signal to generate a reconstructed pixel signal of the target region; and storing the reconstructed pixel signal as a part of a reconstructed picture, by picture recording means, wherein the step of generating the predicted signal comprises: deriving one or more motion information sets out of plural pieces of motion information

stored in the motion information recording means, each consisting of two pieces of motion information between which any one of constituent elements differs in value; selecting, from the one or more motion information sets, a motion information set to be used in generation of the predicted signal of the target region, and directing storage of two pieces of motion information in the selected set in the motion information recording means; and combining two signals acquired from the reconstructed picture on the basis of the two pieces of motion information in the selected set, to generate the predicted signal of the target region.

[0021] An image predictive encoding program according to an aspect of the present invention makes a computer function as the region partition means, the predicted signal generation means, the motion information recording means, the residual signal generation means, the residual signal encoding means, the residual signal restoration means, the addition means, and the picture recording means. The predicted signal generation means comprises the motion information derivation means, the prediction information selection means, and the motion compensation means.

[0022] Image predictive encoding according to an aspect of the present invention permits two pieces of motion information between which any constituent element differs in value to be selected without being restricted by reference picture lists. As a result, the effect of smoothing can be achieved in the generation of the predicted signal.

[0023] In an embodiment, the image predictive encoding may be configured as follows: the picture recording means stores a plurality of

reconstructed pictures, each piece of the motion information contains information to identify the reconstructed picture, and a motion vector, the reconstructed pictures stored in the picture recording means are identified using respective frame numbers in encoding of the reconstructed pictures, and a derived motion information set is a motion information set including two pieces of motion information having an identical frame number of the reconstructed picture and different values of motion vectors.

[0024] In an embodiment, the image predictive encoding may be configured as follows: the picture recording means stores a plurality of reconstructed pictures, the plurality of reconstructed pictures stored in the picture recording means are identified using two reference picture lists each consisting of one or more reconstructed pictures, each piece of the motion information contains information to identify the two reference picture lists, and the motion vector, and a derived motion information set is a motion information set including two pieces of motion information having identical information to identify the reference picture lists, and different values of motion vectors.

[0025] In an embodiment, the image predictive encoding may be configured as follows: the motion information set to be used in generation of the predicted signal of the target region is selected from plural pieces of motion information associated with a region neighboring the target region.

[0026] In an embodiment, the image predictive encoding may be configured as follows: information to indicate the selected motion information set consisting of the two pieces of motion information is

encoded.

[0027] Another aspect of the present invention relates to image predictive decoding.

[0028] An image predictive decoding device according to another aspect of the present invention is an image predictive decoding device comprising: data analysis means which extracts encoded data of motion information for generation of a predicted signal of a target region of a decoding target out of a plurality of regions in a picture, and encoded data of a residual signal from compressed data; motion information recording means which stores the motion information; motion compensation means which generates the predicted signal of the target region on the basis of the motion information; residual signal decoding means which decodes a residual signal of the target region from the encoded data of the residual signal; picture recording means which adds the predicted signal to the decoded residual signal to restore a pixel signal of the target region, and which stores the restored pixel signal as a part of a reconstructed picture; motion information derivation means which derives, from plural pieces of motion information stored in the motion information recording means, one or more motion information sets each consisting of two pieces of motion information between which any one of constituent elements differs in value,; and motion information determination means which determines a motion information set to be used in generation of the predicted signal of the target region, out of the one or more motion information sets derived, and which directs storage of two pieces of motion information in the determined set in the motion information recording means, wherein the

motion compensation means combines two signals acquired from the reconstructed picture on the basis of the two pieces of motion information in the determined set, to generate the predicted signal of the target region.

[0029] An image predictive decoding method according to another aspect of the present invention is an image predictive decoding method comprising: extracting encoded data of motion information for generation of a predicted signal of a target region of a decoding target out of a plurality of regions in a picture, and encoded data of a residual signal from compressed data; storing the motion information with motion information recording means; generating the predicted signal of the target region on the basis of the motion information; restoring a decoded residual signal of the target region from encoded data of the residual signal; adding the predicted signal to the decoded residual signal to restore a pixel signal of the target region, and storing the restored pixel signal as a part of a reconstructed picture with picture recording means; deriving one or more motion information sets out of plural pieces of motion information stored in the motion information recording means, each of the one or more motion information sets consisting of two pieces of motion information between which any one of constituent elements differ in value; and determining a motion information set to be used in generation of the predicted signal of the target region from among the derived motion information sets, and directing storage of two pieces of motion information in the determined set in the motion information recording means, wherein the step of generating the predicted signal comprises combining two signals

acquired from the reconstructed picture on the basis of the two pieces of motion information in the determined set, to generate the predicted signal of the target region.

[0030] An image predictive decoding program according to another aspect of the present invention makes a computer function as the data analysis means, the motion information recording means, the motion compensation means, the residual signal restoration means, the picture recording means, the motion information derivation means, and the motion information determination means.

[0031] The image predictive decoding according to an aspect of the present invention allows a picture to be restored from compressed data generated by the image predictive encoding of the aforementioned aspect. Noise can be reduced in this restored picture.

[0032] In an embodiment, the image predictive decoding may be configured as follows: it comprises extracting encoded data of indication information, the encoded data extracted out of plural pieces of motion information stored in the motion information recording means to identify one motion information set consisting of two pieces of motion information between which any one of constituent elements differ in value, making prediction information decoding means decode the indication information, and extracting a motion information set to be used in generation of the predicted signal of the target region, out of the one or more motion information sets derived by the motion information derivation means, based on the decoded indication information.

[0033] In an embodiment, the image predictive decoding may be configured as follows: the picture recording means stores a plurality of

reconstructed pictures, the motion information contains information to identify the plurality of reconstructed pictures, and motion vectors, the plurality of reconstructed pictures stored in the picture recording means are identified using respective frame numbers in decoding of the respective reconstructed pictures, and a derived motion information set is a motion information set including two pieces of motion information having an identical frame number of the reconstructed picture and different values of motion vectors.

[0034] In an embodiment, the image predictive decoding may be configured as follows: the picture recording means stores a plurality of reconstructed pictures, the plurality of reconstructed pictures stored in the picture recording means are identified using two reference picture lists each consisting of one or more reconstructed pictures, each piece of the motion information is configured containing information to identify the two reference picture lists, and the motion vector, and a derived motion information set is a motion information set including two pieces of motion information having identical information to identify the reference picture lists, and different values of motion vectors.

[0035] In an embodiment, the image predictive decoding may be configured as follows: the motion information set to be used in generation of the predicted signal of the target region is derived from plural pieces of motion information associated with a region neighboring the target region.

Advantageous Effects of Invention

[0036] As described above, several aspects of the present invention provide the image predictive encoding device, image predictive

encoding method, image predictive encoding program, image predictive decoding device, image predictive decoding method, and image predictive decoding program capable of suppressing noise of the predicted signal.

Brief Description of Drawings

[0037] Fig. 1 is a drawing showing a configuration of an image predictive encoding device according to an embodiment.

Fig. 2 is a drawing showing a configuration of a predicted signal generation unit shown in Fig. 1.

Fig. 3 is a flowchart showing a process of a motion information derivation unit shown in Fig. 2.

Fig. 4 is a schematic diagram for explaining an example of a technique for selecting one piece of candidate prediction information from candidate motion prediction information (combination of two pieces of motion information) using neighboring pixels to a target block.

Fig. 5 is a schematic diagram for explaining examples of encoding orders of pictures.

Fig. 6 is a drawing for explaining examples of reference picture lists.

Fig. 7 is a drawing for explaining examples of neighboring blocks.

Fig. 8 is a drawing for explaining other examples of neighboring blocks.

Fig. 9 is a flowchart showing a procedure of an image predictive encoding method according to an embodiment.

Fig. 10 is a drawing showing a configuration of an image

predictive decoding device according to an embodiment.

Fig. 11 is a flowchart showing a procedure of an image predictive decoding method according to an embodiment.

Fig. 12 is a drawing showing a configuration of a program for causing a computer to execute the image predictive encoding method according to an embodiment.

Fig. 13 is a drawing showing a configuration of a program for causing a computer to execute the image predictive decoding method according to an embodiment.

Fig. 14 is a drawing showing a hardware configuration of a computer for executing a program recorded in a recording medium.

Fig. 15 is a perspective view of a computer for executing a program stored in a recording medium.

Fig. 16 is a schematic drawing for explaining a motion estimation process in inter-picture prediction.

Fig. 17 is a schematic diagram for explaining the conventional intra- picture prediction method.

Description of Embodiments

[0038] Some embodiments will be described below in detail with reference to the drawings. In the drawings identical or equivalent portions will be denoted by the same reference signs, without redundant description. In the present specification, "frames," "pictures," and "images" (501 to 511 in Fig. 5) mean the same.

[0039] Fig. 1 is a drawing showing a configuration of an image predictive encoding device according to an embodiment. The image predictive encoding device 100 shown in Fig. 1 is provided with an

input terminal 101, a block partition unit 102, a predicted signal generation unit 103, a frame memory 104, a subtraction unit 105, a transform unit 106, a quantization unit 107, an inverse quantization unit 108, an inverse transform unit 109, an addition unit 110, a quantized transform coefficient encoding unit 111, an output terminal 112, a prediction information memory 113, and a prediction information encoding unit 114. In an embodiment, the transform unit 106 and the quantization unit 107 function as residual signal encoding means, the inverse quantization unit 108 and the inverse transform unit 109 function as residual signal restoration means, and the prediction information memory functions as motion information recording means. It is noted that the prediction information memory 113 may be incorporated in the predicted signal generation unit 103. In an embodiment, the quantized transform coefficient encoding unit 111 may also function as a part of the residual signal encoding means and, in this case, an element that decodes data generated by the quantized transform coefficient encoding unit 111 and outputs the decoded data to the inverse quantization unit 108 constitutes a part of the residual signal restoration means.

[0040] The input terminal 101 is a terminal that implements input of a signal of a video sequence consisting of a plurality of pictures.

[0041] The block partition unit 102 partitions each picture of an encoding target represented by a signal input from the input terminal 101, into a plurality of regions. In an embodiment, an encoding target picture may be partitioned into a plurality of blocks each consisting of 8×8 pixels. These regions or blocks are target regions or target blocks

to which the processing described below is applied. The size and shape of blocks are not limited to this example. Blocks of different sizes may be mixed in each picture.

[0042] The predicted signal generation unit 103 detects motion information necessary for generation of a predicted signal of each prediction block in a target block and generates the predicted signal. There are no particular restrictions on a method of generation of the predicted signal, but it may be the inter-picture prediction or the intra-picture prediction as described in the background art (though the intra-picture prediction is not illustrated). In an embodiment, the motion information may be detected by the block matching shown in Fig. 16. In the block matching, the predicted signal to minimize the sum of absolute differences from the original signal (pixel signal) of the target block is detected from a reconstructed picture acquired via line L104.

[0043] In an embodiment, the motion information may contain a motion vector, an identification of a reference picture list (List0 and List1 in Fig. 5), and a reference index (ref_idx in Fig. 5).

[0044] When each reference picture list is composed of a plurality of reference pictures as shown in Fig. 6, the reference index is detected simultaneously with the motion vector, in order to indicate which reference picture is a target of the motion vector in the motion information. When bi-predictive prediction is used, a prediction type (forward/backward/two-way) is also determined at the same time. When the prediction type is the forward prediction, the motion information with the reference picture list identification of List0 is detected; when the prediction type is the backward prediction, the

motion information with the reference picture list identification of List1 is detected. In the case of the bi-predictive prediction, two pieces of motion information identified by List0 and List1 are detected. The reconstructed pictures registered in the reference picture lists shown in Fig. 6 may be automatically determined in accordance with a predetermined rule, or information to specify the reconstructed pictures registered in the reference picture lists may be explicitly encoded in frame units or in sequence units. In an embodiment, frame numbers may be used as the information, as shown in Figs. 5 and 6.

[0045] Referring back to Fig. 1, the motion information and the prediction type are output via line L103c and via line L103b to the prediction information memory 113 and to the prediction information encoding unit 114, respectively.

[0046] The prediction information memory 113 stores the input motion information and prediction type.

[0047] The prediction information encoding unit 114 entropy-encodes the motion information of the encoding target block and outputs encoded data via line L114 to the output terminal 112. There are no particular restrictions on a method of the entropy encoding, but it may be the arithmetic coding, variable-length coding, or the like.

[0048] The predicted signal generated by the predicted signal generation unit 103 is output via line L103a to the subtraction unit 105 and to the addition unit 110.

[0049] The subtraction unit 105 subtracts the predicted signal of the target block input via line L103a, from the pixel signal of the target block to generate a residual signal. The subtraction unit 105 outputs

the residual signal via line L105 to the transform unit 106.

[0050] The transform unit 106 performs a discrete cosine transform of the input residual signal to generate transform coefficients. The quantization unit 107 quantizes the transform coefficients generated by the transform unit 106, to generate quantized transform coefficients. The quantized transform coefficient encoding unit 111 entropy-encodes the quantized transform coefficients generated by the quantization unit 107. The encoded data generated by the quantized transform coefficient encoding unit 111 is output via line L111 to the output terminal 112. There are no particular restrictions on a method of the entropy encoding in the quantized transform coefficient encoding unit 111, but it may be arithmetic coding, variable-length coding, or the like.

[0051] The output terminal 112 outputs the information received from the prediction information encoding unit 114 and the information from the quantized transform coefficient encoding unit 111, together to the outside.

[0052] The inverse quantization unit 108 performs inverse quantization of the quantized transform coefficients generated by the quantization unit 107, to generate transform coefficients. The inverse transform unit 109 applies an inverse discrete cosine transform to the transform coefficients generated by the inverse quantization unit 108, to restore a residual signal. The addition unit 110 adds the restored residual signal to the predicted signal input provided via line L103a, to generate a reconstructed pixel signal of the encoding target block and stores the reconstructed pixel signal into the frame memory 104. It is also possible to adopt other transform processes replacing the processes of

the transform unit 106 and the inverse transform unit 109. It is also noted that the transform unit 106 and the inverse transform unit 109 are not always essential. In this manner, the reconstructed pixel signal of the encoding target block is stored as a part of a reference picture or reconstructed picture in the frame memory 104, in order to be used for generation of a predicted signal of a subsequent encoding target block.

[0053] In an embodiment, the predicted signal generation unit 103 can also generate the predicted signal of the target block by making use of the motion information associated with blocks neighboring the target block (neighboring blocks). The motion information associated with a neighboring block is motion information used in generation of the predicted signal when the neighboring block was an encoding target. The motion information associated with the neighboring blocks is stored in the prediction information memory 113. For each block, information to indicate that the predicted signal was generated by making use of the motion information associated with the neighboring blocks, or information to indicate that the predicted signal was generated by making use of the motion vector detected by the block matching may be included in an encoded state in compressed data.

[0054] Now, let us consider the case where the predicted signal of the target block is generated by the bi-predictive prediction. It will be described below using an example of (A) in Fig. 7. Neighboring blocks are two blocks 401 and 402 neighboring the target block 400, which are located immediately on the left side of and immediately above the target block 400, and the motion information associated with these blocks is used for the prediction. In this case, if the prediction

type associated with the two neighboring blocks is the bi-predictive prediction, up to four pieces of motion information can be used for generation of the predicted signal of the target block. The predicted signal generation unit 103 selects two pieces of motion information having different values of motion vectors or different values of frame numbers of reference pictures, from the four pieces of motion information as candidates (four pieces of candidate motion information). A frame number (frame_num) of a reference picture can be identified by a combination of a reference picture list identification (List0 or List1) and a reference index (ref_idex). Then two predicted signals are acquired from the frame memory 104 on the basis of the respective pieces of motion information, and the two predicted signals are averaged in pixel units to obtain a predicted signal. If the two predicted signals resemble each other, the prediction performance is improved by the effect of noise removal owing to smoothing.

[0055] The prediction information encoding unit 114 encodes information (indication information) to identify the two pieces of motion information used in the generation of the predicted signal of the target block out of the four pieces of candidate motion information, instead of the motion information. Therefore, the predicted signal may be generated in a bit count that is smaller than that of the encoded data obtained by encoding the motion information. The indication information may be two pieces of indication information to identify the respective pieces of candidate motion information, or may be a piece of indication information to identify the combination of the two pieces of candidate motion information.

[0056] The prediction information memory 113 stores the two pieces of motion information actually used in the generation of the predicted signal of the target block, as motion information associated with the target block. The two pieces of motion information may be stored as they are, but since the existing image predictive encoding and decoding processes have the restriction to refer to the two reference picture lists of List0 and List1, the two pieces of motion information can be stored as described below.

[0057] Namely, when the two pieces of motion information have the same reference picture list identification, the reference picture list identification in the two pieces of motion information can be modified in accordance with a predetermined rule. For example, of the two pieces of motion information, the reference picture list identification of one piece of motion information with a smaller reference index (ref_idx) is assigned to List0, and the reference picture list identification of the other piece of motion information is assigned to List1.

[0058] When the two pieces of motion information also have the same reference index, the reference picture list identification of one piece of motion information with the motion vector having the smaller sum of absolute values of respective components thereof is assigned to List0, and the reference picture list identification of the other piece of motion information is assigned to List1.

[0059] If this assignment results in referring to a frame number different from the frame number before the assignment, the reference index (ref_idx) can be modified so as to enable reference to the frame number before the assignment.

[0060] If the two reference picture lists do not include any identical frame number as shown in (B) of Fig. 6, the motion information may be managed using reference picture lists (523) extended so as to include the same frame numbers as shown in (C) of Fig. 6. The extended reference picture lists may be used directly in the case where the predicted signal of the target block is generated using the motion information associated with neighboring blocks. Since the frame numbers are recorded, they are also applicable to cases where the motion vector is scaled according to the distance in the temporal direction between pictures (e.g., each vector component is doubled when the motion vector to the target of the picture 504 in (A) of Fig. 5 is converted into the motion vector to the target of the picture 503). Since the motion vector storage methods described herein can be converted from one to another, any one of the storage methods can be applied with no influence on the encoding result.

[0061] The processing of the predicted signal generation unit 103 will be further described in detail. Fig. 2 is a drawing showing a configuration of the predicted signal generation unit 103 according to an embodiment. As shown in Fig. 2, the predicted signal generation unit 103 is provided with a motion information derivation unit 121, a motion prediction information selection unit 122, and a motion compensation unit 123. In this drawing, the part of motion detection by the conventional block matching is omitted.

[0062] When both of the prediction types of the immediately above and immediately left blocks used as neighboring blocks are bi-predictive prediction, the motion information derivation unit 121 compares four

pieces of motion information associated with these neighboring blocks. The motion information derivation unit 121 defines motion information having a different value of any one of elements constituting the motion information, which are the motion vector, the reference picture list identification, and the reference index, from those of the other three pieces of motion information, as candidate motion information applicable to the prediction of the target block. If there are two or more pieces of motion information with the same motion vector value and the same frame number value of the reference picture (the frame number of the reference picture can be identified by the combination of the reference picture list and the reference index), one of them is defined as candidate motion information. Then the motion information derivation unit 121 derives a set of any two pieces of candidate motion information, as candidate bi-predictive prediction information. In this case, up to six pieces of candidate bi-predictive prediction information can be derived. In cases where at least one of the immediately above and immediately left blocks to the target block is employed as a neighboring block and where the prediction is bi-predictive prediction, combination patterns of two pieces of motion information as many as the patterns described below can be derived from the neighboring blocks. Namely, the combination patterns are as follows.

(1) 2 patterns: in cases where two pieces of motion information of the immediately above block (402) or the immediately left block (401) are employed.

(2) 4 patterns: in cases where one piece of motion information of the immediately above block (402) and one piece of motion information

of the immediately left block (401) are employed and where the reference picture lists in the selected two pieces of motion information are different.

(3) 2 patterns: in cases where one piece of motion information of the immediately above block (402) and one piece of motion information of the immediately left block (401) are employed and where the reference picture lists in the selected two pieces of motion information are the same.

[0063] Fig. 3 shows a flowchart of the processing carried out by the motion information derivation unit 121. First, in step S301 the number N of neighboring blocks to be used in the prediction of the target block is set and the current neighboring block number n is set to "0." For example, n=0 is set for block 401 and n=1 for block 402. Furthermore, the number L of pieces of candidate motion information is initialized to 0. Then the motion information and prediction type of the n-th block are acquired from the prediction information memory 113 (step S302). Thereafter, the number M of pieces of motion information associated with the n-th neighboring block is set based on the prediction type (step S303).

[0064] Next, the m-th motion information of the n-th neighboring block is compared with the L pieces of candidate motion information (step S304). If all the pieces of candidate motion information are different from the motion information of the comparison target, or if L=0, the processing proceeds to step S305. In step S305, the m-th motion information of the n-th neighboring block is added to the candidate motion information and the value of L is given an increment of 1.

Thereafter, the processing proceeds to step S306.

[0065] If step S304 results in finding the same as the m-th motion information in the candidate motion information, the processing proceeds to step S306. In step S304, if values of two motion vectors in comparison are identical and if frame numbers corresponding to two sets of reference indices and reference picture list identifications in comparison are identical, the motion information is determined to be identical even with difference of the reference picture lists. A reason for it is that when the motion vectors of two pieces of motion information and the frame numbers of reference pictures thereof are identical, predicted signals generated from these pieces of motion information become coincident.

[0066] In step S306, the value of m is given an increment of 1. In subsequent step S307, it is determined whether the processes of steps S304 to S306 are completed, for all the pieces of motion information associated with the n-th neighboring block; if they are completed, the processing proceeds to step S308; if they are not completed, the processing returns to step S304.

[0067] In step S308, the value of the number n of the neighboring block is given an increment of 1. Then in step S309, it is determined whether the processing is completed for all the neighboring blocks. If it is completed, the processing proceeds to step S310; if it is not completed, the processing returns to step S302.

[0068] Finally, in step S310, every set of two pieces of motion information having different values of motion vectors or different values of frame numbers of reference pictures (the frame number of each

reference picture can be identified by the combination of the reference picture list and the reference index) is derived from the L pieces of candidate motion information and is set as candidate bi-predictive prediction information.

[0069] Concerning the motion information from which neighboring blocks the motion information derivation unit 121 uses, it may be preliminarily determined for both of the encoding side and the decoding side. Furthermore, information to specify the motion information from which neighboring blocks the motion information derivation unit 121 used may be transmitted in an encoded state to the decoding side.

[0070] The motion prediction information selection unit 122 selects an optimum piece from the candidate bi-predictive prediction information input via line L121 from the motion information derivation unit 121. Specifically, it selects the candidate bi-predictive prediction information to minimize the sum of absolute differences between the "signal obtained as an average of two signals acquired via line L104 from the frame memory 104 on the basis of two pieces of candidate motion information" and the "pixel signal of the target block acquired via line L102." Then indication information to identify the selected candidate bi-predictive prediction information or to identify the two pieces of candidate motion information constituting the selected candidate bi-predictive prediction information is output via line L103b to the prediction information encoding unit.

[0071] Furthermore, the motion prediction information selection unit 122 outputs the selected two pieces of motion information via line L103c to the prediction information memory 113 so as to make the

information stored therein, and outputs the information via line L122 to the motion compensation unit 123. The motion compensation unit 123 averages two reconstructed pixel signals extracted from the reconstructed pictures acquired via line L104, based on the two input pieces of motion information, to generate a predicted signal, and outputs the predicted signal via line L103a to the subtraction unit 105 and to the addition unit 110.

[0072] There are no particular restrictions on the technique of selecting one optimum piece of candidate bi-predictive prediction information in the motion prediction information selection unit 122. For example, as shown in Fig. 4, instead of a pixel signal of a target block 461 in an encoding target picture 451, a reconstructed region (462) of an inverted L-shape neighboring the block 461 may be used. Since the region 462 is the reconstructed region, the decoding device can also use it. As shown in the drawing, a predicted region (472) corresponding to the region 462 can be acquired from the reference picture 452 on the basis of one piece of motion information 473. For every piece of candidate bi-predictive prediction information, inverted L-shape regions can be acquired from two pieces of candidate motion information and signals thereof are averaged to find the candidate bi-predictive prediction information to minimize the sum of absolute differences from the region 462. This technique allows the decoding device to select the same candidate bi-predictive prediction information as that selected by the encoding device, and thus provides the advantage of eliminating the need for transmission of the indication information.

[0073] In this manner, the predicted signal with a small difference from

the original signal is generated from the motion information associated with the blocks neighboring the target block and two pieces of motion information having different values of motion vectors or different values of frame numbers of reference pictures (the frame number of each reference picture can be identified by the combination of the reference picture list and the reference index) are selected, whereby the bi-predictive prediction of the target block can be implemented without transmission of the motion information. Since the motion information contains the information to identify the reference picture list, it becomes feasible to perform the bi-predictive prediction using the motion information with the same reference picture list. Therefore, even if the reference pictures in two reference picture lists are completely different, the bi-predictive prediction can be carried out using motion vectors of the same reference picture. For example, when the bi-predictive prediction with two similar signals is carried out using the motion vectors of the same reference picture, improvement in prediction performance based on the smoothing effect may be expected.

[0074] Fig. 9 is a flowchart showing a procedure of an image predictive encoding method according to an embodiment. First, the block partition unit 102 partitions an input picture into 8×8 encoding blocks (the input picture may be partitioned in any other block size or shape or blocks of different sizes may be mixed in a picture).

[0075] Next, the motion information derivation unit 121 acquires the motion information associated with blocks neighboring a target block of an encoding target, from the prediction information memory 113 and extracts motion information having a different value of any one of the

constituent elements thereof from the other motion information, as candidate motion information applicable to the prediction of the target block. Then a set of any two pieces of candidate motion information is derived as candidate bi-predictive prediction information (step S300). The details of this step S300 were already described with reference to Fig. 3.

[0076] Next, the motion prediction information selection unit 122 selects one to be used in generation of the predicted signal of the target block, from plural pieces of candidate bi-predictive prediction information. Then the motion prediction information selection unit 122 determines the indication information to identify the selected piece of candidate bi-predictive prediction information or to identify the two pieces of candidate motion information constituting the selected piece of candidate bi-predictive prediction information (step S320). In subsequent step S101, the selected two pieces of motion information are stored into the prediction information memory 113 and the prediction information encoding unit 114 encodes the indication information.

[0077] Next, in step S102 the motion compensation unit 123 acquires pixel signals from two reconstructed pictures acquired from the frame memory 104, based on the selected motion information, and averages these pixel signals to generate a predicted signal of the target block. In the prediction technique not using the motion information of neighboring blocks, the motion information is detected by the block matching described in Fig. 16 and the predicted signal of the target block is generated based on the motion information. The prediction information encoding unit 114 encodes the motion information and

information for switching of the two prediction techniques. The motion information is stored into the prediction information memory 113.

[0078] Next, the residual signal indicative of the difference between the pixel signal of the encoding target block and the predicted signal is transformed and encoded by the transform unit 106, the quantization unit 107, and the quantized transform coefficient encoding unit 111 (step S103). Encoded data of the indication information and quantized transform coefficients is output through the output terminal 112 (step S104).

[0079] After these processes or in parallel with these processes, the encoded residual signal is decoded by the inverse quantization unit 108 and the inverse transform unit 109, for predictive encoding of a subsequent encoding target block. Then the addition unit 110 adds the decoded residual signal to the predicted signal to reconstruct a signal of the encoding target block. The reconstructed signal is stored as a reference picture (reconstructed picture signal) into the frame memory 104 (step S105). If the processing is not completed for all the encoding target blocks, the processing returns to step S101 to carry out the processing for the next encoding target block. If the processing is completed for all the encoding target blocks, the processing is terminated (step S106).

[0080] Next, an image predictive decoding device according to an embodiment will be described. Fig. 10 is a block diagram showing the image predictive decoding device according to an embodiment. The image predictive decoding device 200 shown in Fig. 10 is provided with

an input terminal 201, a data analysis unit 202, an inverse quantization unit 203, an inverse transform unit 204, an addition unit 205, an output terminal 206, a quantized transform coefficient decoding unit 207, a prediction information decoding unit 208, a frame memory 104, a motion information determination unit 209, a motion information derivation unit 121, a prediction information memory 113, and a motion compensation unit 126. The inverse quantization unit 203 and the inverse transform unit 204 function as residual signal decoding means and the prediction information memory 113 functions as motion information recording means. The decoding means by the inverse quantization unit 203 and the inverse transform unit 204 may be any means other than these in other embodiments. Furthermore, the inverse transform unit 204 may be excluded in other embodiments.

[0081] The input terminal 201 implements input of compressed data compressively encoded by the aforementioned image predictive encoding method. This compressed data contains, for each of the blocks, encoded data obtained by entropy encoding of quantized transform coefficients generated by transform and quantization of the residual signal, encoded data of the prediction information, i.e., encoded data of motion information for generation of the predicted signal of the block or indication information to identify the motion information to be used in generation of the predicted signal of the decoding target block from the motion information associated with neighboring blocks, and encoded data of information to indicate which of the motion information and the indication information is encoded.

[0082] The data analysis unit 202 analyzes the compressed data input

into the input terminal 201, separates the compressed data into encoded data of quantized transform coefficients and encoded data of prediction information, for a decoding target block, and outputs these encoded data via line L202a and via line L202b to the quantized transform coefficient decoding unit 207 and to the prediction information decoding unit 208, respectively.

[0083] The prediction information decoding unit 208 entropy-decodes the encoded data of motion information or indication information. The decoded motion information is sent via line L208 to the motion information determination unit 209 and is then output without change via line L207a and line L207b to the motion compensation unit 126 and the prediction information memory 113. The prediction information memory 113 stores the motion information and the motion compensation unit 126 acquires a pixel signal from a reconstructed picture acquired from the frame memory 104, based on the motion information, and generates the predicted signal of each prediction block in the decoding target block, using the pixel signal. The generated predicted signal is output via line L126 to the addition unit 205.

[0084] On the other hand, when the prediction information decoding unit 208 entropy-decodes the encoded data of indication information, the motion information derivation unit 121 functions. The motion information derivation unit 121, as described in detail with reference to Fig. 3, acquires the motion information associated with blocks neighboring the decoding target block, from the prediction information memory 113, and extracts information having different values of any one of constituent elements thereof, as candidate motion information

applicable to the prediction of the decoding target block. Then the motion information derivation unit 121 derives a set of any two pieces of candidate motion information as candidate bi-predictive prediction information and outputs the candidate bi-predictive prediction information via line L121 to the motion information determination unit 209.

[0085] The motion information determination unit 209 selects one piece of candidate bi-predictive prediction information identified by the indication information input via line L208, from plural pieces of candidate bi-predictive prediction information input via line L121. The two pieces of motion information constituting the selected candidate bi-predictive prediction information are output via line L207a and via line L207b to the motion compensation unit 126 and to the prediction information memory 113. The prediction information memory 113 stores the motion information and the motion compensation unit 126 acquires pixel signals from reconstructed pictures acquired from the frame memory 104, based on the two pieces of motion information, and generates a predicted signal of each prediction block in the decoding target block, using the pixel signals. The generated predicted signal is output via L126 to the addition unit 205.

[0086] The quantized transform coefficient decoding unit 207 entropy-decodes the encoded data of quantized transform coefficients of the residual signal in the encoding target block, and outputs the restored quantized transform coefficients via line L207 to the inverse quantization unit 203.

[0087] The inverse quantization unit 203 performs inverse quantization of the quantized transform coefficients of the decoding target block input via line L207, to generate transform coefficients. The inverse transform unit 204 performs an inverse discrete cosine transform of the transform coefficients to restore a residual signal.

[0088] The addition unit 205 adds the predicted signal generated by the predicted signal generation unit 103, to the residual signal restored by the inverse quantization unit 203 and the inverse transform unit 204, to generate a decoded pixel signal of the decoding target block. The decoded pixel signal is output via line L205 to the output terminal 206 and the frame memory 104. The output terminal 206 outputs a reconstructed picture including the decoded pixel signal, to the outside (e.g., a display).

[0089] The frame memory 104 stores as a reference picture the reconstructed picture output from the addition unit 205, which is used as a reconstructed picture for reference for the next decoding process.

[0090] An image predictive decoding method according to an embodiment will be described below with reference to Fig. 11. First, compressed data is input through the input terminal 201 (step S201). Then the data analysis unit 202 performs the data analysis of the compressed data and extracts the encoded data of the motion information for generation of the predicted signal of the block or the indication information to identify the motion information used in the generation of the predicted signal of the decoding target block from the motion information associated with neighboring blocks, the encoded data of the information to indicate which of the motion information and

the indication information is encoded, and the encoded data of quantized transform coefficients. The encoded data of motion information or the encoded data of indication information is decoded by the prediction information decoding unit 208.

[0091] When the motion information is restored, the motion compensation unit 126 generates the predicted signal of the decoding target block, based on the restored motion information. The restored motion information is stored in the prediction information memory 113.

[0092] When the indication information is restored, the motion information derivation unit 121 acquires the motion information of neighboring blocks from the prediction information memory 113. The motion information derivation unit 121 may acquire the motion information from predetermined neighboring blocks. The motion information may be acquired from neighboring blocks, based on the information transmitted from the encoding side.

[0093] Next, the motion information derivation unit 121, as shown in detail in Fig. 3, extracts motion information between which any one of constituent elements differs in value out of the acquired pieces of motion information, as candidate motion information applicable to the prediction of the decoding target block. Then the motion information derivation unit 121 derives a set of any two pieces of candidate motion information as candidate bi-predictive prediction information (step S300).

[0094] Next, the motion information determination unit 209 selects one piece of bi-predictive prediction information from the derived pieces of candidate bi-predictive prediction information, using the restored

indication information, to restore two pieces of motion information (step S203).

[0095] Thereafter, the motion compensation unit 126 generates the predicted signal of the decoding target block, based on the restored two pieces of motion information, and stores the two pieces of motion information into the prediction information memory (S207).

[0096] Next, the quantized transform coefficients decoded by the quantized transform coefficient decoding unit 207 are inversely quantized by the inverse quantization unit 203 and the inverse transform unit 204 further applies the inverse transform to the output from the inverse quantization unit 203 to generate a decoded residual signal (S208).

[0097] Next, the generated predicted signal is added to the decoded residual signal to generate a decoded signal, and this decoded signal is stored into the frame memory 104, for decoding of the next decoding target block (step S209). Thereafter, when the determination in step S210 results in determining that there is the next compressed data, the processes of S202 to S209 are repeated (S210), thereby processing all data to the last.

[0098] Fig. 12 is a drawing showing a configuration of a program for causing a computer to execute the image predictive encoding method according to an embodiment. The image predictive encoding program P100 is provided with a block partition module P101, a predicted signal generation module P103, a storage module P104, a subtraction module P105, a transform module P106, a quantization module P107, an inverse quantization module P108, an inverse transform module P109, an

addition module P110, a quantized transform coefficient encoding module P111, a prediction information storage module P113, and a prediction information encoding module P114. The predicted signal generation module P103 includes a motion information derivation module P121, a motion prediction information selection module P122, and a motion compensation module P123.

[0099] The functions that the block partition module P101, predicted signal generation module P103, storage module P104, subtraction module P105, transform module P106, quantization module P107, inverse quantization module P108, inverse transform module P109, addition module P110, quantized transform coefficient encoding module P111, prediction information storage module P113, prediction information encoding module P114, motion information derivation module P121, motion prediction information selection module P122, and motion compensation module P123 cause a computer to execute are the same as the respective functions of the block partition unit 102, predicted signal generation unit 103, frame memory 104, subtraction unit 105, transform unit 106, quantization unit 107, inverse quantization unit 108, inverse transform unit 109, addition unit 110, quantized transform coefficient encoding unit 111, prediction information memory 113, prediction information encoding unit 114, motion information derivation unit 121, motion prediction information selection unit 122, and motion compensation unit 123.

[0100] Fig. 13 is a drawing showing a configuration of a program for causing a computer to execute the image predictive decoding method according to an embodiment. The image predictive decoding program

P200 is provided with a data analysis module P201, a prediction information decoding module P202, a motion information derivation module P121, a motion information determination module P203, a prediction information storage module P205, a motion compensation module P126, a quantized transform coefficient decoding module P206, an inverse quantization module P207, an inverse transform module P208, an addition module P209, and a storage module P104.

[0101] The functions that the data analysis module P201, prediction information decoding module P202, motion information derivation module P121, motion information determination module P203, prediction information storage module P205, motion compensation module P126, quantized transform coefficient decoding module P206, inverse quantization module P207, inverse transform module P208, addition module P209, and storage module P104 cause a computer to execute are the same as the respective functions of the data analysis unit 202, prediction information decoding unit 208, motion information derivation unit 121, motion information determination unit 209, prediction information memory 113, motion compensation unit 126, quantized transform coefficient decoding unit 207, inverse quantization unit 203, inverse transform unit 204, addition unit 205, and frame memory 104.

[0102] The image predictive encoding program P100 or the image predictive decoding program P200 configured as described above is stored in a recording medium 10 and executed by a computer described below.

[0103] Fig. 14 is a drawing showing a hardware configuration of a

computer for executing a program recorded in a recording medium and Fig. 15 is a perspective view of a computer for executing a program stored in a recording medium. Equipment that executes a program recorded in a recording medium is not limited to computers, but may be a DVD player, a set-top box, a cell phone, or the like provided with a CPU and configured to perform processing and control based on software.

[0104] As shown in Fig. 14, a computer 30 is provided with a reading device 12 such as a floppy disk drive unit, a CD-ROM drive unit, or a DVD drive unit, a working memory (RAM) 14 on which an operating system is resident, a memory 16 storing a program stored in a recording medium 10, a monitor device 18 such as a display, a mouse 20 and a keyboard 22 as input devices, a communication device 24 for transmission/reception of data and others, and a CPU 26 to control execution of the program. When the recording medium 10 is put into the reading device 12, the computer 30 becomes accessible to the image predictive encoding program P100 or the image predictive decoding program P200 stored in the recording medium 10, through the reading device 12, and becomes able to operate as the image predictive encoding device 100 or the image predictive decoding device 200, based on the image predictive encoding program P100 or the image predictive decoding program P200.

[0105] As shown in Fig. 15, the image predictive encoding program and the image predictive decoding program may be those provided in the form of computer data signal 40 superimposed on a carrier wave, through a network. In this case, the computer 30 can save the image

predictive encoding program or the image predictive decoding program received by the communication device 24, into the memory 16 and execute the image predictive encoding program or the image predictive decoding program.

[0106] The various embodiments were described above, but it should be noted that the present invention can further include modifications as described below.

[0107] (Candidates for Motion Information)

In the above embodiment the motion information is composed of the motion vector, the reference picture list identification (List0 and List1 in Fig. 5), and the reference index (ref_idx in Fig. 5), but the motion information may be configured by replacing the reference picture list identification and the reference index with a frame number. A reason for it is that the same result and effect as in the above-described embodiment using the motion information associated with neighboring blocks can be achieved by use of the frame numbers.

[0108] (Restrictions on Candidate Two-way Motion Information)

In the above embodiment the motion information derivation unit used all the combinations of candidate motion information as candidate two-way motion information, but only sets of motion information satisfying the conditions below may be adopted as candidate two-way motion information.

1) Only the immediately above and immediately left blocks to the target block.

2) Two pieces of motion information to indicate the same reference picture.

3) Two motion vectors which are not identical but have close values (the difference of absolute values of which is smaller than a certain value). Namely, the motion vector (MVXa,MVYa) of the immediately above block and the motion vector (MVXb,MVYb) of the immediately left block satisfy the following two relations:

(|MVXa − MVXb| != 0)||(|MVYa − MVYb| != 0); and

(|MVXa − MVXb| <= N) && (|MVYa − MVYb| <= N) (where N is a small value).

If there is no candidate two-way motion information satisfying the conditions, the other prediction process may be used without transmission of the indication information.

[0109] The number of candidate two-way motion information may be designated in frame units or in block units and information to specify the designated number may be included in compressed data. Furthermore, information to indicate such application of the restrictions on the candidate two-way motion information may be encoded in frame units.

[0110] (Identification of Indication Information)

In determining the identification of the indication information, a priority order of the immediately left block (401) and the immediately above block (402) may be determined using the motion information of the above right block (404 in Fig. 8) to the target block (400 in Fig. 8). If the motion vector of the immediately above block (MVXa,MVYa), the motion vector of the immediately left block (MVXb,MVYb), and the motion vector of the above left block (MVXc,MVYc) satisfy the relation of |MVXa − MVXc| + |MVYa − MVYc| < |MVXb − MVXc| +

|MVYb − MVYc|, the motion information of the immediately left block is given an identification with a smaller bit count. If the relation is not met, the motion information of the immediately above block is given an identification with a smaller bit count. When a higher priority is put on the motion information of the neighboring block assumed to have a higher correlation with the target block in this manner, the bit count of the indication information can be reduced.

[0111] (Derivation and Selection of Candidate Motion Vector)

In the above embodiment the derivation and selection of candidate motion vector are carried out for the horizontal and vertical components of the motion vector together, but the derivation and selection of candidate motion vector may be individually performed for the horizontal component and the vertical component of the motion vector. In the above embodiment a single motion vector of a neighboring block was defined as a candidate motion vector, but it is also possible to define as a candidate motion vector an average of two motion vectors or a motion vector scaled according to the distance in the temporal direction between pictures (e.g., each vector component is doubled when the motion vector to the target of the picture 504 in (A) of Fig. 5 is converted into the motion vector to the target of the picture 503).

[0112] (Prediction Type)

In the above embodiment the prediction type for both of two neighboring blocks is the bi-predictive prediction, but the present invention can also be applied to cases where one of them is forward or backward prediction. In such cases, up to three pieces of candidate

bi-predictive prediction information are derived. In these cases, combination patterns of two pieces of motion information as many as the patterns described below can be derived from neighboring blocks. Namely, the combination patterns are as follows:

(1) 1 pattern: in cases where two pieces of motion information of the immediately above block (402) or the immediately left block (401) are adopted;

(2) 1 pattern: in cases where one piece of motion information of the immediately above block (402) and one piece of motion information of the immediately left block (401) are employed and where the reference picture lists about the two pieces of motion information are different;

(3) 1 pattern: in cases where one piece of motion information of the immediately above block (402) and one piece of motion information of the immediately left block (401) are employed and where the reference picture lists about the two pieces of motion information are identical.

[0113] Furthermore, both of the two neighboring blocks may be the forward or backward prediction, or one of them may be the forward prediction and the other the backward prediction.

[0114] (Number of Neighboring Blocks)

There are no particular restrictions on the number of neighboring blocks applicable to the prediction of the target block. As in (B) of Fig. 7, three blocks 401 to 403 may be used as neighboring blocks. It is also possible to use motion information associated with a block of a preceding frame. If the neighboring blocks (401-405) are

further partitioned as shown in Fig. 8, the motion information associated with such smaller blocks may be included in the candidate motion information. It is also possible to use motion information of a block which is not adjacent to the target block.

[0115] (N-predictive Prediction)

In the above embodiment the prediction method of the target block performed using the motion information associated with the neighboring blocks is the bi-predictive prediction, but the prediction method does not have to be limited to it, and may be uni-predictive prediction or tri-predictive prediction. In this case, step S310 in Fig. 3 may be configured to derive candidate N-predictive prediction information and select one therefrom. N pieces of motion information may be stored in that case, but in the case where N is three or more, the motion information to be stored may be limited to two, by the reference indices or the values of the motion vectors.

[0116] (Transform Unit and Inverse Transform Unit)

The transform process of the residual signal may be carried out in a fixed block size or the transform process may be carried out by further partitioning the target region according to a partial region.

[0117] (Color Signal)

In the description of the above embodiment there is no particular description about color format, but the predicted signal generation process may be carried out individually from a luminance signal, for a color signal or a color difference signal. The processing may also be carried out in synchronism with the processing of the luminance signal.

[0118] The present invention was described above in detail on the basis

of some embodiments thereof. However, it should be noted that the present invention is by no means limited to the above embodiments. The present invention can be modified in many ways without departing from the scope and spirit of the invention.

Reference Signs List

[0119] 100 image predictive encoding device; 101 input terminal; 102 block partition unit; 103 predicted signal generation unit; 104 frame memory; 105 subtraction unit; 106 transform unit; 107 quantization unit; 108 inverse quantization unit; 109 inverse transform unit; 110 addition unit; 111 quantized transform coefficient encoding unit; 112 output terminal; 113 prediction information memory; 114 prediction information encoding unit; 121 motion information derivation unit; 122 motion prediction information selection unit; 123 motion compensation unit; 201 input terminal; 202 data analysis unit; 203 inverse quantization unit; 204 inverse transform unit; 205 addition unit; 206 output terminal; 207 quantized transform coefficient decoding unit; 208 prediction information decoding unit; 209 motion information determination unit.

CLAIMS

1. An image predictive encoding device comprising:

region partition means adapted to partition an input picture into a plurality of regions;

predicted signal generation means adapted to determine motion information for acquisition of a pixel signal having a high correlation with a pixel signal of a target region of an encoding target, the target region being from among the plurality of regions partitioned by the region partition means, from a reconstructed picture, and the predicted signal generation means adapted to generate a predicted signal of the target region from the reconstructed picture based on the motion information;

motion information recording means adapted to store the motion information;

residual signal generation means adapted to generate a residual signal, the residual signal being a difference between the predicted signal of the target region and the pixel signal of the target region;

residual signal encoding means adapted to encode the residual signal generated by the residual signal generation means;

residual signal reconstruction means adapted to encoded data generated by the residual signal encoding means, to generate a reconstructed residual signal;

addition means adapted to add the predicted signal to the reconstructed residual signal to generate a reconstructed pixel signal of the target region; and

picture recording means adapted to store the reconstructed pixel signal generated by the addition means, as a part of a reconstructed picture,

wherein the predicted signal generation means comprises:

motion information derivation means adapted to derive one or more motion information sets from plural pieces of motion information stored in the motion information recording means, each of the one or more motion information sets consisting of two pieces of motion information between which any one of constituent elements differ in value;

motion prediction information selection means adapted to select, from the one or more motion information sets, a motion information set to be used in generation of the predicted signal of the target region, the motion prediction information selection means further adapted to direct storage of two pieces of motion information included in the selected set in the motion information recording means; and

motion compensation means adapted to combine two signals acquired from the reconstructed picture based on the two pieces of motion information in the selected set, to generate the predicted signal of the target region.

2. The image predictive encoding device according to claim 1,

wherein the picture recording means is adapted to store a plurality of reconstructed pictures,

wherein each piece of the motion information contains information to identify the reconstructed picture, and a motion vector,

wherein the reconstructed pictures stored in the picture recording means are identified using respective frame numbers assigned during encoding of the reconstructed pictures, and

wherein the motion information derivation means is adapted to derive a motion information set including two pieces of motion information which contain motion vectors whose frame number of the reconstructed picture is identical, but whose values are different.

3. The image predictive encoding device according to claim 1,

wherein the picture recording means is adapted to store a plurality of reconstructed pictures,

wherein the plurality of reconstructed pictures stored in the picture recording means are identified using two reference picture lists each consisting of one or more reconstructed pictures,

wherein each piece of the motion information contains information to identify the two reference picture lists, and a motion vector, and

wherein the motion information derivation means is adapted to derive motion information sets to include two pieces of motion information having identical information to identify the reference picture lists, and different values of motion vectors.

4. The image predictive encoding device according to any one of claims 1 to 3, wherein the motion information derivation means is adapted to derive the motion information set to be used in generation of the predicted signal of the target region from plural pieces of motion information associated with a region neighboring the target region.

5. The image predictive encoding device according to any one of claims 1 to 4, further comprising prediction information encoding means adapted to encode information to indicate the motion information set consists of the two pieces of motion

information selected by the predicted signal generation means.

6. An image predictive decoding device comprising:

data analysis means adapted to extract encoded data of motion information, the encoded data of motion information extracted out of a plurality of regions in a picture for generation of a predicted signal of a target region of a decoding target, and the data analysis means further adapted to extract encoded data of a residual signal from compressed data;

motion information recording means adapted to store the motion information;

motion compensation means adapted to generate the predicted signal of the target region based on the motion information;

residual signal decoding means adapted to decode a residual signal of the target region from the encoded data of the residual signal;

picture recording means adapted to add the predicted signal to the decoded residual signal to restore a pixel signal of the target region and to store the restored pixel signal as a part of a reconstructed picture;

motion information derivation means adapted to derive, from plural pieces of motion information stored in the motion information recording means, one or more motion information sets each consisting of two pieces of motion information between which any one of constituent elements differ in value; and

motion information determination means adapted to determine, from among the one or more derived motion information sets, a motion information set to be used in generation of the predicted signal of the target region and the motion information determination means further adapted to direct storage of two pieces of motion information included in the determined set in the motion information recording means,

wherein the motion compensation means is adapted to combine two signals acquired from the reconstructed picture to generate the predicted signal of the target region, the two signals combined based on the two pieces of motion information in the determined set.

7. The image predictive decoding device according to claim 6,

wherein the data analysis means is adapted to extract, from plural pieces of motion information stored in the motion information recording means, encoded data comprising indication information that identifies at least one motion information set consisting of two pieces of motion information between which any one of constituent elements differ in value,

the image predictive decoding device further comprising prediction information decoding means adapted to decode the indication information,

wherein the motion information determination means is adapted to extract, based on the decoded indication information, a motion information set to be used in generation of the predicted signal of the target region, the motion information set extracted out of the one or more motion information sets derived by the motion information derivation means.

8. The image predictive decoding device according to claim 6 or 7,

wherein the picture recording means is adapted to store a plurality of reconstructed pictures,

wherein the motion information contains information to identify the plurality of reconstructed pictures, and motion vectors,

wherein the plurality of reconstructed pictures stored in the picture recording means are identified using respective frame numbers during decoding of the respective reconstructed pictures, and

wherein the motion information derivation means is adapted to derive a motion information set including two pieces of motion information which contain motion vectors whose frame number of the reconstructed picture is identical but whose values are different.

9. The image predictive decoding device according to claim 6 or 7,

wherein the picture recording means is adapted to store a plurality of reconstructed pictures,

wherein the plurality of reconstructed pictures stored in the picture recording means are identified using two reference picture lists, each consisting of one or more reconstructed pictures,

wherein each piece of the motion information is configured to contain information identifying the two reference picture lists, and a motion vector, and

wherein the motion information derivation means is adapted to derive a motion information set including two pieces of motion information having identical information to identify the reference picture lists, and different values of motion vectors.

10. The image predictive decoding device according to any one of claims 6 to 9, wherein the motion information derivation means is adapted to derive the motion information set to be used in generation of the predicted signal of the target region, from

plural pieces of motion information associated with a region neighboring the target region.

11. An image predictive encoding method comprising:

partitioning an input picture into a plurality of regions;

generating a predicted signal of a target region of an encoding target out of the plurality of regions, said generating including determining motion information for acquisition of a pixel signal from a reconstructed picture having a high correlation with a pixel signal of the target region, and generating the predicted signal from the reconstructed picture based on the motion information;

storing the motion information by motion information recording means;

generating a residual signal, the residual signal being a difference between the predicted signal of the target region and the pixel signal of the target region;

encoding the residual signal;

decoding encoded data generated by said encoding of the residual signal, to generate a reconstructed residual signal;

adding the predicted signal to the reconstructed residual signal to generate a reconstructed pixel signal of the target region; and

storing, with picture recording means, the reconstructed pixel signal as a part of a reconstructed picture,

wherein said generating the predicted signal comprises:

deriving one or more motion information sets out of plural pieces of motion information stored in the motion information recording means, each of the one or more sets consisting of two pieces of motion information between which any one of constituent elements differ in value;

selecting, from the one or more motion information sets, a motion information set to be used in generation of the predicted signal of the target region, and storing two pieces of motion information in the selected set in the motion information recording means; and

combining two signals acquired from the reconstructed picture based on the two pieces of motion information in the selected set, the two signals combined to generate the predicted signal of the target region.

12. An image predictive encoding program for causing a computer to function as:

region partition means adapted to partition an input picture into a plurality of

regions;

predicted signal generation means adapted to determine motion information for acquisition of a pixel signal from a reconstructed picture, the pixel signal from the reconstructed picture having a high correlation with a pixel signal of a target region of an encoding target included in the plurality of regions partitioned by the region partition means, and the predicted signal generation means further adapted to generate a predicted signal of the target region from the reconstructed picture based on the motion information;

motion information recording means adapted to store the motion information;

residual signal generation means adapted to generate a residual signal, the residual signal being a difference between the predicted signal of the target region and the pixel signal of the target region;

residual signal encoding means adapted to encode the residual signal generated by the residual signal generation means;

residual signal reconstruction means adapted to decode encoded data generated by the residual signal encoding means to generate a reconstructed residual signal;

addition means adapted to add the predicted signal to the reconstructed residual signal to generate a reconstructed pixel signal of the target region; and

picture recording means adapted to store the reconstructed pixel signal generated by the addition means, as a part of a reconstructed picture,

wherein the predicted signal generation means comprises:

motion information derivation means adapted to derive one or more motion information sets out of plural pieces of motion information stored in the motion information recording means, each of the motion information sets consisting of two pieces of motion information between which any one of constituent elements differs in value;

motion prediction information selection means adapted to select a motion information set to be used in generation of the predicted signal of the target region, from the one or more motion information sets, and the motion prediction information selection means adapted to direct storage of two pieces of motion information in the selected set in the motion information recording means; and

motion compensation means adapted to combine two signals acquired from the reconstructed picture based on the two pieces of motion information in the selected set, the two signals combined to generate the predicted signal of the target region.

13. An image predictive decoding method comprising:

extracting encoded data of motion information, the encoded data of motion information extracted out of a plurality of regions in a picture for generation of a predicted signal of a target region of a decoding target, and extracting encoded data of a residual signal from compressed data;

storing the motion information with motion information recording means;

generating the predicted signal of the target region based on the motion information;

restoring a decoded residual signal of the target region from encoded data of the residual signal;

adding the predicted signal to the decoded residual signal to decode a pixel signal of the target region, and storing, with picture recording means, the decoded pixel signal as a part of a reconstructed picture;

deriving, out of plural pieces of motion information stored in the motion information recording means, one or more motion information sets each consisting of two pieces of motion information between which any one of constituent elements differ in value; and

determining, from among the derived motion information sets, a motion information set to be used in generation of the predicted signal of the target region, and directing storage of two pieces of motion information in the determined set in the motion information recording means,

wherein said generating the predicted signal comprises combining two signals acquired from the reconstructed picture based on the two pieces of motion information in the determined set, to generate the predicted signal of the target region.

14. An image predictive decoding program for causing a computer to function as:

data analysis means adapted to extract encoded data of motion information, the encoded data of motion information extracted out of a plurality of regions in a picture for generation of a predicted signal of a target region of a decoding target, and the data analysis means further adapted to extract encoded data of a residual signal from compressed data;

motion information recording means adapted to store the motion information;

motion compensation means adapted to generate the predicted signal of the target region based on the motion information;

residual signal restoration means adapted to decode a residual signal of the

target region from the encoded data of the residual signal;

picture recording means adapted to add the predicted signal to the decoded residual signal to decode a pixel signal of the target region, and the picture recording means further adapted to store the decoded pixel signal as a part of a reconstructed picture;

motion information derivation means adapted to derive, from plural pieces of motion information stored in the motion information recording means, one or more motion information sets each consisting of two pieces of motion information between which any one of constituent elements differ in value; and

motion information determination means adapted to determine, out of the one or more derived motion information sets, a motion information set to be used in generation of the predicted signal of the target region, and the motion information determination means further adapted to direct storage of two pieces of motion information in the determined set in the motion information recording means,

wherein the motion compensation means is adapted to combine two signals acquired from the reconstructed picture based on the two pieces of motion information in the determined set, the two signals combined to generate the predicted signal of the target region.

15. An image predictive decoding device comprising:

data analysis means adapted to extract, from compressed data, encoded data of motion information or encoded data of indication information, and extracting encoded data of a residual signal from the compressed data, the encoded data of motion information extracted out of a plurality of regions in a picture for generation of a predicted signal of a target region of a decoding target, the encoded data of indication information identifying a motion information set utilized for generation of a predicted signal of the target region out of a plurality of pieces of motion information associated with neighboring regions neighboring the target region;

prediction information decoding means adapted to decode (a) the encode data of motion information to restore two pieces of motion information which comprise a reference picture list identification, a reference picture index and a motion vector, or (b) the encoded data of indication information to restore the indication information;

motion information derivation means adapted to derive, one or more motion information sets out of plural pieces of motion information stored in memory, in response to restoration of the indication information, each of the one or more motion information sets including two pieces of motion information between which respective

frame numbers identified with respective reference picture list identification and respective reference picture index are different from each other, or between which respective motion vectors are different from each other;

motion information determination means adapted to determine, from the one or more motion information sets, a motion information set to be used in generation of the predicted signal of the target region based on the indication information;

motion compensation means adapted to generate the predicted signal of the target region by combining two signals obtained from one or more reconstructed pictures stored in the memory based on the two pieces of motion information which are the determined motion information set or the restored two pieces of motion information;

motion information recording means adapted to store motion information utilized for generation of the predicted signal of the target region in the memory;

residual signal restoration means adapted to restore a residual signal of the target region from the encoded data of the residual signal;

picture recording means adapted to reconstruct a pixel signal of the target region based on the predicted signal and the restored residual signal, and to store the reconstructed pixel signal as a part of a reconstructed picture.

16. An image predictive decoding method comprising:

extracting from compressed data, encoded data of motion information or encoded data of indication information, and extracting encoded data of a residual signal from the compressed data, the encoded data of motion information extracted out of a plurality of regions in a picture for generation of a predicted signal of a target region of a decoding target, the encoded data of indication information identifying a motion information set utilized for generation of a predicted signal of the target region out of a plurality of pieces of motion information associated with neighboring regions neighboring the target region;

decoding either (a) the encode data of motion information to restore two pieces of motion information which comprise a reference picture list identification, a reference picture index and a motion vector, or (b) the encoded data of indication information to restore the indication information;

deriving one or more motion information sets out of plural pieces of motion information stored in memory, in response to restoration of the indication information, each of the one or more motion information sets including two pieces of motion information between which respective frame numbers identified with respective reference picture list identification and respective reference picture index are different

from each other, or between which respective motion vectors are different from each other;

determining from the one or more motion information sets, a motion information set to be used in generation of the predicted signal of the target region based on the indication information;

generating the predicted signal of the target region by combining two signals obtained from one or more reconstructed pictures stored in the memory based on the two pieces of motion information which are the determined motion information set or the restored two pieces of motion information;

storing motion information utilized for generation of the predicted signal of the target region in the memory;

restoring a residual signal of the target region from the encoded data of the residual signal;

reconstructing a pixel signal of the target region based on the predicted signal and the restored residual signal; and

storing the reconstructed pixel signal as a part of a reconstructed picture.

17. An image predictive decoding program for causing a computer to function as:

extracting from compressed data, encoded data of motion information or encoded data of indication information, and extracting encoded data of a residual signal from the compressed data, the encoded data of motion information extracted out of a plurality of regions in a picture for generation of a predicted signal of a target region of a decoding target, the encoded data of indication information identifying a motion information set utilized for generation of a predicted signal of the target region out of a plurality of pieces of motion information associated with neighboring regions neighboring the target region;

decoding either (a) the encode data of motion information to restore two pieces of motion information which comprise a reference picture list identification, a reference picture index and a motion vector, or (b) the encoded data of indication information to restore the indication information;

deriving one or more motion information sets out of plural pieces of motion information stored in memory, in response to restoration of the indication information, each of the one or more motion information sets including two pieces of motion information between which respective frame numbers identified with respective reference picture list identification and respective reference picture index are different from each other, or between which respective motion vectors are different from each

other;

determining from the one or more motion information sets, a motion information set to be used in generation of the predicted signal of the target region based on the indication information;

generating the predicted signal of the target region by combining two signals obtained from one or more reconstructed pictures stored in the memory based on the two pieces of motion information which are the determined motion information set or the restored two pieces of motion information;

storing motion information utilized for generation of the predicted signal of the target region in the memory;

restoring a residual signal of the target region from the encoded data of the residual signal;

reconstructing a pixel signal of the target region based on the predicted signal and the restored residual signal; and

storing the reconstructed pixel signal as a part of a reconstructed picture.